ISBA World Meeting

July 1-7, 2024
San Giobbe
Economics Campus
Ca' Foscari University
of Venice

Thursday 4th, 2024

Poster Session 2

Esmail Abdul Fattah	Scalable Bayesian Inference using Integrated Nested Laplace Approximations

Integrated Nested Laplace Approximations (INLA) stand as a pivotal tool for performing approximate Bayesian inference, with broad applications in statistical modeling and various data-science fields. However, the current INLA implementation often hits computational and scalability ceilings, particularly when dealing with high-dimensional spatial-temporal models and intricate matrices. To overcome these limitations, this talk introduces a new versatile framework, designed specifically to leverage the scalability and processing power of manycore systems. This framework employs new numerical solver methods that take advantage of the dense block structure found in several data-science applications, thereby harmoniously blending the benefits of both dense and sparse computational techniques. By making optimal use of manycore architectures, the new framework effectively bypasses previous computational barriers, allowing for Bayesian inference on models with an extensive scope of latent parameters. We validate the efficacy of the new framework through comprehensive simulation tests and its application to real-world datasets. Our results confirm that the new framework enhances computational performance while maintaining the rigor and accuracy of the inference, thereby pushing the envelope of what is feasible with the existing INLA implementation.

Luca Aiello	Detecting Spatial Health Disparities Using Disease Maps

Epidemiologists commonly use regional aggregates of health outcomes to map mortality or incidence rates and identify geographic disparities. However, to detect health disparities across regions, it is necessary to identify ``difference boundaries'' that separate neighboring regions with significantly different spatial effects. This can be particularly challenging when dealing with multiple outcomes for each unit and accounting for dependence among diseases and across areal units. In this study, we address the issue of multivariate difference boundary detection for correlated diseases by formulating the problem in terms of Bayesian pairwise multiple comparisons by extending it through the introduction of adjacency modeling and disease graph dependencies. Specifically, we seek the posterior probabilities of neighboring spatial effects being different. To accomplish this, we adopt a class of multivariate areally referenced Dirichlet process models that accommodate spatial and interdisease dependence by endowing the spatial random effects with a discrete

probability law. Our method is evaluated through simulation studies and applied to detect difference boundaries for multiple cancers using data from the Surveillance, Epidemiology, and End Results Program of the National Cancer Institute.

Louise Alamichel	Bayesian mixture models (in)consistency for the number of clusters

Bayesian nonparametric (BNP) mixture models are commonly utilised for modeling intricate data structures. The posterior density of these models demonstrates robust theoretical properties, converging at the optimal minimax rate to the true data-generating density. Extensive research has focused on developing this theory. However, the consistency of the posterior density does not guarantee the consistency of the inferred number of clusters. Recent studies have revealed that both the Dirichlet process and Pitman-Yor process mixture models can exhibit inconsistency in estimating the number of clusters. We extend these results to additional Bayesian nonparametric priors such as Gibbs-type processes and finite-dimensional representations thereof. In the case of Dirichlet process mixture models, this problem can be avoided when a prior is put on the model's concentration hyperparameter α, a practice commonly adopted. We explore whether this result extends to more general BNP mixture models. We prove that for the Pitman-Yor mixture models with a prior on α and a fixed discount parameter $\sigma \in (0, 1)$, these models remain inconsistent in estimating the number of clusters. Finally, we investigate the utilisation of summarization techniques to obtain a point estimator for the partition distribution. We study the asymptotic properties of these estimators, focusing in particular on the maximum a posteriori estimator.

Angelos Alexopoulos	Variance reduction for Metropolis-Hastings samplers

We introduce a general framework that constructs estimators with reduced variance for random walk Metropolis and Metropolis-adjusted Langevin algorithms. The resulting estimators require negligible computational cost and are derived in a post-process manner utilising all proposal values of the Metropolis algorithms. Variance reduction is achieved by producing control variates through the approximate solution of the Poisson equation associated with the target density of the Markov chain. The proposed method is based on approximating the target density with a Gaussian and then utilising accurate solutions of the Poisson equation for the Gaussian case. This leads to an estimator that uses two key elements: (1) a control variate from the Poisson equation that contains an intractable expectation

under the proposal distribution, (2) a second control variate to reduce the variance of a Monte Carlo estimate of this latter intractable expectation. Simulated data examples are used to illustrate the impressive variance reduction achieved in the Gaussian target case and the corresponding effect when target Gaussianity assumption is violated. Real data examples on Bayesian logistic regression and stochastic volatility models verify that considerable variance reduction is achieved with negligible extra computational cost.

Julyan Arbel	Rapture of the deep: highs and lows of Bayes in a world of depths

Bayesian deep learning is appealing as it combines the coherence and natural uncertainty quantification of the Bayesian paradigm together with the expressivity and compositional flexibility of deep neural networks. Besides, it has the potential to provide learning mechanisms endowed with certain interpretability guarantees. In this talk, I will make an overview of distributional properties of Bayesian neural networks. This journey will start with early work by Radford Neal in the 90s. This led to the so-called Gaussian hypothesis of the pre-activations, which can be justified when the number of neurons per layer tends to infinity. I will then contrast this hypothesis with recent work on heavy-tailed pre-activations. Finally, I will describe a set of constraints that a neural network should fulfil to ensure Gaussian pre-activations.

Somnath Bhadra	Variational Bayes and Truncation approximations for Enriched Dirichlet process mixtures

A common impediment in conducting inference for Bayesian nonparametric models is either the need for complex MCMC algorithms and/or computational run-time for large datasets. We propose solutions here for Enriched Dirichlet process mixtures (EDPM). We derive a variational Bayes estimator based on a previously developed truncation approximation for EDPMs. The variational Bayes estimator can be used in two ways: 1) to develop a more efficient truncation approximation; 2) as good inital values for a blocked gibbs sampler based on this more efficient truncation approximation or for a polya urn sampler. We derive the accuracy of this more efficient truncation approximation and demonstrate how this allows for simple implementation of EDPMs in standard software for Hamiltonian MC. We confirm the validity of the approximations by simulations.

Srijato Bhattacharyya	A Bayesian Spatially Contiguous Partitioning Method with User-Defined Soft Constraints

This study develops a novel Bayesian random partition distribution as a prior model for latent cluster memberships, taking into account contiguity constraint with respect to a neighborhood graph such that elements in each cluster are adjacent to each other in space. The prior model is built upon random spanning tree cuts, which is strategically designed to also allow for the soft enforcement of additional user-prespecified constraints, such as shrinkage towards a specified clustering or other covariate-constrained conditions. We define suitable graph-agnostic as well as graph-aware measures in order to quantify the extent of enforcement of the constraint in our prior. Moreover, we introduce hyperparameters in our prior model to provide users with flexibility to adjust the extent of enforcement of the prespecified clustering constraint. The integration of our proposed prior into a suitably designed Bayesian hierarchical model results in spatially contiguous clustering with a strong tendency to respect the pre-defined clustering constraint. We introduce an efficient and fast Markov chain Monte Carlo (MCMC) algorithm utilizing spanning tree algorithms for model inference. Finally, we illustrate the performance of our method using simulated and real data examples for spatial clustering.

Ranran Bian	Adaptive Representation Learning for Large-Scale Dynamic Heterogeneous Networks

Representation learning generates the embedding vector of an object based on its relationships with others in a network. The generated vectors are inputs to various downstream machine learning tasks, such as classification, clustering and similarity search. The research area has attracted great interest and effort in recent years. However, due to its complexity, few of the existing representation learning methods have been developed for dynamic heterogeneous networks. Comparing with a static homogeneous network (graph), which contains single-typed objects (nodes) and relationships (edges) and remains unchanged over time, a dynamic heterogeneous network contains multiple-typed objects and relationships and evolves with time. We develop a novel adaptive representation learning algorithm, named Ada2vec, to address the challenges of embedding learning in large-scale dynamic heterogeneous networks. The key challenge is how to efficiently and effectively handle the network dynamics and heterogeneity features simultaneously. Ada2vec employs a statistical bound to capture network dynamics, and metapath-guided random walks to capture

network heterogeneity. To the best of our knowledge, Ada2vec is the first approach that leverages the Hoeffding bound for modeling changes in dynamic heterogeneous network embedding. Extensive experiments demonstrate that Ada2vec significantly outperforms the state-of-the-art benchmarks in terms of embedding learning accuracy and efficiency. Compared to other dynamic heterogeneous network representation learning models, the size of the experiment datasets that we use are orders of magnitude larger (millions instead of tens of thousands).

Daniel Bonnéry	A tempered Sequential Monte Carlo Method for estimating the number of variants in metagenomic samples

Antibiotic-resistance genes in various bacteria are the focus of intense attention given the global concern around the rise of multi-resistant pathogens. These genes come in several variants, and the precise identification of variants in samples is very useful to track antibiotic-resistance gene propagation. We identify the variants of antibiotic resistance genes and their relative abundances using shotgun metagenomic data collected on multiple samples. We consider a probabilistic approach, as it is essential to model DNA sequencing errors to distinguish actual variants from noise. We adapt ideas from the DESMAN software by Quince et al. (2017). The model is essentially a hierarchy of finite mixture submodels, which allows sharing of components (sharing of variants) among the different metagenomic samples. A challenge for efficient posterior sampling comes from the fact that the mixture components (which correspond to the variants' genome) take values on a large dimensional discrete space (of long nucleotide sequences), which is difficult to explore and thwarts standard Gibbs sampling or MCMC strategies by causing severe mixing issues. Drawing on recent advances in tempered Sequential Monte Carlo (Dau & Chopin 2022), we take advantage of natural tempering parameters in the mixture model to build efficient posterior sampling algorithms and leverage Sequential Monte Carlo estimates of marginal likelihoods to build a model choice framework for estimating the number of variants present in a collection of metagenomic samples.

Jordan Bryan	Empirical Bayes methods for linear models with heteroscedasticity

When the errors in a linear model have non-identity covariance, the ordinary least squares (OLS) estimate of the model coefficients is less efficient than the generalized least squares (GLS) estimate. However, the latter depends on the error covariance

matrix, which is unknown in practice. Feasible generalized least squares (FGLS) estimates first approximate the error covariance matrix and then use this approximation as a substitute for the ground truth. FGLS estimates often outperform OLS estimates, but they are not guaranteed to do so. In fact, they can be less efficient than OLS estimates if the approximation is poor. We characterize certain regimes in which FGLS is more efficient than OLS in linear models with independent, heteroscedastic errors. We show that an empirical Bayes (EB) procedure yields an FGLS estimate that outperforms that of OLS under sufficient heteroscedasticity. Under homoscedasticity, the EB FGLS estimate is nearly as efficient the OLS estimate. The EB FGLS estimate can be computed using simple linear regression tools, and its asymptotic variance has a closed form expression, which may be used for standard errors. Our results demonstrate how frequentist robustness can be obtained via Bayes procedures.

David Burt	Consistent Predictive Validation of Bayesian Spatial Models Under Infill Asymptotics

Bayesian methods have enjoyed success in a variety of spatial applications (including oceanography, air pollution, ecology, and remote sensing problems) due to their flexibility, ability to incorporate prior knowledge, and coherent uncertainty quantification. Since the models involved are inevitably misspecified, though, one must check predictive performance empirically. While many checks exist, these checks either (A) make the assumption that data is generated independently and identically, and thus can dramatically underestimate generalization error when data exhibits spatial correlation or (B) are not transparent about their assumptions or when they can be trusted. In the present work, we focus on the special case where: (1) we aim to predict at a particular set of test points and (2) infill asymptotics hold – that is, we assume validation data is available in a neighborhood of each test point. We demonstrate via simulations that under these assumptions, directly computing the average loss on the validation set can provide a poor estimate of risk. Conversely, under these assumptions, we provide the first proof of consistency for risk estimation in a spatial validation procedure. Under weak assumptions about the smoothness of the model and the spatial field being modeled, we provide a finite-sample error bound on the error of a validation method based on nearest neighbor regression. We support the consistency of this validation method with simulations using Gaussian process regression. And we illustrate the benefits of this method on Bayesian spatial regression applied to real weather data.

Giulia Carallo	Generalized Poisson Difference INGARCH

This paper introduces a novel stochastic process with signed integer values. Its autoregressive dynamics captures persistence in the conditional moments and is a convenient model feature for forecasting applications. The increments are Generalized Poisson differences and they capture over- and under-dispersion in the conditional distribution, extending standard Poisson difference models. We derive some properties of the process, such as stationarity conditions, stationary distribution, and conditional and unconditional moments, which are useful in forecasting. We further provide a Bayesian inference framework and an efficient posterior approximation based on Markov Chain Monte Carlo, allowing us to easily incorporate the inherent parameter uncertainty smoothly into the predictive distributions. Applications to benchmark datasets on car accidents and an original dataset on cyber threats show that the proposed model provides a better fitting and has superior forecasting abilities compared to standard Poisson models.

Lilia Carolina Carneiro da Costa	Learning Structure and Parameters of Dynamic Graphs: the mdmr package for R

A Multiregression Dynamic Model (MDM) is a Bayesian class of multivariate time series that represents various dynamic causal processes in a graphical way. Because the marginal likelihood has a closed form, model selection across many potential connectivity networks is easy to perform. With the application of the Integer Programming Algorithm, we can quickly find optimal models that satisfy acyclic graph constraints, and due to a factorization of the marginal likelihood, the search over all possible directed (acyclic or cyclic) graphical structures is even faster. The package mdmr implements the method developed by Costa et al. (2015) in R language, open source, and can be accessed at https://github.com/arzevedo/mdmr. In this work, we will present two applications that show the method's versatility using this package. In the first, we will investigate the relationship between new cases of COVID-19 in Brazilian macro-regions, and in the second one, we will use it to estimate inter-brain connectivity.

Jack Carter	Separable priors for Gaussian graphical models

A detailed examination of the use of separable priors, in particular spike and slab priors, for a Gaussian precision matrix. Specific considerations needed for such priors will be highlighted as well as some best practices for the choice of prior.

Particular attention will be made to the issues of scale invariance and positive definiteness of the prior, along with the importance of the diagonal entries of the precision matrix and how to best choose the prior for these diagonals.

Arhit Chakrabarti	Graphical Dirichlet Process for Clustering Non-Exchangeable Grouped Data

We consider the problem of clustering grouped data with possibly non-exchangeable groups whose dependencies can be characterized by a known directed acyclic graph. To allow the sharing of clusters among the non-exchangeable groups, we propose a Bayesian nonparametric approach, termed graphical Dirichlet process, that jointly models the dependent group-specific random measures by assuming each random measure to be distributed as a Dirichlet process whose concentration parameter and base probability measure depend on those of its parent groups. The resulting joint stochastic process respects the Markov property of the directed acyclic graph that links the groups. We characterize the graphical Dirichlet process using a novel hypergraph representation as well as the stick-breaking representation, the restaurant-type representation, and the representation as a limit of a finite mixture model. We develop an efficient posterior inference algorithm and illustrate our model with simulations and a real grouped single-cell dataset.

