ISBA

2018 ISBA World Meeting

Edinburgh, UK

June 24-29, 2018

ABSTRACTS

Oral Presentations

Sponsored by:

Office of Naval Research
USN
ONR
Science & Technology

RStudio

STATA®

Google

ASA Section on Bayesian Statistical Science

The Alan Turing Institute

CONVENTION EDINBURGH

ICMS
International Centre for Mathematical Sciences

Abraham, Kweku
Posterior contraction rates for diffusions with high-frequency data.
We investigate Bayesian methods for scalar stochastic diffusions with high-frequency data (where asymptotics are taken as the increments shrink but maximal time horizon goes to infinity). We find priors for which the posterior contracts at the (minimax) optimal rate around the true drift parameter, hence guaranteeing that estimators based on the posterior have good worst-case performance. Contraction rates have previously been found for low frequency data (where the increments are fixed), so these results demonstrate that the Bayesian method unifies these two sampling schemes.

Aitken, Colin
Discussant
Since the late 1970's, Prof. Aitken has been one of the leading researchers in Bayesian methods in forensic science and related areas. He has lead the recent effort by the Royal Statistical Society in providing statistical guidance to the judiciary, published numerous papers on computational and foundational methods for forensic statistics, and co-authored the main text books that are used in the training of new forensic statisticians. He is considered one of the founders of the forensic statistics sub-discipline.
Prof . Aitken will discuss the three proposed presentations-
1- Inference about relationships from DNA mixtures
2- Approximate Bayesian Computational and Kernel Based methods for extending the Bayesian paradigm for forensic evidence to high dimensional settings
3-The Interplay between Bayesian and Frequentist Methods for the Presentation and Interpretation of Evidence in a Forensic Laboratory Setting.
prof . Aitken will discuss the three proposed presentations-
1- Inference about relationships from DNA mixtures
2- Approximate Bayesian Computational and Kernel Based methods for extending the Bayesian paradigm for forensic evidence to high dimensional settings
3-The Interplay between Bayesian and Frequentist Methods for the Presentation and Interpretation of Evidence in a Forensic Laboratory Setting

Alaimo Di Loro, Pierfrancesco
Bayesian Population Size Estimation with A Single Sample
The estimation of the size of a population is a problem encountered in a variety of applications. One standard statistical approach relies on capture-recapture sampling, which requires the population to be sampled at least twice. he necessity of further capture occasions leads to increasing costs and, furthermore, can cause an ever-increasing annoyance to the population of interest. These considerations have motivated the search for procedures that allow to estimate the size of a population from a single capture occasion rather than repeated occasions. Hettiarachchige [Hettiarachchige C.K.H., PhD Thesis, University of Melbourne, http://hdl.handle.net/11343/118636, (2016)] proposes a method that is viable when the population is made of only two generations: a group of generators and one of generated units. We investigate Bayesian methods alternative to the frequentist estimators used by the original author. Preliminary results, based on simulation studies, give evidence of competing performance of the Bayesian approach, which in some cases sensibly outperforms the frequentist alternatives. Moreover, the flexibility of the Bayesian approach allows for the development of some promising extensions that may improve even more on the model performances.

Albert, Jim

A Bayesian Redesign of the First Probability/Mathematical Statistics Course

The traditional calculus-based introduction to statistical inference consists of a semester of probability followed by a semester of frequentist inference. Cobb (2015) challenges the statistical education community to rethink the undergraduate statistics curriculum. In particular, he suggests that we should focus on two goals: making fundamental concepts accessible and minimizing prerequisites to research. Using five underlying principles of Cobb, we describe a new calculus-based introduction to statistics based on simulation-based Bayesian computation.

Alberto Fuquene Patino, Jairo

On choosing mixture components via non-local priors

Choosing the number of components in mixture models remains a central but elusive challenge. Traditional model selection criteria can be either overly liberal or conservative when enforcing parsimony. They may also result in poorly separated components of limited practical use. Non-local priors (NLPs) are a family of distributions that encourage parsimony by enforcing a separation between the models under consideration.