Cliburn Chan, Lynn Lin, Haiqi Zhang	Probabilistic Antigen Cartography Using Bayesian Multidimensional Scaling Procedure

Antigenic cartography is a technique used to visually represent and measure the antigenic relationships across various viral strains or serums. It aids in comprehending the evolutionary path of viruses by offering a distinct and measurable representation of the antigenic relationship between different virus strains. This is achieved by preserving as much antigenic information as possible while minimizing any potential loss during the reduction of dimensions. Antigenic cartography also enables the prediction of vaccination effectiveness and immune responses, facilitating the efficient management and control of viral illnesses like influenza. Antigenic mapping is not only restricted to influenza viruses, but may also be utilized for other diseases that exhibit antigenic variability, such as HIV, rabies virus, and Campylobacter jejuni.

Traditional methods for antigenic cartography, such as Classical Multidimensional Scaling (CMDS), are inadequate in accurately capturing the intricate and nuanced relationships between antigens. This is due to the inherent complexity of antigen

data and the chances of having missing data, or lower-than-detection-limit data in the assay results. This constraint can lead to inadequate or deceptive interpretations, particularly when dealing with highly changeable infections or complex immunological interactions. Therefore, this study intends to address these challenges by utilizing a novel Bayesian method, as Bayesian MDS is a suitable method for managing missing values and detecting thresholds. We can establish precise priors to account for the fundamental biological assumptions and uncertainties included in the data. Furthermore, Bayesian MDS has the capability to perform prediction, whereas CMDS lacks this ability. This makes Bayesian MDS a very suitable approach to employ on antigenic datasets. This study intends to discover disparities in SARS-CoV-2 information during the past two years that may have been overlooked by CMDS, using a comprehensive dataset. This research is significant since it presents a sophisticated method for mapping antigens. This method might be critical in evaluating the antigenic properties of emerging SARS-CoV-2 or influenza variants and influencing the design of future vaccines.

Andrew Chin	Bouncy Hamiltonian samplers: A unifying framework for Hamiltonian Monte Carlo and piecewise deterministic processes

Piecewise-deterministic Markov process (PDMP) samplers constitute a state of the art paradigm in Bayesian computation, with examples including the zig-zag and bouncy particle sampler. Recent work on the zig-zag has indicated the paradigm's potential connection to Hamiltonian Monte Carlo, a version of the Metropolis algorithm that exploits Hamiltonian dynamics. We demonstrate that the connection between the paradigms extends far beyond the specific instance. The key lies in the fact that any time-reversible deterministic dynamic provides a valid Metropolis proposal, and how PDMPs' characteristic velocity changes constitute an alternative to the usual acceptance-rejection. We turn this observation into a rigorous framework for constructing rejection-free Metropolis proposals based on bouncy Hamiltonian dynamics which, albeit having Hamiltonian-like properties, generate trajectories more similar in appearance to PDMPs. In fact, when combined with periodic refreshment of the inertia, the dynamics converge strongly to PDMP equivalents in the limit of increasingly frequent refreshment. To demonstrate the practical implications of this new paradigm, we present a sampler based on a bouncy Hamiltonian dynamic closely related to the bouncy particle sampler. The proposed

Hamiltonian bouncy particle sampler exhibits competitive performances on challenging real-data posteriors involving tens of thousands of parameters.

Stefano Cortinovis	Automatic inductive bias selection in dynamical systems

The assumptions that a statistical model makes about the underlying data-generating process, commonly known as its inductive biases, are often crucial to enable good generalisation beyond the training set. While inductive biases are usually either directly incorporated into the model architecture or encouraged through data augmentation, recent work has shown that the marginal likelihood constitutes a sensible training objective to learn them from data. In this work, we investigate the effectiveness of this approach in the context of dynamical systems. In particular, we develop a suitable Gaussian process model equipped with a kernel parameterising the range of structural properties that might arise in a dynamical system, such as energy conservation or dissipation. We train this model using a variational inference scheme and study the correspondence between the inductive biases inferred from data and the actual physical properties of the underlying system in a variety of settings.

Vaidehi Dixit	Anytime valid and asymptotically optimal statistical inference driven by predictive recursion

Distinguishing two classes of candidate models is a fundamental and practically important problem in statistical inference. Error rate control is crucial to the logic but, in complex nonparametric settings, such guarantees can be difficult to achieve, especially when the stopping rule that determines the data collection process is not available. My talk is based on our novel e-process construction that leverages the so-called predictive recursion (PR) algorithm. The proposal is based on constructing a marginal likelihood by mixing over a specified class of distributions. Such a likelihood could be constructed in a Bayesian way by introducing a prior on the class of distributions in the alternative and finding the corresponding Bayesian marginal likelihood. But implementing a purely Bayesian strategy to account for nonparametric aspects of applications can be computationally demanding. The PR algorithm stems as an approximation to the posterior mean of the mixing distribution under the Dirichlet process prior and hence is able to rapidly and recursively fit nonparametric mixture models. The resulting PRe-process affords anytime valid inference uniformly over stopping rules and is shown to be efficient in the sense that

it achieves the maximal growth rate under the alternative relative to the mixture model being fit by PR.

Shourya Dutta, Janet Van Niekerk, Haavard Rue	Correcting the Laplace Method with Variational Bayes: Marginal Skewness

Within the domain of fast and scalable approximated Bayesian Inference, the Laplace Method and Variational Bayes have been key contenders. However, accuracy can be compromised in pursuit of speed with these standalone methods. To address this challenge, we propose an innovative fusion of the Laplace Method with a Low-Rank Variational Bayes Correction (VBC) that particularly targets the correction of marginal skewness. Our approach not only refines the approximation of the posterior mean and variance but also introduces a novel distribution paradigm, liberating us from the constraints of Gaussianity prevalent in the Laplace Method. Notably, our method allows tailored and precise adjustments exclusively for user-specified covariates, showcasing impressive scalability without significant computational costs. Experimental validation, using both simulated and real-world data, highlights the efficacy of our approach in addressing the issue of marginal skewness. The proposed method is anticipated to be integrated into the widely used R-INLA package, ensuring accessibility and practical implementation for the research community.

Jonas Esser	Seemingly unrelated Bayesian additive regression trees for cost-effectiveness analysis in healthcare

In recent years, theoretical results and simulation evidence have proven BART (Bayesian Additive Regression Trees) to be a highly performant method for nonparametric regression. We propose a multivariate generalization of BART, to be applied in regression analyses with several dependent outcome variables. In the case with continuous outcomes, our model is essentially a nonparametric version of seemingly unrelated regression (SUR). Likewise, our proposal for binary outcomes is a nonparametric generalization of the the multivariate probit model. We give suggestions for easily interpretable prior distributions, which allow the specification of both informative and uninformative priors. For all proposed models, we provide a detailed discussion of MCMC methods to estimate the posterior distribution. Our ideas are implemented in an R package. We further use this implementation to assess the performance of our models, on both simulated and empirical datasets.

Anna Flowers	Auxiliary Latent Inputs for Jump Gaussian Processes

Gaussian Process (GP) regression is a flexible modeling apparatus that is attractive both for its predictive accuracy and its ability to provide uncertainty estimates. However, their distance-based covariance means that they are typically ill-suited for nonstationary data. We present two methods to improve prediction for data that comes from a mixture distribution whose components are independent but whose component membership is spatially dependent. This separates the space into regions whose response is dependent within region, but independent across regions. The optimal neighbor GP is an extension of the local GP that can vary the neighborhood size for each predictive location. This allows predictive locations near boundaries to choose smaller neighborhoods than those well inside of a region. The GP with auxiliary latent inputs is a sequential process that first estimates a point's component membership and then uses that information as a new input in the data. This decreases the distance between points that are likely to be in the same components and increases the distance between points on opposite sides of the boundary. Both methods have been shown to increase predictive accuracy, with the greatest improvement seen when they are used in combination.

Augusto Fasano	Scalable expectation propagation for probit models and beyond

Bayesian binary regression is a prosperous area of research due to the computational challenges encountered by currently available methods either in high-dimensional settings (large p) or large datasets (large n), or both. In the present work, we mainly focus on the expectation propagation (EP) approximation of the posterior distribution in Bayesian probit regression under a multivariate Gaussian prior distribution.

We first derive an efficient implementation having per-iteration cost scaling linearly in n (and quadratically in p), by leveraging results on the extended multivariate skew-normal distribution and showing that multiple improvements over state-of-the-art implementations are possible. Secondly, we show that the EP routine can be implemented at the per-iteration cost scaling linearly in p (and quadratically in n). This novel result opens the possibility of employing EP in settings where it was believed to be practically impossible. We show also that, as a byproduct, approximate EP predictive probabilities admit a simple closed form in terms of the parameters outputted by the algorithm. A simulation study shows the order of the improvements over state-of-the-art approaches. Finally, we adapt such derivations to

other generalized linear models, paving the way to the use of EP for a wide variety of routinely-used models for Bayesian inference.

Julie Fendler	Bayesian profile regression mixture models to assess the combined health effects of multiple correlated exposures: How to deal with an inconsistent estimated number of clusters?

We focus on the problem of estimating the combined health effects of multiple correlated environmental exposures from a censored time-to-event outcome. We consider the Bayesian profile regression mixture (BPRM) model proposed by Belloni et al. (2020): it can be seen as an extension of Dirichlet process mixture models (DPMs) where a survival outcome and multiple exposures are linked via a shared clustering structure. Interestingly, these mixture models allow clustering individuals with similar exposure characteristics and estimating the associated excess hazard risk for each group in a unique step. However, as illustrated in Belloni et al (2020) and Rouanet et al. (2023), when BPRM models are fitted using standard Metropolis-within-Gibbs algorithms, this may lead to an unstable estimated number of clusters: the Markov chains are most certainly stuck in local minima. Moreover, DPMs may overestimate the number of clusters, even in finite samples (Chaumeny et al. (2022)). Here, we propose and compare, from simulated data, different algorithms to overcome these issues: they include different sampling schemes to fit BPRM models (MCMC with slice sampler, parallel tempering, SMC algorithm, …) and different post-processing techniques of the MCMC (or SMC) outputs to identify a representative clustering structure (K-meloids, variation of information, merge-trunc-merge algorithm, …). Then, our algorithms are applied to fit the BPRM model proposed by Belloni et al. (2020) on the post-55 French cohort of uranium miners who were chronically and occupationally exposed to multiple and correlated sources of ionizing radiation (radon, gamma rays and uranium dust) in order to estimate their potential radiation-related risk of death by lung cancer.

Nina Fischer	History Matching As a Calibration Method for Carbon Cycle Models

Simulators, complex physical models implemented in computer code, are a fundamental tool to assess ecosystem dynamics. Coupled with observational data, models can infer unobserved ecosystem properties. An example of this is the carbon cycle model DALEC Crop, which models the effect of Nitrogen fertilisation on wheat

growth in a field over a crop growing season. Currently, DALEC Crop's parameters are calibrated using a computationally costly Adaptive Proposal Markov Chain Monte Carlo algorithm. The current framework further does not allow for exploration of the effects of structural discrepancy and parametric uncertainty within the modelling framework. This in turn limits our understanding of the fields' physical processes due to an underestimation of uncertainties when predicting yield or Leaf Nitrogen content.

To facilitate quick model calibration and a thorough analysis of the uncertainties involved with DALEC Crop, this work presents history matching combined with emulation as an alternative to the current calibration framework. Using observations of Leaf Area Index when history matching allows us to find the set of input combinations for which the simulator gives acceptable matches to observed data. We demonstrate the suitability of the methodology building on previous work and data presented in [Revill et al., Agronomy, 2021]. An increase in possible outputs' ranges highlights the need to explicitly account for uncertainties when assimilating observational data.

Jared D. Fisher	A Bayesian Classification Trees Approach to Treatment Effect Variation with Noncompliance

Estimating varying treatment effects in randomized trials with noncompliance is inherently challenging since variation comes from two separate sources: variation in the impact itself and variation in the compliance rate. In this setting, existing frequentist and machine learning methods are quite flexible but are highly sensitive to the so-called weak instruments problem, in which the compliance rate is (locally) close to zero. Parametric Bayesian approaches, which account for noncompliance via imputation, are more robust in this case, but are much more sensitive to model specification. In this paper, we propose a Bayesian semiparametric approach that combines the best features of both approaches. Our main methodological contribution is to present a Bayesian Causal Forest model for binary response variables in scenarios with noncompliance. In this Bayesian noncompliance framework, we repeatedly impute individuals' compliance types, allowing us to flexibly estimate varying treatment effects among compliers. We then apply the method to detect and analyze heterogeneity in a study of workplace wellness, where there are a plethora of binary outcomes of interest.

Jacob Fontana	Scalable M-Open Model Selection in Large Data Settings

We consider the variable selection problem for linear models in the M-open setting, where the data generating process is outside the model space. We focus on the novel problem of Model Superinduction, which refers to the tendency of model selection procedures to exponentially favor larger models as the sample size grows, resulting in overparametrized models which induce severe computational difficulties. We examine popular model selection priors, such as mixtures of g-priors and the family of spike and slab priors. We demonstrate that when utilizing these priors and comparing nested models, the larger model will always be asymptotically selected. We further show this behavior is inescapable for any KL-divergence minimizing model selection procedure, so we seek to minimize its effects for large n, while preserving posterior consistency. We propose variants of the aforementioned priors that result in a slowly diminishing rate of prior influence on the posterior, which favors simpler models while preserving consistency. We further propose a model space prior which induces stronger model complexity penalization for large sample sizes. We demonstrate the aforementioned phenomena, and the efficacy of our proposed solutions, via synthetic data examples and a case study using albedo data from GOES satellites.

Catherine Forbes	Summarizing case influence from Bayesian moment condition models

Moment condition models are common in Economics and related disciplines where theory implies a set of moment constraints, yet a full generative probability model is not specified. Case influence diagnostics provide a way to assess the sensitivity of the posterior distribution to perturbations of the data – something that is especially important in this context because the empirical moment conditions can be greatly affected by the presence of extreme observations. Measures of divergence between the full-data and case-deleted moment condition posterior are considered, as are measure of local influence from infinitesimal perturbations to the moment-based likelihood function. These more traditional methods are compared against a new technique that embeds data perturbations directly inside the moment conditions, resulting in posterior information regarding case-influence.

Mariano Gabitto	Hierarchical Bayesian Inference with repulsive priors to model continuous phenotypical progression in Alzheimer's Disease

The evolution of biological systems follows a continuous dynamic that many times is revealed through several biophysical processes that can be observed simultaneously. Inferring the latent phenotypical dynamic is a prominent topic of interest that can be model by assigning a score - termed pseudotime - to each observation as an alternative measure of progression that can be used to order observations and create a temporal series. This paper proposes a hierarchical Bayesian model for analyzing observations randomly distributed across time of multiple biophysical proposes. We devise a Bayesian repulsive prior to model points spread across a bounded interval and use it to represent samples from the underlying phenotypical dynamic. We provide Markov chain Monte Carlo estimation algorithms and use simulation studies to show the effectiveness of the proposed methods across data conditions. We apply our methodology to study the progression of Alzheimer's Disease in a brain area. We demonstrate that by carefully modeling the progression of pathological proteins in a single region, we can predict the progression of disease throughout the brain. We characterize inferential uncertainty and demonstrate how multiple measurements can be used to constrain inferential estimates through hierarchical modeling. We further investigate how predictions inform optimal experimental design for cohort selection by maximizing the observation of the entire phenotypical diversity.

Valentina Ghidini	Weighting covariates in Bayesian nonparametric clustering: an application to transportation networks

In clustering, observed individual data are often accompanied by covariates that can assist the clustering process itself. This is the case, for example, of transportation networks, where each node has spatial coordinates, and it is often desirable that clusters of nodes are spatially cohesive. In fact, the obtained clusters may be used to inform public policy decisions, and it may be preferable that such policies are uniform over neighboring areas. Naturally, depending on the application, different notions of closeness can be used to define such neighborhoods, thus potentially requiring to transform the spatial covariates.
Motivated by real-world data about subscriptions to the public transportation system of Bergamo (Italy) and its surroundings, we show how to incorporate properly transformed spatial covariates into a state-of-the-art stochastic block model, while inferring the weight of covariates.
(joint work with Sirio Legramanti and Raffaele Argiento)

Matteo Giordano	Nonparametric Bayesian intensity estimation for covariate-driven inhomogeneous point processes

The talk will consider nonparametric Bayesian estimation of the intensity function of an inhomogeneous Poisson point process in the important case where the intensity depends on covariates, based on the observation of a single realisation of the point pattern over a large area. It is shown how the presence of covariates allows to borrow information from far away locations in the observation window, enabling consistent inference in the growing domain asymptotics. In particular, minimax-optimal posterior contraction rates under both global and point-wise loss functions are derived. The rates in global loss are obtained under conditions on the prior distribution resembling those in the well established theory of Bayesian nonparametrics, here combined with concentration inequalities for functionals of stationary processes to control certain random covariate-dependent loss functions appearing in the analysis. The local rates are derived with an ad-hoc study that builds on recent advances in the theory of Pólya tree priors, extended to the present multivariate setting with a novel construction that makes use of the random geometry induced by the covariates. Joint work with Alisa Kirichenko and Judith Rousseau.

Ryan Giordano	Evaluating Sensitivity to the Stick-Breaking Prior in Bayesian Nonparametrics

Bayesian models based on the Dirichlet process and other stick-breaking priors have been proposed as core ingredients for clustering, topic modeling, and other unsupervised learning tasks. However, due to the flexibility of these models, the consequences of prior choices can be opaque. And so prior specification can be relatively difficult. At the same time, prior choice can have a substantial effect on posterior inferences. Thus, considerations of robustness need to go hand in hand with nonparametric modeling. In the current paper, we tackle this challenge by exploiting the fact that variational Bayesian methods, in addition to having computational advantages in fitting complex nonparametric models, also yield sensitivities with respect to parametric and nonparametric aspects of Bayesian models. In particular, we demonstrate how to assess the sensitivity of conclusions to the choice of concentration parameter and stick-breaking distribution for inferences under Dirichlet process mixtures and related mixture models. We provide both theoretical and empirical support for our variational approach to Bayesian sensitivity analysis.

Yiwei Gong	Detecting State Changes in Dynamic Neuronal Networks

Our brain function depends on temporal modulation of neuronal networks. Neuronal developmental disorders (NDDs) such as autism spectrum disorder and Rett syndrome are associated with certain patterns of modulation. Our research goal is to identify state changes in neuronal activity associated with the onset of NDDs. Existing statistical models for spike-train data focus on changes of amplitudes and correlations, instead of node connectivity; exploratory analysis on spike train data from in-vitro mouse MEA recordings with known timings of light stimulation suggests that they are insufficiently flexible for our purposes. We present a flexible nonlinear model based on factorial hidden Markov Model that is designed to capture the interactive behavior among nodes while characterizing the changes in spike intensity. We propose a variational inference algorithm using Gumbel-Softmax relaxation to infer the latent states and model parameters. The model is designed but not limited to reveal the switching process of brain neurons.

Henrik Häggström	Simulation-based Bayesian inference for intractable mixed-effects stochastic models of cell variation

Simulation-based inference is an important area in statistical inference, allowing parameter estimation for complex models. We construct a novel sequential Bayesian sampler, employing Gaussian locally linear mappings, to infer model parameters in models for cell-to-cell variation. Hierarchical mathematical models are used to model the dynamics of multiple cells, via mixed-effects modelling. Our inference methodology is able to produce surrogate models of the posterior and the likelihood, in closed form, returning Bayesian inference at a much lower computational cost compared to exact state-of-art Bayesian methodology that uses expensive particle MCMC approaches or approximate inference via neural density estimation. This reduction in computational cost has an important impact on the feasibility of mixed effects modelling applied to possibly hundreds of cell dynamics.
Joint work with Marija Cvijovic and Umberto Picchini.

Ofir Hanoch	Bayesian GMM with informative priors for the identification of sudden gainers in MDD treatment

Mental disorders such as major depressive disorder (MDD) affect a large number of individuals, families, informal and formal caretakers worldwide.

Growth Mixture Models (GMMs) are often used for MDD treatment analysis, to assess dynamics of latent classes and to account for heterogeneity in individuals'

response to MDD treatment. One of the relatively less researched aspects of MDD analysis using GMMs is patients who undergo sudden improvements in MDD symptoms, identified as sudden gainers, which constitute around 30% of cases. Identification in GMMs in general, and potentially for a GMM with sudden gainers, is complex due to typically small sample sizes in each study and large numbers of variables. We propose a new GMM for identifying depression patient subgroups with different change patterns and key discontinuities in symptom course such as sudden gainers. This is achieved with an absorbing state in the mixture model. We strengthen the inference of the model and identification of parameters using informative priors based on earlier studies. We apply the proposed method to a longitudinal dataset of MDD treatment. We show that particularly the dynamics of the latent groups, including sudden gainers, are estimated more accurately under the proposed approach. The results provide further insights into the underlying dynamics of sudden gainers from MDD treatment.

Brian Hassett	Latent position models for dynamic multiplex networks

Networks are used to represent interactions between entities. Examples include people connecting on social media, countries trading commodities, and protein-protein interactions. Statistical network analysis casts these interactions in a probabilistic framework which facilitates inference on the network. Real network data is often dynamic; over time, people change who they connect with on social media, and countries adjust their trading patterns from year to year. Multiplex networks contain multiple layers. In a social network, each layer could be a different social media company. In a trading network, each layer could be a different commodity. Networks with both of these properties are referred to as dynamic multiplex networks, and such networks typically exhibit correlation both temporally and between the layers. Latent position models are a family of models for network data. These models position the nodes of the network throughout a latent geometric space with the assumption that the probability of a connection between two nodes depends on their respective positions in the latent space. This work extends the original latent position model to simultaneously account for correlations between time points and layers in a dynamic multiplex network. The resulting model is estimated in a Bayesian framework and tested on simulated data from both binary and real-valued networks. The work is motivated using a real dataset involving international trade of agricultural products.

Disha Hegde	Learning to Solve Related Linear Systems

Solving multiple large linear systems defined across a parameter space is central to many numerical tasks, such as solving nonlinear PDEs arising from applications like fluid dynamics, optimising hyperparameters of Gaussian processes and finding coefficients for statistical models. The computational expense of solving these linear systems can be lowered if their interdependence across the parameter space is exploited efficiently. Viewing this problem from a probabilistic numerics perspective provides a better uncertainty quantification.

Our work extends the idea of probabilistic linear solvers like the Bayes Conjugate Gradient method across a parameter space, providing a way to utilize the underlying connection between the parameters of the linear systems. We use the Bayesian framework by placing a multi-output Gaussian process prior on the solution vector to model its correlation across the parameter space. This is then conditioned on information obtained by previously solved linear systems to construct a posterior. We discuss different choices of prior covariances, both separable and non-separable, and their properties. Our work also discusses the different choices of information to be conditioned on, and their properties. Motivated by the warm-restarting approach for solving multiple linear systems, we also discuss a filtering formulation for our approach.

Wei Jin

A Bayesian Decision Framework for Optimizing Sequential Combination Antiretroviral Therapy in People with HIV

Numerous adverse effects (e.g., depression) have been reported for combination antiretroviral therapy (cART) despite its remarkable success in viral suppression in people with HIV (PWH). To improve long-term health outcomes for PWH, there is an urgent need to design personalized optimal cART with the lowest risk of comorbidity in the emerging field of precision medicine for HIV. Large-scale HIV studies offer researchers unprecedented opportunities to optimize personalized cART in a data-driven manner. However, the large number of possible drug combinations for cART makes the estimation of cART effects a high-dimensional combinatorial problem, imposing challenges in both statistical inference and decision-making. We develop a two-step Bayesian decision framework for optimizing sequential cART assignments. In the first step, we propose a dynamic model for individuals' longitudinal observations using a multivariate Gaussian process. In the second step, we build a probabilistic generative model for cART assignments and design an uncertainty-penalized policy optimization using the uncertainty quantification from

the first step. Applying the proposed method to a dataset from the Women's Interagency HIV Study, we demonstrate its clinical utility in assisting physicians to make effective treatment decisions, serving the purpose of both viral suppression and comorbidity risk reduction.

Kaniav Kamary	Reference prior driven by reparameterisations of multivariate location-scale mixtures

Multivariate location-scale mixture distributions allow for modeling complex relationships in multivariate data by combining multiple distributions, each with its own location and scale parameters. Kamary et al. (2018) proposed a genuine, weakly informative prior for univariate location-scale mixtures through reparameterization centered on the global mean and variance. The reparameterization technique facilitated weakly informative prior modeling and the sampling procedure from the complex posterior distributions of the mixture parameters. In this paper, we extend the reparameterization idea to mixtures of multivariate location-scale distributions and demonstrate that the resulting posterior is proper when there is at least one data. Dependent/independent variable selection in the Bayesian framework is proposed, and by providing MCMC implementations, we illustrate the good performance of the model estimation throughout simulated and real data sets. Keywords and phrases: Mixture distribution, Reference prior, Multivariate location-scale distribution, Bayesian analysis.

Jennifer Kampe	Link prediction in dynamic plant-pollinator networks in an arctic community

Networks represented as adjacency matrices provide simple, interpretable summaries of highly complex ecological systems. The expansion of ecological monitoring technologies has dramatically increased the availability of interaction network data with spatial and dynamic replicates, yet sampled networks are still plagued by extensive missing data. We develop a bipartite network-valued stochastic process with temporal dependencies between networks described via a latent space model in which accommodates high levels of missing data. We implement this model via an efficient Gibbs sampler and illustrate its performance with a case study of plant-pollinator networks from the high arctic Zackenberg Valley.

Joseph Kang, Adam Hall	Bayesian multinomial logit models for the U.S. Voting Rights Act (VRA) determination

The U.S. Voting Rights Act (VRA) of 1965 is a landmark piece of federal legislation that prohibits discrimination at the polls based on race, color, or language minority status. Section 203 of the VRA requires the U.S. Census Bureau to provide data on the voting-age population in each state and county, including the number of citizens, limited English proficient (LEP) individuals, and those with limited education. In the 2021 cycle, Census Bureau statisticians developed Bayesian models to determine the number of language minority groups. In this talk, we present an updated version of our Bayesian model with the Polya-Gamma (PG) distribution (Polson et al., 2003), which transforms a non-linear computational optimization problem into a linear one. Using the PG distribution, we were able to perform stochastic search variable selection (George and McCulloch, 1993) for high-dimensional covariates and sequential regression multiple imputation (Raghunathan et al., 2001) for covariates with missing data.

Hyotae Kim	Marked Hawkes processes for earthquake occurrences: A Bayesian nonparametric modeling approach

We propose a Bayesian nonparametric model for marked Hawkes processes (MHPs). The processes' conditional intensity function is decomposed into the ground process intensity and the mark density function. We concentrate primarily on modeling the process intensity but offer several choices for the mark density as well. The prior probability model for intensity has been carefully designed to provide model flexibility and tractable posterior inference using a novel mixture modeling method. This model was motivated by seismology applications, where magnitude is regarded as a mark associated with a time point for earthquake occurrence. Accordingly, the mixture model basis is a function of occurrence time and magnitude, with its functional form selected considering not only model flexibility but also earthquake data characteristics, for example, the fact that earthquakes of greater magnitude cause more subsequent shocks than earthquakes of smaller magnitude. The model is illustrated with two earthquake occurrence datasets (Japan / Southwestern USA) and several synthetic examples.

Seongmin Kim	Bayesian analysis of the eigenstructure of the high-dimensional covariance with shrinkage inverse Wishart prior.

The eigenstructure of covariances is essential in understanding the covariance of multivariate observations. The inverse Wishart prior is commonly used for the

covariance estimation in Bayesian inference. However, in the high-dimensional setting where the number of covariates increases with the sample size, it is well-known that the estimates of eigenvalues tend to spread out. Surprisingly, there has been little prior study in Bayesian statistics on eigenstructures of unconstrained covariance. Recently, the shrinkage inverse Wishart (SIW) prior had been proposed to mitigate the limitation of the inverse Wishart. In this study, we investigate the asymptotic properties of the posterior of both the eigenvalues and eigenvectors. We discuss the pros and cons of the SIW in the high-dimensional setting. Specifically, we examine the behavior of the eigenvalues estimated using the shrinkage inverse Wishart prior in comparison to the inverse Wishart prior.

David Kohns	Shrinkage priors over auto-regressive dynamics

We present the AR-R2 prior, a joint prior over the auto-regressive time-series components in Bayesian models and their induced R^2. The prior is motivated by decomposing the variance contributions of the auto-regressive dynamics in Bayesian models, which we link to the expected predictive performance through its coefficient of determination, thus providing a means of avoiding over-fitting. Compared to other priors designed for times-series models, the AR-R2 prior allows for flexible and intuitive shrinkage which nests the behaviour of competing methods. We extend and derive the R^2 based prior framework for pure auto-regressive models, auto-regressive models with exogenous inputs, and state-space models. Through both simulations and real-world modelling exercises, we elucidate the efficacy of the AR-R2 prior in improving sparse and reliable inference, all while reducing over-fitting compared to competing shrinkage priors. We find that latent state-space models benefit the most from a joint prior on the predictive space.