We formalize NLPs in the context of mixtures and show how they lead to well-separated components that have non-negligible weight, hence interpretable as distinct subpopulations. We derive tractable expressions and suggest default prior settings, for instance aimed at detecting multi-modal densities in Normal and T mixtures. We also give a theoretical characterization of the sparsity induced by NLPs and propose easily implementable algorithms to obtain the integrated likelihood and parameter estimates. Although the framework is generic we fully develop multivariate Normal, Binomial and product Binomial mixtures based on a novel family of exchangeable moment priors. The theory and underlying principles hold more generally, however. Our results show a serious lack of sensitivity of the Bayesian information criterion (BIC), which is ameliorated but still present in the singular BIC, and insufficient parsimony of the AIC and local prior counterparts to our formulation. We also offer comparisons to overfitted and repulsive overfitted mixtures.

Alexander, Monica

A Bayesian hierarchical model for estimating subnational female populations in developing countries.

Accurate estimates of subnational-level populations are important for policy formulation and monitoring key population health indicators. In particular, estimates of the number of women of reproductive age affect measures of maternal mortality, contraceptive prevalence and fertility. However, in many developing countries, data on population counts are limited and are of poor quality, and so levels are unclear. We present a Bayesian hierarchical model to estimate female populations at the subnational level. The model incorporates available data on population counts and migration, and builds on characteristic mortality schedules to obtain robust population estimates and uncertainty levels. The model is applied to estimate and predict populations by county in Kenya for 1979-2020.

Alston-Knox, Clair

Search and annoy: Using Bayesian analysis to inform better surveillance of the red Fireant

As an applied statistician, some of the projects we work on evolve over a period of many years. These projects often involve humble beginnings, where knowledge of the subject and expected outcomes is vague. As time passes, knowledge grows, involving information from deliberately recorded data through to anecdotal experience of various team members. Using a Bayesian approach to reflect this evolution of knowledge is a very natural and beneficial modelling strategy. This talk presents a case study of the invasive red fireant in Brisbane, Australia. Here, the aim was to provide a suitability of habitat indicator to assist in the allocation of surveillance resources. Although the war against fireants had been 10 years in the making at the beginning of this project, the objective of eradication left the project team with very scant "data", with only location and year of find available. The useful data was contained in the collective experience of the team. In this work, Bayesian analysis was a useful tool for both using expert knowledge as a form of data and as a convenient computational tool. We will show how the outcomes of the Bayesian analysis can provide technical staff and managers in very different fields of research with an appropriate understanding of probability and risk, which in turn can aid in their decision-making process.

Anacleto, Osvaldo

A stochastic transmission model to estimate social genetic effects in disease transmission

Current stochastic epidemic models ignore genetic heterogeneity in infectivity, which is the propensity of an infected individual to transmit diseases. Variation in this social interaction trait leads to the common superspreading phenomenon, where a minority of highly infected hosts transmit the majority of infections. To date, is not known whether infectivity is genetically controlled. We present a novel stochastic transmission model which, by combining individual-level Poisson processes with bivariate random effects, can fully capture genetic variation in infectivity. Not only can this Bayesian model accurately estimate heritable variation in both infectivity and the propensity to be infected, but it also can identify parents more likely to generate offspring that are disease superspreaders. We will also present how a Bayesian analysis of a recent large-scale fish experiment reveals, for the very first time, that genetics does indeed contribute to variation in infectivity and therefore affects the spread of diseases.