Julia Kowalska	Bayesian regression discontinuity design with unknown cutoff

(Joint work with Stéphanie van der Pas and Mark van de Wiel).
Regression discontinuity design (RDD) is a quasi-experimental approach used to estimate the causal effects of an intervention assigned based on a cutoff criterion. Regression discontinuity design exploits the idea that close to the cutoff units below and above are similar; hence, they can be meaningfully compared. Consequently, the causal effect can be estimated only locally at the cutoff point. This makes the cutoff point an essential element of the RDD. However, especially in medical applications, the exact cutoff location may not always be disclosed to the researcher, and even

when it is, the actual location may deviate from the official one. Estimating the causal effect at an incorrect cutoff point leads to meaningless results, but since the cutoff criterion often acts as a guideline rather than a strict rule, the location of the cutoff may be unclear from the data. We use a Bayesian approach to incorporate prior knowledge and uncertainty about the cutoff location in the causal effect estimation. Yet, RDD is a local, boundary point estimation problem, whereas the Bayesian model is fitted globally to the whole data. In this work, we address a natural question that arises: how to make Bayesian inference more local?

Debamita Kundu	Bayesian Inference of Chemical Mixtures in Risk Assessment Incorporating the Hierarchical Principle

Analyzing health effects associated with exposure to environmental chemical mixtures is a challenging problem in epidemiology, toxicology, and exposure science. In particular, when there are a large number of chemicals under consideration it is difficult to estimate the interactive effects without incorporating reasonable prior information. Based on substantive considerations, researchers believe that true interactions between chemicals need to incorporate their corresponding main effects. In this paper, we use this prior knowledge through a shrinkage prior that a priori assumes an interaction term can only occur when the corresponding main effects exist. Our initial development is for logistic regression with linear chemical effects. We extend this formulation to include non-linear exposure effects and to account for exposure subject to the detection limit. We develop an MCMC algorithm using a shrinkage prior that shrinks the interaction terms closer to zero as the main effects get closer to zero. We examine the performance of our methodology through simulation studies and illustrate an analysis of chemical interactions in a case-control study of cancer.

Michele Lambardi di San Miniato	A Bayesian Two-Parameter Normal Ogive Model for Crowdsourced Fact-Checking

In the internet age, it has become unprecedentedly easy for people to contribute to the spread of news, over social media and other new means of self-expression. The demand for fast and accurate fact-checking has increased with respect to the past when information was more under the control of fewer subjects (newspapers, television, etc). As a possible solution, crowdsourcing in fact-checking relies on the masses themselves to rate the validity of claims. The task is related to information technology, so it has long been in the domain of machine learning and data mining.

We aim to provide a probabilistic framework that can quantify the uncertainty in classification. We present a statistical perspective on the problem based on the two-parameter normal ogive model, which we borrow from item response theory (IRT). The ogive model disentangles worker-specific rating behaviour from the intrinsic truthfulness of claims. The Bayesian framework is especially suited for handling the latent variables that are needed to model workers' behaviour and claims' features. The method is necessarily supervised, as experts must provide samples of correct judgments. The proposal implies an aggregation of crowd judgements that surrogates the expert implicitly. The workers contributions are automatically weighted in differently, depending on their ability to emulate expert rating. Here, we illustrate the method based on a dataset of political claims, each rated by an expert along with crowd workers.

Changwoo Lee	Logistic-Beta Processes for Modeling Dependent Random Probabilities with Beta Marginals

We propose a novel stochastic process called the logistic-beta process (LBP), whose finite-dimensional marginal is a multivariate generalized logistic distribution with a highly flexible dependence structure. The logistic transformation of the LBP leads to a stochastic process with common beta marginals, and we propose to use the LBP as a prior for logit-transformed dependent random probabilities for Bayesian inference and prediction. The LBP induces marginal beta priors on random probabilities while having a flexible dependence structure, which is applicable to multiple existing dependent Bayesian nonparametric models involving beta distributions. We study several dependence cases, including temporal, spatial, and a flexible model class relying on feature maps. The LBP facilitates posterior computation due to a normal variance-mean mixture representation distinct from copula-type constructions, and we show its computational benefits through nonparametric binary regression simulation studies. We apply LBP in modeling dependent Dirichlet processes, and illustrate its application and benefits such as for Bayesian density regression problems in a toxicology study.

Alice L'Huillier	A scalable Bayesian method for estimating a fixed number of coordinates in the high-dimensional sparse linear regression model.

We study the problem of inferring one or a fixed number of coordinates in the high-dimensional linear regression model under sparsity constraints via a Bayesian

approach. We define an appropriate sparse prior on the whole regression vector that targets the coordinates of interest. Under compatibility conditions on the design matrix, the posterior distribution induced on the coordinates of interest is shown to have an asymptotically Gaussian shape in the form of a Bernstein-von Mises theorem. As simulating from the posterior distribution itself can be computationally intensive, we propose a variational class and consider the variational approximation to the posterior within this class. This variational approximation is shown to have the same desired asymptotic properties as the posterior, is easy to simulate from and can be computed by a CAVI algorithm, thereby providing a tractable method to build confidence intervals. The method is implemented on simulated data illustrating that our Bayesian procedure can be computed in comparable time to other frequentist methods while exhibiting favorable performance in a number of different settings. This is joint work with Ismaël Castillo (Sorbonne), Kolyan Ray (Imperial College) and Luke Travis (Imperial College).

Karina Lilleborge	Spatial modeling of public transport data using SPDEs on metric graphs

Public transport (PT) is recognized as one of the key contributors to sustainability in densely populated areas by reducing CO2 emissions from shorter inter-city journeys. PT companies automatically collect an increasing amount of data, which has the potential to reveal important spatio-temporal patterns in urban traffic, thereby aiding the planning of future sustainable cities.

The peculiarity of road traffic data lies in the fact that the spatial domain is not continuous but is defined by the road network. Currently, there is still a lack of computationally efficient statistical models that can properly describe spatio-temporal dependencies on metric graphs. In this work, we apply a recent proposal by Bolin et al. (2023), which introduces a Gaussian random field on metric graphs through a stochastic partial differential equation. This approach allows for efficient and inexpensive computations by leveraging the sparsity of the precision matrix. We extend the model by introducing a new observational model where data are not collected pointwise but as averages over segments of the graph. This reflects the structure of the PT data, where bus speed is derived from the arrival and departure times at different stops. We demonstrate, through a simulation experiment and a case study set in the city of Trondheim, Norway, how our model enables the recovery of the main traffic patterns in the area of interest.

Benoit Liquet	Nonstationary Spatial Process Models with Spatially Varying Covariance Kernels

Spatial process models for capturing nonstationary behavior in scientific data present several challenges with regard to statistical inference and uncertainty quantification. While nonstationary spatially-varying kernels are attractive for their flexibility and richness , their practical implementation has been reported to be overwhelmingly cumbersome because of the high-dimensional parameter spaces resulting from the spatially varying process parameters. Matters are considerably exacerbated with the massive numbers of spatial locations over which measurements are available. With limited theoretical tractability offered by nonstationary spatial processes, overcoming such computational bottlenecks require a synergy between model construction and algorithm development. We build a class of scalable nonstationary spatial process models using spatially varying covariance kernels. We present some novel consequences of such representations that befit computationally efficient implementation. More specifically, we operate within a coherent Bayesian modeling framework to achieve full uncertainty quantification using a Hybrid Monte-Carlo with nested interweaving. We carry out experiments on synthetic data sets to explore model selection and parameter identifiability and assess inferential improvements accrued from the nonstationary modeling. We illustrate strengths and pitfalls with a data set on remote sensed normalized difference vegetation index.

Ximena Castañeda-Ochoa, Yojan Alcaraz-Pérez, YeisonY. Ocampo-Naranjo, Isabel C. Ramírez-Guevara, Carlos M. Lopera-Gómez	"conjugate models": Understanding the conjugate models in a Bayesian statistics undergraduate course

Conjugate models are the first models to be studied in a Bayesian statistics undergraduate course. However, the limited time in classes only allows us to analyze a few examples, restricting a deeper understanding of the concepts. For this reason, a Shiny application is proposed in which students can interact with the discrete conjugate models, specifically the Poisson and binomial cases; and the normal conjugate models, in the cases: 1) Unknown mean and known variance, 2) Known mean and unknown variance, 3) Unknown mean and variance with independent a priori distributions, and 4) Both unknown mean and variance with a priori distribution of the mean conditioned on the variance. In any case, the students can either simulate both the data distribution and the prior distribution by setting the

corresponding parameters' values, or loading a set of data associated with the distribution of interest, and simulating the prior distribution through setting the parameters' values.

Nancy Lopes Garcia	Modelling correlated funcional data

The motivation for this study comes from a randomized placebo controlled depression clinical trial of sertraline. Although several factors have been studied in relation to treatment response in depression, it is well known that individual responses to treatment can vary widely. The dataset consists of major depressive disorder patients, randomized to either a drug or placebo treatment followed over a period of eight weeks. For each subject, several scalar and categorical covariates are available, as well as their resting state electroencephalography (EEG) under a closed eyes condition. The goal is to verify if the EEG measurements can be used to tailor treatment to patients. This EEG data contains the current source density amplitude spectrum values at a total of 14 electrodes located in occipital and parietal brain regions. It is necessary to include in the modelling the correlation among the EEGs. This can be done using dimension reduction techniques such as principal component analysis. Although this method is powerful, it requires all functional covariates to be observed at the same points across observations. Also, it does not take advantage of the intrinsic ordering of observations (e.g., in time or space) given by the functional structure of the covariate. We propose a functional data approach to analyse the EEG data based on a mixed-effect model across individuals. This is a joint work with Mariana R. Motta, Eva Petkova, Thaddeus Tarpey and R. Todd Ogden.

Freddy López Quintero	How wide is the Pay Disparity Between Men and Women in Chile? A Bayesian GMM approach

Gender gap is a concept that consider the differences among women and men in terms of their levels of participation, access, rights and maybe more importantly, remuneration or benefits measured through salaries. In this regard, we inspect the behavior of men and women salaries separately in Chile by using grouped counties data in order to understand their distributions. The main variable we are using is the proportion of men and women, respectively, that earn more than the mean (or median) salary in each county. We model them using the simplex distribution employing the Generalized Method of Moment under a Bayesian framework and compare this approach with several other including the traditional Maximum Likelihood.

Yiyong Luo	Non-linearity in Factor-Augmented Factor Autoregression

Prompted by events such as the global financial crisis and the COVID-19 pandemic, recent behaviour of the economy has revealed limitations in linear time series models when used for forecasting and structural analysis. Among linear models, Factor-Augmented Vector Autoregressive (FAVAR) class leverages extensive information through latent linear dynamic factors and models these factors with a VAR. This study investigates two distinct forms of non-linearity - neural networks and time-varying parameterization - within the framework of FAVAR. We extract dynamic factors using an autoencoder (rather than the traditional Principal Component Analysis), and these factors are modelled using a time-varying parameter VAR (TVP-VAR). We train the autoencoder with gradient descent and estimate the parameters in the TVP-VAR using Bayesian inference. We employ dynamic model averaging to mitigate overfitting by choosing between models with different types of non-linearity and dynamic. A simulation study then sheds light on the merits of both non-linear approaches, and we will apply the method to a real data.

Roberto Macri Demartino	Bayesian Methods for the Integration of Historical Data

In recent years, there has been an increasing interest and acceptance in utilizing and integrating historical data. Given the inherent nature of sequential information updating, Bayesian methods are natural for this purpose. Moreover, the process of informative prior elicitation is widely recognized as a complex and multifaceted task.

Within this context, the concept of power priors (Chen and Ibrahim, 2000) has emerged as a popular approach for incorporating historical data into the prior distribution. The power prior methodology heavily relies on a power parameter, ranging between 0 and 1, that is a crucial factor in determining the degree to which the historical data influences the prior distribution. The optimal prior distribution for the power parameter should balance the aims of encouraging borrowing when the data are compatible and limiting borrowing when they are in conflict. Consequently, the ability to effectively elicit an optimal initial prior for this parameter is a crucial step not fully investigated.

Additionally, analyzing replication studies involves per definition the use of historical data. Notably, power priors (Pawel et al., 2023) and hierarchical models (Bayarri and Mayoral 2002; Pawel and Held 2020) are two leading approaches in this domain.

However, the development of mixture models for analyzing replication studies is still an unexplored field. We propose a novel and conceptually intuitive Bayesian approach for quantifying replication success based on the use of mixture priors. Moreover, we explore the use of a Simulation-based Calibrated Bayes Factor, employing hypothetical replications generated from the posterior predictive distribution, to discriminate between competing initial Beta prior specifications for the power parameter.

Francesco Mascari	Measuring dependence with characteristic symmetric positive definite kernels under partial exchangeability

Partial exchangeability is a natural assumption in Bayesian nonparametrics when the observations are grouped in a finite number of blocks with similar features. The distribution of each block is modeled with a random probability measure, leading to a vector of joint de Finetti measures. Given two groups, the dependence structure can show a continuous spectrum of behaviors between two extremes: either total independence or almost sure equality of the two random probability measures. We propose a correlation index based on characteristic symmetric positive definite kernels that measures how far a partially exchangeable model is to the two extreme cases under general assumptions. Moreover, a) it generalizes the set-wise linear correlation, which is the most common measure used in the Bayesian nonparametric literature for this scope; b) it can be applied to both prior and posterior distributions; c) it preserves analytical tractability under mixing. These results are then specialized for various popular Bayesian nonparametric models, such as the Hierarchical Dirichlet Process and Dependent Dirichlet Processes. Numerical simulations confirm our theoretical understanding of our proposal as a measure of dependence.

Jeffrey W Miller	Compressive Bayesian non-negative matrix factorization for mutational signatures

Non-negative matrix factorization (NMF) is widely used in many applications. Inferring an appropriate number of factors for NMF is a challenging problem, and sparse Bayesian models have been proposed for this problem. However, inference in these models is difficult due to the complicated multimodal posterior distributions that arise. We introduce a novel methodology for overfitted Bayesian NMF models using ``compressive hyperpriors'' that force unneeded factors to epsilon while imposing mild shrinkage on needed factors. The basic idea is to use simple semi-conjugate priors but set the strength of the hyperprior in a data-dependent

way in order to achieve compressivity. This yields an easy-to-implement Gibbs sampler that has improved mixing and accuracy compared to state-of-the-art alternatives. We demonstrate in simulations and on real data for mutational signatures analysis in cancer genomics.

Hossein Moradi Rekabdarkolaee	Soybean Prediction using Computationally Efficient Bayesian Spatial Regression Models and Satellite Imagery

Remote sensing-based models can be used to create yield maps for precision treatments. However, many remote sensing-based yield models have relatively low accuracy and are computationally inefficient. Many current yield prediction models are based on machine learning techniques that behave as a black box. An alternative approach is to use of Bayesian statistics is to create models that are computationally efficient and produce coefficients that have meaning. The objective was to develop a computationally efficient Bayesian regression model for predicting soybean (Glycine max) yields using PlanetScope imagery. Because computational efficiency can be improved by reducing the computational steps, this paper investigated the importance of considering short-range and long-range spatial correlation and the data set population distribution. In the model, the Bayesian approach was employed to reduce uncertainty and the analysis showed that considering short-range spatial correlation was more important than considering the population distribution and long-range spatial correlation. Among the models tested, the Bayesian Nearest Neighbor Gaussian Process (R2=0.77) model which captures the short-range correlation outperformed the Bayesian Multiple Linear Regression (R2=0.29), the Bayesian Spatial Regression (R2=0.57), and the Bayesian skewed Spatial Regression (R2= 0.56) models. However, associated with improved accuracy was an increase in run time from 449 seconds for the Bayesian multiple linear regression model to 2013 seconds for the Bayesian Nearest Neighbor Gaussian Process model. A comparison between the Bayesian nearest neighbor process model and the non-Bayesian random forest machine learning model showed that the Bayesian model had higher computation time and an 18% improvement in forecasting accuracy. These data show that relatively accurate within-field yield estimates can be obtained without sacrificing computational efficiency and coefficients that have biological meaning.

Alexander Mozdzen	Non-local mixture models with repulsive weights

Mixture models are a popular tool for studying data which is assumed to arise from multiple clusters. Common problems when employing standard mixture models to

uncover clusters within a sample are the estimation of overlapping clusters and interpretability. Repulsive mixtures aim to ameliorate this drawback by favouring the estimation of well separated component locations. In a Bayesian framework, this is achieved by incorporating the distance between the locations into their prior distribution. Such a distribution can be constructed by multiplying a standard, potentially conjugate prior with a repulsive term. Numerous ways of incorporating a repulsive term into the locations of mixture models as well as computational methods for their estimation have been proposed in the literature. We extend this development and propose a mixture model with a repulsive potential acting on both locations and weights. This strategy leads not only to well separated components but also to non-negligible weights, greatly improving interpretability and reducing redundancy in the estimates.
Posterior inference is performed through a tailored MCMC scheme, including a birth-and-death step to update the random number of components. The performance of the model is demonstrated in a simulation study as well as on real data applications.

Sayan Mukherjee	Perturbed Bayesian Best Response in Continuum Games

The notion of perturbed Bayesian best response dynamic for continuum strategy Bayesian population games is introduced. Fundamental properties of the dynamic such as existence of perturbed Bayesian equilibrium, convergence of the perturbed equilibrium to the Bayesian equilibrium of the underlying game, as well as existence, uniqueness and continuity of solutions from arbitrary initial conditions is established. As applications to the theory, convergence of solutions to the perturbed Bayesian equilibria is shown to hold for two classes of games, namely, Bayesian potential games and Bayesian negative semidefinite games.

Roi Naveiro	Simulation based Bayesian Optimization

Optimizing black-box functions of categorical variables has important applications. Bayesian optimization is widely used in this type of problem. It involves adjusting a probabilistic predictive model of the objective and using an acquisition function to guide the optimization process.

In contrast to traditional approaches in Bayesian Optimization, we present a novel simulation-based method for sequentially optimizing the acquisition function, that

eliminates the need for analytical access to the posterior predictive distribution of the objective given the covariates.

This allows us to use a wide variety of surrogate probabilistic models that could be potentially better suited for categorical variables than the commonly used Gaussian Processes. We address convergence issues and demonstrate its effectiveness empirically.

Riccardo Omenti	A Bayesian Model to Estimate Male and Female Fertility Patterns at a Subnational Level

Accurate subnational fertility estimates are crucial for shaping policy decisions across diverse sectors, including education, health care, and social welfare. One of the major challenges in producing these estimates is the presence of small populations, in which information about birth counts stratified by the age of the parent at the birth of the child may be lacking or inadequate. In this research paper, we describe a Bayesian model tailored to estimate the period Total Fertility Rates (TFR) for both men and women at a subnational level. Building on previous work by Schmertmann & Hauer (2019), the model utilizes population counts from age-sex pyramids and models age-specific mortality and fertility patterns allowing for uncertainty. We present a real data application focusing on fertility estimation in US counties for the historical period 1982-2019. Preliminary results reveal distinctive fertility trajectories for men and women across different US counties. Furthermore, the proposed model exhibits significant potential for the examination of male and female fertility behaviors across diverse regions and time frames in multiple countries.

Paolo Onorati, Brunero Liseo	Modified Gaussian Processes for Nonparametric Binary Regression

Gaussian processes represent a standard tool for nonparametric regression. Their popularity in Bayesian statistics is motivated by the fact that they provide a manageable and flexible prior on a function space.

A fundamental property of the Gaussian processes - in a Bayesian regression framework - is the conjugacy when the noise is also Gaussian. However, for several applications, the use of a model with continuous response is not appropriate. Also, it is well known that, in a parametric binary regression model, as the logistic regression

and/or the probit model, the Gaussian prior on the coefficients is no longer conjugate.

This lack of conjugacy also affects the nonparametric version, and the resulting "posterior process" is no longer Gaussian, independently of the prior process. Recently, there were introduced conjugate families for the probit and logit models, at least in the parametric case. Similar results were also obtained in the nonparametric version of the probit model. The gist of our work is to fill the gap and to extend the conjugacy property to the case of the nonparametric logit model.

We introduce a new class of distributions, that is the Extended Quasi SUN family which generalizes the SUN class; furthermore, we show that this new class of densities is still conjugate - in a parametric binary regression set-up with a large family of link functions including the logit and probit ones. In addition, it maintains some closure properties of the SUN distribution and it is then easy to generalize those properties to an infinite dimensional case. In other words, we define a new family of stochastic processes, and we refer to it as the Extended Quasi SUN process. In this way all finite collections of random variables are distributed as an Extended Quasi SUN distribution. Furthermore, we show that this new stochastic process is conjugate with a nonparametric binary regression model for a large family of link functions including the logit and probit ones.

In addition, if one adopts a Gaussian process prior, the resulting posterior process is an Extended Quasi SUN process.

From a computational perspective, we also introduce a novel variational Bayes approach based on the stochastic representation of the Extended Quasi SUN process. This new method is available both for logit and probit links and it allows some improvements with respect to the existing approaches. In particular, the new approximating density is no longer Gaussian, like in Laplace approximation and expectation propagation algorithms. A scale mixture of the new approximating density can be easily implemented, and it can be used as a proposal density in an importance sampling algorithm, in order to produce a weighted and independent sample from the posterior process. We perform simulation studies and a real data analysis both for the logit and probit links in order to compare our approach with

some competitors such as Laplace approximation, expectation propagation, previous variational Bayes methods and a MCMC method.

Lorenzo Pacchiardi	Bayesian inference of capabilities of artificial intelligence systems

Artificial Intelligence (AI) systems (such as reinforcement-learning agents and large language models) are typically evaluated by testing them on a benchmark and reporting an aggregated score. As benchmarks are constituted of tasks demanding various capability levels to be completed, the aggregated score is uninformative of performance in new settings. We present a new framework aiming to infer the level of multiple capabilities of an AI system based on its granular performance on a set of tasks. In practice, our method relies on specifying a likelihood connecting capabilities and task demands to the observed task performance, using which the posterior distribution of the capability levels can be inferred. The likelihood determines how different capabilities and task demands interact and can be informed by insights from psychometrics and domain knowledge, thus making the method scientifically informative and interpretable. Moreover, disentangling the evaluation of performance from a specific benchmark allows to predict performance on new tasks, while the Bayesian framework seamlessly provides uncertainty quantification. We demonstrate this framework on large language models, showing how the performance prediction generalises to out-of-distribution tasks.

Dario Palumbo	Bayesian Dynamic Calibration of Models Predictions

This paper proposes a calibration model that combines and dynamically calibrates predictive densities. While the weights are statically estimated the time-varying calibration is introduced giving an observation driven dynamics to the parameters of the calibrating function which is driven by the score of the assumed conditional likelihood of the data generating process. The model is very flexible and can handle different shapes, instability and model uncertainty in the data generating process density. We show its effectiveness on various simulated datasets. Two empirical applications are also introduced, one on financial index density forecasts and one on short-term wind speed predictions. Both the simulations and the empirical applications documents the large instability of individual model performance in comparison to the properties of the combined and calibrated forecasts, favouring our model in terms of predictive accuracy.

Francesca Panero	Forecasting food insecurity with Gaussian processes on auxiliary data

Estimating the share of population that is food insecure is one of the major tasks that governments and international organisations need to address in order to direct their humanitarian aid efficiently. Typically, the estimation is conducted through surveys that are collected via internet, phone or in person. For certain countries, the data collection procedure can be expensive, potentially dangerous and subject to irregular time patterns. To address this, we will forecast food insecurity in different regions of selected countries using auxiliary information which is readily available, such as economical, weather and conflict data. Gaussian processes are a strong candidate to perform this estimation, thanks to their flexibility and interpretability in the description of spatio-temporal data, the integration of expert opinions into the choice of kernels and priors, the natural uncertainty quantification and the possibility of introducing hierarchies at spatial level. We will present the model framework and present our forecasting results on some real datasets, comparing with other approaches in literature. This work is joint with the United Nations World Food Programme Hunger Monitoring Unit.

Tamas Papp	Scalable couplings for Metropolis-Hastings algorithms

There has been a recent surge of interest in coupling methods for Markov chain Monte Carlo (MCMC) algorithms. These methods enable principled parallel MCMC with many short chains by removing the bias of MCMC estimates. Furthermore, they facilitate robust convergence quantification through estimated upper bounds on metrics such as the total variation distance.

We tackle the problem of designing practical couplings that scale well with the dimension of the target measure. Methodologically, for the Random Walk Metropolis kernel, we introduce certain low-rank corrections of the synchronous coupling which are provably optimal in standard high-dimensional asymptotic regimes. Our proposed couplings improve upon the status-quo by orders of magnitude when the target is both eccentric and of moderate dimension. Analytically, we bridge the gap to the optimal scaling literature and derive ordinary differential equation limits in the stylized, but important, case of asymptotically Gaussian targets. In particular, these limits enable us to infer optimal step size scalings under our proposed couplings. We discuss extensions of our couplings and results to Metropolis-adjusted Langevin and Hamiltonian Monte Carlo kernels.

Antonio Peruzzi	A Bayesian Dynamic Latent-Space Model for Asset Clustering

Periods of financial turmoil are not only characterized by higher correlation across assets but also by modifications in their overall clustering structure. In this work, we develop a bayesian dynamic Latent-Space mixture model for capturing changes in the clustering structure of financial assets at a fine scale. Through this model, we are able to project stocks onto a lower dimensional manifold and detect the presence of clusters. The infinite-mixture assumption ensures tractability in inference and accommodates cases in which the number of clusters is large. The Bayesian framework we rely on accounts for uncertainty in the parameters' space and allows for the inclusion of prior knowledge. After having tested our model's effectiveness and inference on a suitable synthetic dataset, we apply the model to the cross-correlation series of two reference stock indices. Our model correctly captures the presence of time-varying asset clustering. Moreover, we notice how assets' latent coordinates may be related to relevant financial factors such as market capitalization and volatility. Finally, we find further evidence that the number of clusters seems to soar in periods of financial distress.

Peter Potaptchik	Parametric Bayes + Infinitesimals = Nonparametric Bayes

Nonparametric Bayesian models are Bayesian models with infinite-dimensional parameter spaces. Historically, many prior distributions in this setting were derived by taking limits of priors on sequences of related, finite-dimensional models. The same approach was taken to deriving conditional distributions. These derivations were viewed, by some, as informal arguments. Considerable effort was expended re-deriving the same results "rigorously" via measure theory. In this work, we show that several "informal" limiting constructions are, in fact, rigorous nonstandard analytic arguments. To do so, we prove a key result demonstrating that conditioning can be commuted with the pushdown operation, which allows us to relate infinite-dimensional stochastic processes to elementary, (hyper)finite-dimensional ones.

Estevao Prado	Metropolis-Hastings with fast, flexible sub-sampling

Metropolis-Hastings (MH) is a flexible and commonly used MCMC algorithm which requires all observations in a dataset to correctly sample from the target distribution. For large datasets, however, it can be unviable to evaluate the likelihood on the full data for thousands of iterations. We propose a new minibatch MH which scales the

MH algorithm by using control variates and performs exact Bayesian inference on large datasets. We prove that our method satisfies detailed balance with respect to the target distribution as well as establish the optimal scaling parameter in the proposal distribution. Through theoretical results, simulation experiments and results on real-world datasets, we show that our method outperforms other minibatch MH algorithms on a subset of generalised linear models.

Rakesh Ranjan	Bayesian Variable Selection in Weibull Regression using Gaussian Process Prior

This paper considers the Bayesian variable selection in the Weibull nonparametric regression model. The variable selection in the model is derived using a Gaussian process prior. The model comparison uses the Bayes factor obtained from the Trans-dimensional Transformation-based Markov Chain Monte Carlo (TTMCMC) sample. We also derived the theoretical results related to the convergence of the Bayes Factor. Numerical illustrations are provided for simulated and real datasets in low as well as high-dimensional setups. Real data illustrates the selection of genes that can predict Uterine Serous Carcinoma Patient Survival.

Lizanne Raubenheimer	Bayesian Estimation of Cronbach's Alpha in the Case of More than One Balanced Random Effects Model

Cronbach's alpha (α) is used as a measure of reliability in for example psychological research. The reason for Cronbach's alpha's popularity is that it is computationally simple. Only the sample size and the variance components are needed and it can be computed for continuous as well as binary data. For m experiments with equal α but possibly different variance components the combined Bayesian estimation of Cronbach's alpha is considered. Since our model is the one way balanced random effects model, the assumption of equicorrelated normal data is satisfied. Our approach is therefore different from the method used in literature where the assumption of equicorrelation was made. It is shown that for α, the parameter of interest, the reference and the probability-matching priors are the same. The Bayesian theory and results are applied to two examples. The frequentist coverage of the credibility intervals are for all practical purposes 95%. It is also clear that the posterior distribution of α for the combined sample becomes more important as the number of samples increase.

Isaac Ray	DensityBAST: Constrained Spatial Point Process Intensity Estimation using Additive Ensembles of Spanning Trees

Point process intensity estimation on complex domains is a difficult problem as most existing methods, such as kernel or spline methods, are not suitable for domains with irregular boundaries, sharp concavities, or interior holes. This article proposes a Bayesian additive ensemble of spanning trees model for intensity and density function estimation (DensityBAST) via inhomogeneous Poisson point processes on complex domains. Our model uses a random spanning tree weak learner which can produce flexible and contiguous partitions of the domain while respecting its geometry and constraints, allowing it to handle both smooth and discontinuous density functions. Using the a data thinning strategy and Poisson-Gamma conjugacy, we develop an exact and fast Bayesian inference algorithm to fit DensityBAST. We demonstrate the advantages of our model over other methods in simulation studies and in an application to basketball data.

Ruchira Ray	Theoretical Guarantees for Data-Dependent Posterior Tempering

Power posteriors reduce the influence of the likelihood in the calculation of the posterior by raising the likelihood to a constant fractional power. This procedure, also known as posterior tempering, has been shown to exhibit appealing properties, including robustness to model misspecification and asymptotic normality (Bernstein-von Mises). However, practical recommendations for selecting the power and theoretical guarantees for data-driven power selection methods remain open questions. We engage with these issues by connecting posterior tempering to penalized estimation. Data-driven approaches to tuning the power in these settings suggest a novel asymptotic regime where the power decays to 0. In this new regime, we provide sufficient conditions for (i) asymptotic normality of the power posterior and its variational approximation, (ii) convergence of the power posterior moments to those of the limiting normal distribution in (i), and (iii) asymptotic equivalence between the power posterior mean and the maximum likelihood estimator.

Lorenzo Rimella	Simulation Based Composite Likelihood

Inference for high-dimensional hidden Markov models is challenging due to the exponential-in-dimension computational cost of the forward algorithm. To address this issue, we introduce an innovative composite likelihood approach called ``Simulation Based Composite Likelihood'' (SimBa-CL). With SimBa-CL, we approximate the likelihood by the product of its marginals, which we estimate using Monte Carlo sampling. In a similar vein to approximate Bayesian computation (ABC),

SimBa-CL requires multiple simulations from the model, but, in contrast to ABC, it provides a likelihood approximation that guides the optimization of the parameters. Leveraging automatic differentiation libraries, it is simple to calculate gradients and Hessians to not only speed-up optimization, but also to build approximate confidence sets. We conclude with an extensive experimental section, where we empirically validate our theoretical results, conduct a comparative analysis with SMC, and apply SimBa-CL to real-world Aphtovirus data.

Alan Riva-Palacio	Bayesian Semiparametric General Hazard Models

We propose a novel structure for Bayesian semiparametric hazard regression models which contains the Cox and accelerated failure time models as particular cases. We employ a Bayesian nonparametric approach based on neutral to the right priors where we allow for time dependent covariates. A posterior characterization and a fully Bayesian inference scheme are provided. We discuss asymptotic properties of the proposed model. Simulated and real data studies show the effectiveness of the general hazard construction.

Stefano Rizzelli	Asymptotic properties of Bayesian analysis of Peaks-over-Threshold

The Peaks Over Threshold (POT) method is the most popular statistical approach for the analysis of univariate extremes. Even though there is a rich applied literature on Bayesian analysis of POT, there is no asymptotic theory for the existing procedures. Even more importantly, the challenging problem of predicting future extreme events according to a proper probabilistic forecasting approach has received no attention to date. In this work, we develop the asymptotic theory for Bayesian inference (estimation and prediction) within the POT method. We extend such an asymptotic theory to cover posterior inference on the tail properties of the conditional distribution of a response random variable with respect to a vector of random covariates, as well as prediciton of future extremes in such a conditional context. This is a joint work with Clément Dombry (Université de Franche-Comté, Besançon) and Simone A. Padoan (Bocconi University, Milan).

Tomas Rodriguez Taborda	Bayesian analysis for the geochemical risk in the Province of Huelva, Spain

Global concern about soil quality has risen in recent decades since most human activities depend on it. Soils are associated with food production, one of the main

challenges of sustainable development. The geochemical composition of soils is determined mainly by the geological environment and the economic activities carried out on them. From a geological point of view, the province of Huelva, located in the southwest of Spain, is divided into three main blocks. From north to south is the Ossa Morena Zone (ZOM) with the presence of metamorphic rocks, followed by the Sub-Portuguese Zone (ZPS), where the Iberian Pyritic Belt is located, and in the southernmost part are the sediments of the Guadalquivir Basin of Miocene-Pliocene age. The province of Huelva has been subjected to intense anthropogenic activities such as mining, industry, and agriculture, which are closely related to geology. For this work, 229 soil samples were collected from 3 locations, Aroche, El Campillo, and Huelva Capital, which differ from each other from a geological point of view, soil characteristics, and economic activity. The samples were analyzed using the inductively coupled plasma mass spectrometry (ICP-MS) technique for 53 elements. The objective of this research is to evaluate the geochemical relationship among the three study sites by using Bayesian modeling strategies to determine the geochemical risk. Specifically, non-parametric and parametric models are employed to determine potential similarities related to land use and the geology of the region under study. The results obtained will be a significant contribution to the knowledge of the territory and will help in making decisions associated with land use.

Daniel Kjellevold	Rolandsgard	Using marked Cox point processes for turbulence statistics on free surface fluid mechanics

For free surface fluid flow, one of the main descriptive properties is the surface divergence, which is an important factor in the heat-mass transfer between the ocean and the atmosphere. However, the surface divergence is not easily observed. Vortex imprints, or simply vortices, are visually distinct points on the free surface of a fluid, which can be easily observed from a video of the free surface, and it has been demonstrated that there is a dependency between vortices and surface divergence. This motivates the choice of using such vortices as a proxy for surface divergence. In this work, we model vertices with a Bayesian log Gaussian Cox marked point process model, where the centers of the observed vortices are the points and the size of each vortex as the marks. The analyses is based on a case study from a numerically simulated dataset, of a turbulent water flow in a square tank with a free surface with 12500 time steps. We consider snapshots in time for our analysis, focusing on the spatial structure with repeated observation, in order to draw conclusions about the ability learn surface divergence from vortices observations.

Riccardo Rossetti	An approximate message passing algorithm for the matrix tensor product model

We propose and analyze an approximate message passing (AMP) algorithm for the matrix tensor product model, which is a generalization of the standard spiked matrix models that allows for multiple types of pairwise observations over a collection of latent variables. A key innovation for this algorithm is a method for optimally weighing and combining multiple estimates in each iteration. Building upon the approach of Berthier et al. (2020), we prove a state evolution for non-separable functions that provides an asymptotically exact description of its performance in the high-dimensional limit. We leverage this state evolution result to provide necessary and sufficient conditions for recovery of the signal of interest. Such conditions depend on the singular values of a linear operator derived from an appropriate generalization of a signal-to-noise ratio for our model. We show that, under a wide range of priors, an appropriately tuned version of our algorithm is capable of achieving Bayes-optimal performance under squared error loss.

Diego Martínez Salmerón	On integral priors and posteriors for multiple comparison in Bayesian model selection

It is well known that noninformative priors for estimation usually are not appropriate for model selection. The methodology of Integral Priors for the comparison of two models has been developed to get appropriate priors for Bayesian model selection modifying the initial noninformative priors, and it has been successfully applied in many situations. We propose a generalization of the methodology for $q\geq 2$ models. This approach adds a copy of each model constraining the parametric space to a compact set in such a way that i) we obtain the Integral Prior for each of the $2q$ models as the proper priors to which certain Markov chains converge at a geometric rate, and ii) the initial noninformative priors for the copied models are Integral Priors. We study the behavior of the methodology as the compact parametric space goes to the original one. We also propose new numerical procedures to approximate the Bayes factors for Integral Priors, and a Metropolis-Hastings algorithm to simulate from the Integral Posteriors. The methodology is illustrated with several examples.

Nilotpal Sanyal	Variable selection and estimation using nonlocal prior mixtures for data with highly dispersed effect sizes

Different nonlocal priors are seen to provide better support for smaller and larger effect sizes. Motivated by this observation, we focus on mixtures of nonlocal priors that should provide a better fit compared to the individual priors for sparse data where effect sizes vary considerably in magnitude, such as in public health data with biological, clinical, and environmental features. Particularly, we examine the use of a mixture prior which is a sum of a point mass and two nonlocal prior components. In addition to considering existing alternatives, we propose a novel objective way to set the tuning parameters of the nonlocal priors so as to allow maximum support to diverse effect sizes. Following a full Bayes approach, we develop a joint selection and estimation procedure based on Gibb's sampling and Laplace approximation. Variable selection was A preliminary data analysis based on sensible hyperpriors shows a competitive performance of our method when compared to established methods like LASSO and horseshoe prior analyses. We demonstrate the utility of our method through real-data applications.

Partha Sarkar	High-Dimensional Bernstein Von-Mises Theorems for Covariance and Precision Matrices

This paper aims to examine the characteristics of the posterior distribution of covariance/precision matrices in a ``large p, large n'' scenario, where p represents the number of variables and n is the sample size. Our analysis focuses on establishing asymptotic normality of the posterior distribution of the entire covariance/precision matrices under specific growth restrictions on p_n and other mild assumptions. In particular, the limiting distribution turns out to be a symmetric matrix variate normal distribution whose parameters depend on the maximum likelihood estimate. Our results hold for a wide class of prior distributions which includes standard choices used by practitioners. Next, we consider Gaussian graphical models which induce sparsity in the precision matrix. Asymptotic normality of the corresponding posterior distribution is established under mild assumptions on the prior and true data-generating mechanism.

Terrance Savitsky, Luca Sartore, Lu Chen	A Privatized Bayesian Hierarchical Copula Framework for the Synthetic Multivariate Mixed-Type Data Generation

When statistical agencies consider the release of synthetic record-level data to the public, simulation algorithms are used as powerful tools for maintaining the anonymity of the survey and census respondents. However, the variables used to generate synthetic data are usually mixed types and not independent. Copulas are

typically used to elicit the multivariate covariance structure under mixed-type data, and they are less often applied to generate multivariate synthetic data. Therefore, a privatized Bayesian hierarchical copula model is proposed to generate multivariate mixed-type data. The risk-weighted pseudo-posterior mechanism guarantees asymptotic differential privacy (aDP) under the copula formulation. The pseudo posterior mechanism maintains privacy by downweighting the likelihood contributions of records using weights inversely proportional to their disclosure risks. Disclosure risks are formulated as a function of the individual log-likelihood contributions, so that high disclosure risks are associated with less likely observations. Therefore, the distribution of synthetic data will exhibit good utility properties and have minimal distortion when compared to the distribution of private data. This paper provides a demonstration of how to generate multivariate synthetic data that are aDP guaranteed. An application based on the data collected by the United State Department of Agriculture's National Agricultural Statistics Service for the U.S. Census of Agriculture is presented, and the results are discussed.

Maximilian Scholz	Stabilizing Bayesian Simulations with Inverse Forward-Sampling

The power of generative models is a cornerstone of Bayesian data analysis, allowing for predictive simulations and model checking. However, in practice the difficulties of specifying independent priors can lead to unreasonable simulation outcomes, a common obstacle for full Bayesian simulations. In this, we present inverse forward-sampling, a technique to stabilize Bayesian simulations by use of precondition. While the underlying concepts are not novel themselves, we formalize the idea and show its utility for practitioners in different modeling scenarios.

Christine Shen	Target matched and average discrepancy p-values for composite null models

We investigate issues of model checking in Bayesian contexts, with a focus on situations where it is desirable to use parameter-dependent "statistics" to measure model and data compatibility. A commonly utilized approach is the posterior predictive p-value (Gelman et el, 1996). However, the distribution of such p-value deviates significantly from uniformity in both finite samples and the asymptotic setting due to double use of data. Other suggested methods for computing p-values in such situations include the calibrated p-value (Hjort et el, 2006), posterior twice predictive p-value (Wang and Xu 2021), and split predictive checks (Li and Huggins

2022, Moran et el, 2022). However, these alternatives exhibit drawbacks such as high computational costs and inefficient data utilization. We propose two alternatives which strike a balance between computational efficiency and uniformity, the target-matched p-value and the average discrepancy p-value. We demonstrate the advantages of our proposed methods from both Bayesian and frequentist perspectives through simulations and real-world data examples.

Ami Sheth (Presenter); Aaron Smith; Andrew Holbrook	Sparse Bayesian multidimensional scaling

Bayesian multidimensional scaling (BMDS) is a probabilistic dimension reduction tool that allows one to model and visualize data consisting of dissimilarities between pairs of objects. Whereas BMDS has proven useful within, e.g., Bayesian phylogenetic inference, its log-likelihood and log-likelihood gradient calculations require a burdensome O(N^2) floating-point operations, where N is the number of data points. Thus, BMDS becomes impractical as N grows large. We propose a sparse version of BMDS (sBMDS) that applies log-likelihood and gradient computations to a subset of the observed data. Through simulations, we demonstrate that these subsets are sufficient to recover the objects' latent locations with reasonable accuracy. To obtain these latent location estimates, we implement three different MCMC algorithms: 1) HMC under the full model, 2) HMC under the sBMDS model and 3) surrogate-trajectory HMC using the sBMDS gradient (surrogate-trajectory sBMDS-HMC). We compare the computational efficiency across these various method and observe, e.g., 10-fold speedups when we apply sBMDS to 100 data points; this speedup only increases with the size of the data considered. Finally, we apply sBMDS and surrogate-trajectory sBMDS-HMC to the phylogeographic modeling of multiple influenza subtypes to better understand how these strains spread through global air transportation networks.

Giovani Silva	Joint modeling of longitudinal binary responses: A nonparametric Bayesian approach using acute and chronic malnutrition data

In the analysis of longitudinal data, Gaussian random effects are usually adopted to deal with the individual dependence over time. We question that Gaussian assumption especially for analyzing longitudinal binary response variables that are natural dependent and jointly distributed according to Bayesian Nonparametric

(BNP) random effects. That is, conditionally on common random effects, the two longitudinal binary response variables are considered independent, facilitating the construction of the corresponding likelihood function and therefore the inference on the model parameters. Nonparametric and semiparametric approaches for statistical data analysis have the advantage of not imposing a given distribution for either responses or common random effects. For a longitudinal binary two outcomes, we assume Bernouilli sampling model and BNP random effects for liking the two binary responses. Regarding BNP models, Dirichlet process, Pitman-Yor process, and some of their variations are performed. These models were implemented in R package NIMBLE, evaluating costs and benefits on computational time and fitting BNP models with pre-clustering of the patients based on similarities matrices. Markov Chain Monte Carlo (MCMC) methods have been also employed for inferential purposes of some parameters such as regression coefficients.Finally, our motivation is due to a binary data study from a longitudinal four-arm randomized parallel trial conducted in Bengo Angolan province to explore the effects of four interventions on wasting and stunting outcomes. A total of 121 children with intestinal parasitic infections were allocated to the four arms. This is joint work with Andre Nunes and Luzia Goncalves.

Luca Alessandro Silva	Robust leave-one-out cross-validation for high-dimensional Bayesian models

Leave-one-out cross-validation (LOO-CV) is a popular method for estimating out-of-sample predictive accuracy. However, computing LOO-CV criteria can be computationally expensive due to the need to fit the model multiple times. In the Bayesian context, importance sampling provides a possible solution but classical approaches can easily produce estimators whose asymptotic variance is infinite, making them potentially unreliable. Here we propose and analyze a novel mixture estimator to compute Bayesian LOO-CV criteria. Our method retains the simplicity and computational convenience of classical approaches, while guaranteeing finite asymptotic variance of the resulting estimators. Both theoretical and numerical results are provided to illustrate the improved robustness and efficiency. The computational benefits are particularly significant in high-dimensional problems, allowing to perform Bayesian LOO-CV for a broader range of models, and datasets with highly influential observations. The proposed methodology is easily implementable in standard probabilistic programming software and has a computational cost roughly equivalent to fitting the original model once.

Mike K.P. So	Graphical copula GARCH modeling with dynamic conditional dependence

The aim is to develop a graphical copula GARCH model for volatility modeling. To allow high-dimensional modeling for large portfolios, the complexity of the modeling is greatly reduced by introducing conditional independence among stocks given the market risk factors, such as the S&P500 index in the United States. The market risk factors are modeled using a directed acyclic graph (DAG) model with a pairwise-copula construction to allow flexible distributional modeling. The use of the DAG model gives a topological order to the market risk factors, which can be regarded as a list of directions of the flow of information or disturbance. The conditional distributions among stock returns are also modeled through pairwise-copula constructions for flexibility. We adopt dynamic conditional dependence structures to allow the parameters in the copulas to be time-varying such that we can model dynamically the tail dependence between any two stocks. Three-stage estimation is used for estimating parameters in the marginal distributions, the copulas of the DAG of the market risk factors, and the copulas of the stocks. Bayesian inference is used to learn the structure of the DAG. In the simulation study, we show that these estimation procedures can be used to recover the parameters and the DAG accurately. With Bayesian inference, we can allow the structure of the market risk factors to be random, and model averaging can be done to obtain robust predictions of volatility.

Aldo Solari (joint with Jelle J. Goeman)	Bonferroni-Bayes: A Fully Bayesian Argument For Multiplicity-Adjustment of Credible Intervals

While the Bayesian and frequentist philosophies are markedly different, the methods derived from these philosophies can be surprisingly similar. For example, Bayesian and frequentist methods can be identical when non-informative priors are chosen, and may converge asymptotically otherwise. It has been argued that differences between the two styles of inference are negligible enough in practice, and that Bayesian credible intervals and frequentist confidence intervals may be used interchangeably.

However, Bayesian and frequentist practice is certainly not in agreement everywhere. One major difference is in the way Bayesian and frequentist methods deal with multiplicity, i.e., the situation in which inference is done on more than one parameter simultaneously.

Here, the frequentist consensus is that multiplicity increases error rates, and should therefore be accounted for. A large literature of multiple testing methods exists to deal with this problem. The larger part of this literature deals with multiplicity in hypothesis testing, but there are also many methods adjusting confidence intervals for multiplicity, usually by using increased confidence levels. One such method is a Bonferroni correction, in which confidence intervals are calculated at level $1-\alpha/m$ rather than at level $1-\alpha$ if m confidence intervals are presented. This method is easy to use, but can be inefficient if the confidence intervals are correlated.

In the Bayesian literature the dominant view is that there is no need for multiplicity correction. There is, however, a dissident stream of thought that does argue for multiplicity correction in the Bayesian framework. The Bayesian arguments for or against multiplicity correction tend to be formulated in the context of Bayesian hypothesis testing. The impact, or lack of it, of multiplicity on credible intervals is seldom considered.

In this contribution we look at the impact of multiplicity of parameters on Bayesian credible intervals. We will adopt a fully Bayesian perspective, focusing on credibility only and avoiding consideration of frequentist error rates. Still, we will find that multiplicity impacts credibility intervals in much the same way as it impacts confidence intervals: multiplicity negatively impacts the credibility of credible intervals, and this lack of credibility may be repaired by increasing the credibility level of each of the intervals. We even suggest a context in which a Bonferroni adjustment of credible intervals would be appropriate.

Ingelin Steinsland — Bayesian spatial model for rutting

Pavement maintenance is essential to ensure road safety and prevent excessive depreciation of roads and it is one of the major causes why pavements need asphalt resurfacing. Many road agencies or road maintenance contractors map their road network regularly and In Norway, all state roads are scanned every year since 2010. The Norwegian Pavement Management System can provide automatically generated reports, for example which sections exceed the maximum allowed rut depth defined in the maintenance standards.

However, deeper diagnostic analysis to identify areas that exhibit unexpectedly high rutting rates still requires manual labor. Such type of data analysis is typically only initiated based on suspicion and are very time consuming. Our working hypothesis

has been that advanced Bayesian analysis of the available data can provide a better understanding of how the pavement condition of a whole network develops and allow automated identification of segments that perform better or worse than expected. Some explanatory variables such as annual average daily traffic, road width, rut depth from the previous year and pavement type are available, while others such as weather and construction quality are not. This call for a Bayesian spatial model with the explanatory variables and both annual spatial effects (due to e.g. the spatial dependency in weather) and a common spatial effect for all years (e.g. due to construction). The model is set up and evaluated for eleven years (2010-2020) of pavement rutting data for one road stretch; the European highway route E14, stretching 67.1 kilometres from Stjørdal, Norway to Storlien on the Swedish border. We identify areas with unexplained or accelerated rutting. Further, when comparing to a model without spatial effects, the effect of the explanatory variables are considerable different.

Tyrel Stokes	A generative approach to frame-level multi-competitor races

Multi-competitor races often feature complicated within race strategies that are difficult to capture and may confound inferences and predictions when training data on race outcome level data. In this work we develop a general generative Bayesian model for multi-competitor races which allows analysts to explicitly model certain strategic effects such as changing lanes or drafting and separate these impacts from competitor ability. The generative model allows one to simulate full races using the posterior predictive distribution from any starting real or created starting position which opens new avenues for attributing value to within race actions and to perform counter-factual analyses. We apply this methodology to one-mile horse races using data provided by the New York Racing Association (NYRA) and the New York Thoroughbred Horsemen's Association (NYTHA) for the Big Data Derby 2022 Kaggle Competition. This data features granular tracking data for all horses at the frame-level (occurring at approximately 4hz). We demonstrate how this model can yield new inferences, such as the estimation of horse-specific speed profiles which vary over phases of the race, and examples of posterior predictive counterfactual simulations to answer questions of interest.

Bingjing Tang	Tractable Gaussian Cox Processes

A Gaussian Cox process is a popular Bayesian nonparametric method for modeling point process data, in which the intensity function is a transformation of a Gaussian process. Posterior inference of this intensity function involves an intractable integral in the likelihood resulting in doubly intractable posterior distribution. Here, we propose an approach for estimating the inhomogeneous intensity function without reliance on large data augmentation or approximations of the likelihood function. We propose to jointly model the intensity and its integral as a truncated Gaussian process, allowing us to directly bypass the integral calculation in the likelihood. We propose an efficient MCMC sampler for posterior inference and show that it is faster than competing approaches on a number of synthetic and real datasets. Finally, we discuss extensions of our proposed method to other inhomogeneous point processes.

The Tien Mai	A reduced-rank approach to predicting multiple binary responses through machine learning

This paper investigates the problem of simultaneously predicting multiple binary responses by utilizing a shared set of covariates. Our approach incorporates machine learning techniques for binary classification, without making assumptions about the underlying observations. Instead, our focus lies on a group of predictors, aiming to identify the one that minimizes prediction error. Unlike previous studies that primarily address estimation error, we directly analyze the prediction error of our method using PAC-Bayesian bounds techniques. In this paper, we introduce a pseudo-Bayesian approach capable of handling incomplete response data. Our strategy is efficiently implemented using the Langevin Monte Carlo method. Through simulation studies and a practical application using real data, we demonstrate the effectiveness of our proposed method, producing comparable or sometimes superior results compared to the current state-of-the-art method.

Giovanni Toto	Local Bayesian clustering for functional data

We propose a Bayesian approach to perform indirect local clustering of functional data through B-splines basis expansion and a dependent random partition model for the basis coefficients. We exploit the local property of B-splines to obtain partially coincident functions when sharing clusters for enough contiguous basis parameters. To allow for a sequential evolution of the partitions, we introduce a novel dependent random partition model inducing sequences of random partitions exhibiting semi-Markovian dependence. By matching the order of the B-spline with that of the

semi-Markovian dependence, the dependent random partition model can be used as a prior distribution in a functional model which fully exploits the local property of B-splines. Additionally, we demonstrate that our novel prior can be effectively applied to any context where the central objective is the direct modeling of a sequence of dependent partitions for which Markovian assumptions may be too restrictive, such as in the context of time series exhibiting higher-order autoregressive dependence.

Gilad Turok	Delayed rejection generalized Hamiltonian Monte Carlo for multiscale densities

Multiscale distributions, with varying curvature in their log density functions, cannot be effectively preconditioned, and thus are often impractical or even impossible to sample with standard Euclidean Hamiltonian Monte Carlo (HMC) methods such as the no-U-turn sampler (NUTS). Neal (2020) showed how generalized Hamiltonian Monte Carlo (G-HMC), with single steps with partial momentum refresh, could make HMC-like directed movement by coupling a non-reversible, sawtooth-shaped acceptance probability update. Hoffman and Sountsov (2022) showed how to automatically adapt a step size and metric for Neal's approach using an ensemble of parallel chains. Modi et al. (2023) showed that delayed rejection HMC (DR-HMC) effectively samples multiscale distributions by retrying rejected trajectories at smaller step sizes. In this note, we provide evaluations to demonstrate that delayed rejection generalized HMC (DR-G-HMC) provides efficient, directed HMC-like movement, handles multiscale distributions, and can be adapted in a parallel ensemble. Joint work with Gilad Turok, Chirag Modi, Edward Roualdes, and Alex Barnett.

Lampis Tzai	Bayesian MANOVA for the combined evaluation of handwriting evidence

Forensic science is a broad field that uses scientific principles and technical methods to help with the evaluation of evidence in legal proceedings of criminal, civil, or administrative nature. Forensic scientists examine recovered traces that can be given by glass fragments, fingerprints, body fluids, textile fibers, digital device data and handwriting. The handwriting examination is a well-known field of analysis. Consider a case involving a handwritten document of unknown origin. Handwritten features extracted from this questioned document will be compared to those extracted from a document written by a person that is suspected of being at the origin of the anonymous document. The propositions of interest are the following:

Hp: the suspect is the author of the manuscript;
Hd: the suspect is not the author of the manuscript.

Handwriting individualization is still largely dependent on the experience of the document examiner, though studies have been conducted with the aim of supporting handwriting examiners to reduce the degree of subjectivity of their expertise. Marquis et al. (2005) proposed to increase the degree of objectivity of handwriting analyses by implementing elements of Fourier analysis in order to describe the contour shape of loops of characters. Specifically, the characters that contain loops can be described by means of Fourier coefficients, which can be used to characterize the shape complexity and other geometric attributes. The analyses conducted showed that these features have a good discriminating power.

With the aim of implementing the use of these handwriting features for handwriting identification, Bozza et al. (2008) proposed a Bayesian probabilistic approach by modeling the data with multivariate Normal- inverse-Wishart distribution (NIW). The value of the evidence is subsequently assessed by means of the Bayes factor, which can be interpreted as a measure of the strength of support provided by the evidence in favor of the hypothesis Hp against the hypothesis Hd. This approach was accomplished to take into account the correlation between variables, the variability between-writers and within-writer variability. However, the above model is implemented separately for each different type of handwritten character. This can be problematic, as different conclusions can be obtained for each character examined.

To reduce the impact of this model constraint, it is proposed the implementation of a Bayesian Multivariate Analysis of Variance (MANOVA) via using the loop characters as predictors. The indicator of the loop character is transformed into a dummy variable (corner-point representation), so that it is possible to model variables describing the handwriting characters jointly taking into account the variability between characters, the variability between-writers for every character and within-writer variability. The Bayesian MANOVA is compared with the two-level random effect model (NIW) proposed by Bozza et al. (2008). Three different methods for estimating the marginal likelihoods are used; the Generalized Harmonic Mean, the Laplace-Metropolis and the Bridge Sampling. Furthermore, the performance of each of the two models under consideration are compared with those of an

alternative closed form approach (i.e. a conjugate approach). This conjugate approach does not allow to model the within-writer and between-writers variability, separately, but the marginals can be obtained analytically.

Firstly, the Bayes factor (BF) which compares the two models under consideration is calculated for each writer separately in order to determine which of the two competing model performs better. Secondly, there have been selected handwriting features originating from the same writer or from different writers to evaluate the rate of false negatives (that is cases where the BF is smaller than one for characters originating from the same source) and false positives (that is cases where the BF is greater than one for characters originating from different sources).

Finally, the sensitivity of models is examined under two critical issues: the impact of the available background information to inform model parameters and the impact of the choice of the number of the degrees of freedom of the Wishart-inverse distribution that is used to model handwriting variability. Concerning the first aspect, the prior distribution is elicited by selecting different sets of writers based on their similarity or dissimilarity of their loop features.

Antoine Van Biesbroeck	Generalized mutual information for prior elicitation

Based on the mutual information criterion, the reference prior theory proposes a non informative prior whose choice can be called objective. We propose an enrichment of the theory by suggesting a generalization of the mutual information definition, arguing our interpretation on an analogy with Global Sensitivity Analysis. Some of our generalized metrics are studied to derive their associated generalized reference priors, under the consideration of different dissimilarity measures such as the f-divergence in a first hand, and of different sub-classes of prior in a second hand. Our work reinforces the robustness of the Jeffreys' prior consideration within such bayesian modeling. Our results are available on arXiv (arXiv:2310.10530).

Cristiano Villa

The construction of objective priors is, at best, challenging for multidimensional parameter spaces. A common practice is to assume independence and set up the joint prior as the product of marginal distributions obtained via “standard" objective methods, such as Jeffreys or reference priors. However, the assumption of

independence a priori is not always reasonable, and whether it can be viewed as strictly objective is still open to discussion. In this paper, by extending a previously proposed objective approach based on scoring rules for the one dimensional case, we propose a novel objective prior for multidimensional parameter spaces which yields a dependence structure. The proposed prior has the appealing property of being proper and does not depend on the chosen model; only on the parameter space considered.

Cecilia Viscardi	Approximate Bayesian Computation for Probabilistic Damage Identification

Structural monitoring and damage detection analyses are fundamental to guarantee the safety of civil structures. In the civil engineering field, they are often formalised as inverse problems whose solution is found by minimising a distance measure between experimental data and data produced by a Finite Element Model (FEM). This approach ignores any source of uncertainty, such as unobserved system characteristics, measurement errors and variations of the material properties. In this framework, a probabilistic approach is essential to be aware of the uncertainty around estimates and predictions and to conduct conscious decision-making processes. However, the inclusion of several sources of uncertainty often leads to complex statistical models with an intractable likelihood function. To address this problem, we propose a fully Bayesian approach based on Approximate Bayesian Computation (ABC) algorithms. The presented ABC method overcomes problems related to the computational cost of FEMs by exploiting their deterministic nature and relying on an emulator based on Artificial Neural Networks (ANNs). We test the proposed method in a simulation study on a damaged beam. Finally, we show some preliminary results from a real-world example of a masonry tower.

Tomoya Wakayama	Similarity-based Random Partition Distribution for Clustering Functional Data

Random partitioned distribution is a powerful tool for model-based clustering. However, the implementation in practice can be challenging for functional spatial data such as hourly observed population data observed in each region. The reason is that high dimensionality tends to yield excess clusters, and spatial dependencies are challenging to represent with a simple random partition distribution (e.g., the Dirichlet process). This paper addresses these issues by extending the generalized Dirichlet process to incorporate pairwise similarity information, which we call the

similarity-based generalized Dirichlet process (SGDP), and provides theoretical justification for this approach. We apply SGDP to hourly population data observed in 500m meshes in Tokyo, and demonstrate its usefulness for functional clustering by taking account of spatial information.

Richard Warr	Bayesian Mixture Models for Histograms: with Applications to Large Datasets

It is not uncommon for privacy or summarization purposes to receive data in a table or in histogram format with bins and associated frequencies. In this work we present a method that fits a mixture distribution to model the probability density function of the underlying population. We focus on a mixture of normal distributions, however the method could be generalized to mixtures of other distributions. A prior is placed on the number of mixture components which could be finite or countably infinite and inference is obtained using reversible jump MCMC. We demonstrate attractive properties of the method, which show a great deal of promise to modeling large data problems using a Bayesian nonparametric approach. Additionally, we consider the case of multiple histograms and cluster them using the Dirichlet process. This clustering allows for the sharing of information between populations and provides a posterior probability of homogeneity between populations.

Ganchao Wei	Flow Matching from a Bayesian Decision Theoretic Perspective

Flow matching (FM) is a family of training algorithms for fitting continuous normalizing flows (CNFs), a class of ODE-based generative models for complex densities. In contrast to diffusion models, whose training relies on simulating the full path from data to noise, training a CNF using FM does not require simulating the so-called flow path, and hence enjoys computational efficiency and is often referred to as ``simulation-free". A CNF is specified by a corresponding marginal vector field along with the latent noise distribution from which a target data distribution is generated through applying continuous transforms specified by the vector field. The key idea behind the standard approach to FM, referred to as conditional flow matching (CFM), is that the marginal vector field of a CNF can be learned through the much easier task of fitting least-square regression to the so-called conditional vector field specified given one or both ends of the flow path. While traditional justification of CFM is the matching of the gradient for the CFM regression problem and that for the original problem of fitting the marginal vector field under L_2 loss, we show that CFM training can be readily justified from a Bayesian decision theoretic perspective of parameter estimation under squared error loss whether or not the gradients match.

Inspired from this perspective we introduce two extensions of CFM training that broaden the applicability of CNFs to generative modeling of time series and to learning conditional densities. To the first goal, we introduce a new algorithm called Gaussian process CFM (GP-CFM), which fits regression to the conditional vector field given not only the end points but the whole flow path, which is endowed with a Gaussian process (GP) "prior" (or marginal distribution). The distributional properties of GPs allow drawing "posterior" (or conditional) samples of the conditional vector field to be achieved without simulating the full paths, maintaining the key ``simulation-free'' feature of CFM training. Moreover, adopting the GP on the flow paths also allows the integration of intermediate observations and additional prior information all while maintaining analytical tractability and the simulation-free nature. We apply CNFs tackle nonparametric density regression (i.e., learning covariate-dependent densities) through incorporating the conditioning on covariates into the CFM fitting to the conditional vector field. We provide two examples from the neuroscience application involving local field potential recordings of mouse.

Manuel Wiesenfarth, Annette Kopp-Schneider and Silvia Calderazzo	Quantification of prior impact in terms of effective sample size targeting test decisions

Bayesian methods allow borrowing historical information through prior distributions but domination of the prior information on posterior inference is commonly unwanted. To quantify and communicate prior informativeness, the prior effective sample sizes (ESS) has gained widespread use, equating information provided by a prior to a sample size. The common approach equates the prior ESS to the number of samples in a (hypothetical) historical data set. However, this measure is independent from the newly observed data, and does not capture potential negative impact on posterior inference induced by the prior in case of prior-data conflict.

In previous work (1), we built on Reimherr et al (2) to relate prior information to a number of (virtual) samples from the current data model. Thereby, the number of samples needed to obtain posterior inference under a baseline prior equivalent to that under the prior of interest is quantified. To this end, an inferential target measure of interest needs to be defined and an MSE type measure was investigated.
In this work, we argue that an impact measure should be tailored to the inferential question of interest which is often a test decision in the context of clinical trials. The integrated risk under the 0-k loss (e.g. 3) is a possible target measure in this situation. We illustrate that tailoring the effective sample size to the test decision problem

provides desirable behavior, for example by showing a decreasing impact of an informative prior in case the data is clearly in the null or alternative hypothesis. In this case the data may be evidently conclusive such that the prior does not play any role for the test decision.

Furthermore, we illustrate how such ESS measure can be helpful in interpreting and selecting mixture weights in robust mixture priors (4) which adaptively discard prior information in case of prior-data conflict. It is well known that the rate at which information is discarded depends on both the mixture weight and the dispersion of the vague component. This makes interpretation of the mixture weight difficult creating needs for an intuitive impact measure for such priors.

An R implementation for the beta-binomial and normal-normal model is provided.

1 Wiesenfarth, M., & Calderazzo, S. (2020). Quantification of prior impact in terms of effective current sample size. Biometrics, 76(1), 326-336.
2 Reimherr, M., Meng, X. L., & Nicolae, D. L. (2021). Prior sample size extensions for assessing prior impact and prior-likelihood discordance. Journal of the Royal Statistical Society Series B: Statistical Methodology, 83(3), 413-437.
3 Calderazzo, S., Wiesenfarth, M., & Kopp-Schneider, A. (2022). A decision-theoretic approach to Bayesian clinical trial design and evaluation of robustness to prior-data conflict. Biostatistics, 23(1), 328-344.
4 Schmidli, H., Gsteiger, S., Roychoudhury, S., O'Hagan, A., Spiegelhalter, D., & Neuenschwander, B. (2014). Robust meta-analytic-predictive priors in clinical trials with historical control information. Biometrics, 70(4), 1023-1032.

Myung Won Lee

Bayesian Regularisation with Smoothing for Tail Index Regression

Extreme events can be better comprehended through the lens of regression models tailored for extreme values. We propose a novel approach for modelling the extreme values via a Bayesian regularisation with smoothing for a tail index regression framework. Our method is based on a conditional Pareto-type specification which is regularised with Bayesian Lasso-type shrinkage priors and smoothed with low-rank thin plate splines. To validate the performance of the proposed method through a simulation study that recovers the true tail index, α(x) over a variety of scenarios along with regularising the covariates. Our illustration continues on analysing extreme wildfires in Portugal and identifying its drivers via our proposed method.

Stephanie Wu	Survey-weighted Bayesian clustering methods for characterizing dietary acculturation among older Mexican-born individuals living in the United States

Dietary acculturation is the process by which immigrants adopt the dietary habits of a host country and is an important factor to consider when creating culturally competent health promotion programs. For instance, it influences the extent to which traditional foods should be emphasized in healthy eating patterns. Dietary acculturation is challenging to quantify, both in terms of differences in food intake and the underlying factors driving the process, because both are latent constructs that cannot be measured directly. Current approaches use single-item or investigator-derived acculturation metrics that may not adequately capture this multidimensional process or are limited by sample size constraints of the research study. To address this, we propose using a data-driven model-based clustering framework with triangulation to measure the two latent constructs using large-scale survey data. We develop Bayesian clustering methods with the ability to: 1) characterize latent patterns and account for acculturation-linked heterogeneity using a robust profile clustering formulation, 2) derive the number of patterns using a sparsity-inducing Dirichlet prior, and 3) accommodating survey data through a weighted Gibbs sampler and a post-processing variance correction. We validate the models using a simulation study and examine dietary acculturation among older Mexican-born individuals living in the United States using data from the Hispanic Community Health Study/Study of Latinos.

Jay Xu	Bayesian Inference for Log-Binomial Regression Models with Missing Covariate and Outcome Values

In clinical and epidemiological research, the relative risk is a widely used quantity to characterize the association between an exposure or risk factor on a dichotomous outcome, and it is often preferred to the odds ratio because it is easier to interpret and communicate to others. Relative risks are typically modeled using a generalized linear model with a log link, commonly referred to as a log-binomial model. The coefficients of a log-binomial model are the log adjusted relative risks, and as such, relative risk estimation comprises fitting a log-binomial model to data. Missing data are ubiquitous in clinical and epidemiological datasets, and failure to properly address missing data can result in biased estimation and misleading conclusions. Sullivan et al. (2017) considered a simulation scenario where a dichotomous outcome and two Gaussian covariates are generated from a log-binomial model where the

relative risks were known and the outcome and one covariate was subject to missingness under a missing at random mechanism. Using simulations, Sullivan et al. (2017) explored the use of a Fully Conditional Specification (FCS) multiple imputation strategy, using a mis-specified logistic regression model and a mis-specified linear regression model to impute the missing outcome and covariate values, respectively. Sullivan et al. (2017) found that this approach resulted in biased estimation. To date, no one has successfully proposed a Frequentist or Bayesian estimation strategy for this simulation scenario. Here, I propose a successful estimation strategy for this simulation scenario, using a clever Metropolis-Hastings (MH) algorithm that never generates proposal values for the unobserved covariate values that fall outside their support given the parameter values from the most recent iteration. I implement the MH algorithm using the Julia programming language, and I demonstrate that it achieves desirable frequentist properties (e.g., coverage) for the estimation of the log-binomial parameter values.

Qingzhao Yu	High-Dimensional Bayesian Mediation Analysis with Adaptive Laplace Priors

The mediation analysis method is used to investigate effects of mediators that intervene in the pathways between an exposure variable and an outcome variable. Bayesian methods are naturally used in mediation analysis due to the hierarchical structure of Bayesian models. This paper introduces an innovative adaptive Bayesian mediation analysis method that incorporates adaptive Laplace priors into the predictive model to account for high-dimensional mediators. This approach introduces a penalization function on the estimated direct and indirect effects rather than solely on the coefficients of predictive models. Consequently, estimated effects that lack statistical significance may shrink to zero, facilitating a more robust analysis. We demonstrate the efficacy of our adaptive mediation analysis method on simulations and on a Louisiana triple negative breast cancer (TNBC) dataset to examine racial disparity in diagnosed stage among TNBC patients diagnosed between 2010 and 2017. The dataset is linked to the 2017 hazardous air pollutant emissions burden estimation database using patients' residential address. We effectively explain a portion of the disparity using currently collected variables. The analysis identifies crucial mediators and confounders, highlighting the significance of variables such as age of diagnosis, insurance status, tumor grades, and the concentration of Naphtha in the air.

Zhaoxi Zhang	The underlap coefficient as measure of a biomarker's discriminatory ability in a multi-class disease setting

The first step when evaluating a diagnostic test is to determine the variation in its values across different disease groups. In the three-class disease setting, the volume under the receiver characteristic surface (VUS) and the three-class Youden index (YI) are the commonly used summary measures of a test's discriminatory ability. However, these measures are only appropriate under a stochastic ordering assumption for the distributions of test outcomes in the three groups. This assumption is stringent, not always plausible, particularly when covariates are involved. Violating this assumption can lead to incorrect conclusions about a test's performance to distinguish between the three classes. To address this, we propose the underlap coefficient, study its properties, as well as its relationship with the VUS and YI when a stochastic order is enforced. We further propose Bayesian nonparametric estimators for both the unconditional underlap coefficient and for its covariate-specific version. A simulation study reveals a good performance of the proposed estimators across a range of conceivable scenarios. We also illustrate the proposed approach with an application to an Alzheimer's disease (AD) dataset to assess how different potential AD biomarkers distinguish between individuals with normal cognition, mild impairment, and dementia, and how age and gender impact this discriminatory ability.

Yichen Zhu	Radial Neighbors for Provably Accurate Scalable Approximations of Gaussian Processes

In geostatistical problems with massive sample size, Gaussian processes can be approximated using sparse directed acyclic graphs to achieve scalable $O(n)$ computational complexity. In these models, data at each location are typically assumed conditionally dependent on a small set of parents which usually include a subset of the nearest neighbors. These methodologies often exhibit excellent empirical performance, but the lack of theoretical validation leads to unclear guidance in specifying the underlying graphical model and sensitivity to graph choice. We address these issues by introducing radial neighbors Gaussian processes (RadGP), a class of Gaussian processes based on directed acyclic graphs in which directed edges connect every location to all of its neighbors within a predetermined radius.

We prove that any radial neighbors Gaussian process can accurately approximate the corresponding unrestricted Gaussian process in Wasserstein-2 distance, with an

error rate determined by the approximation radius, the spatial covariance function, and the spatial dispersion of samples.

We offer further empirical validation of our approach via applications on simulated and real world data showing excellent performance in both prior and posterior approximations to the original Gaussian process.

Alessandro Zito	Bayesian nonparametric modeling of latent partitions via Stirling-gamma priors

Dirichlet process mixtures are particularly sensitive to the value of the so-called precision parameter, which controls the behavior of the underlying latent partition. Randomization of the precision through a prior distribution is a common solution, which leads to more robust inferential procedures. However, existing prior choices do not allow for transparent elicitation, due to the lack of analytical results. We introduce and investigate a novel prior for the Dirichlet process precision, the Stirling-gamma distribution. We study the distributional properties of the induced random partition, with an emphasis on the number of clusters. Our theoretical investigation clarifies the reasons of the improved robustness properties of the proposed prior. Moreover, we show that, under specific choices of its hyperparameters, the Stirling-gamma distribution is conjugate to the random partition of a Dirichlet process. We illustrate with an ecological application the usefulness of our approach for the detection of communities of ant workers.

Alex Ziyu Jiang (presenter), Abel Rodriguez	Flexible excitation modeling for Multivariate Hawkes Processes based on Dependent Dirichlet process

Multivariate Hawkes Processes (MHPs) capture the complex dynamics among temporal event sequences on multiple dimensions. While strong parametric assumptions are usually made about the excitation functions of MHP, the excitation patterns in real-world applications tend to be flexible. Further, they tend to show strong similarities despite being from different dimensions. Motivated by the reasons above, we propose MHP-DDP (MHP based on dependent Dirichlet process), a novel hierarchical nonparametric Bayesian modeling framework for MHP. MHP-DDP flexibly estimates the excitation function via a mixture of scaled Beta distributions, and borrows strengths across dimensions by modeling such mixing distribution as a

mixture of a shared Dirichlet process (DP) and a group-specific idiosyncratic DP. We develop two algorithms using Markov chain Monte Carlo (MCMC) and the stochastic variational inference (SVI) algorithm. We study the estimation and prediction performance via simulations and show that MHP-DDP outperforms the benchmark methods in terms of lower estimation error for both algorithms, with SVI being computationally efficient than MCMC. Finally, we apply our model to study the risk dynamics in the Standard & Poor's 500 intraday index prices among its 11 sectors.