Argiento, Raffaele

A hierarchical nonparametric approach for robust graphical modelling in omics data

Useful tools to express multivariate network structures in gene expression studies are graphical models. However, alternative models are needed when data are strongly overdispersed. An interesting proposal has been introduced by Finegold and Drton (2014) who used a Dirichlet process to cluster data-components and accommodate for overdispersion. In this work, we consider a more general class of nonparametric distributions, namely the class of normalised completely random measures (NormCRM's), which yields a more flexible component clustering. Moreover, in order to borrow information across the data, we model the dependence among the NormCRM's through a nonparametric hierarchical structure. At data level, each NormCRM is centred on the same base measure, which is a NormCRM itself. The discreteness of the shared base measure implies that the processes at data level share the same atoms. This desired feature allows to cluster together components of different data. We will compare the performances of the proposed model with competitors via a simulation study. Moreover, we will explore genomic expression patterns in the yeast Saccharomyces cerevisiae responding to diverse environmental transitions. We will identify the multivariate network structure of the data and meanwhile cluster components according their degree of over dispersion.

Bacallado, Sergio

Adaptive designs for Clinical Trials that Accept New Arms

The majority of randomized oncology trials are two-arm studies that test the efficacy of new therapies against a standard of care, thereby assigning a large proportion of patients to non-experimental therapies. In contrast, multi-arm studies efficiently share a common control arm while evaluating multiple experimental therapies. A major bottleneck for traditional multi-arm trials is the
requirement that all therapies - often drugs from different companies - have to be available at the same time when the trial starts. New drugs are rarely at the same stage of development and traditional multi-arm designs may delay the clinical evaluation of new treatments. These limitations motivate our study study of statistical methods for adding new experimental arms after a clinical trial started enrolling patients. We consider both balanced and Bayesian outcome-adaptive randomization for experimental designs that allow investigators to add new arms, discuss their application in a tuberculosis trial, and evaluate the proposed experimental designs using a set of realistic simulation scenarios.

Bakka, Haakon

Non-stationary Gaussian models with physical barriers

When modeling spatial data near the coast, we need to consider which assumptions to make on the Gaussian field with respect to the coastline, i.e. what kind of boundary effect to assume. One possibility is to have no boundary effect, modeling both water and land, but with observation and prediction locations only in water, leading to a model with a stationary correlation structure. However, a stationary field smooths over islands and peninsulas, inappropriately assuming that observations on two sides of land are highly correlated. Other approaches in the literature range from simple use of Dirichlet or Neumann boundary conditions, to being quite complex and costly. In this talk we showcase a new approach, the Barrier model, implemented in R-INLA, that is intuitive in the way correlation follows the coastline, and is as easy to set up and do inference with as a stationary field, with computational complexity O(n sqrt(n)). We compare this model to two others, showing significant improvement at reconstructing a test function. A real data application shows that the Barrier model smooths around peninsulas, and that inference is numerically stable. We also detail how the stochastic partial differential equations (SPDE) approach was used to construct the Barrier model.

Balocchi, Cecilia

Prior Choice and Ensemble Particle Optimization for Clustering Areal Data

Bayesian hierarchical models of spatial areal data utilize priors that induce spatial smoothing across geographically proximate regions. However, it is often interesting to find distinct clusters of areal units that exhibit markedly different trends. To express our uncertainty about this unknown latent clustering, we must specify a prior distribution over the space of partitions that respects the spatial proximity of the areal units. In this work, we explore the behavior induced by different choices of prior distributions that generate spatial partitions. Rather than directly sampling from the posterior distribution, we introduce an ensemble particle-based optimization procedure that targets the partitions with largest posterior probability. We run greedy searches over the posterior distribution of partitions where the particles are made ``mutually aware'' through an entropy penalty that repels particles that appear headed to the same point. We demonstrate our prior exploration and implementation method with simulated data and also a case study on estimating the change in crime frequencies over time in the city of Philadelphia.

Banerjee, Sayantan

A Bayesian Integrated Graphical Model Framework for Personalized Network Estimation in Cancer

The functional cancer genomic and proteomic data provide rich sources of information to identify variations in signaling pathways and activities within and across tumor lineages. However, current analytic methods lack the ability to exploit the diverse and layered architecture of biological networks to provide coherent metrics for inferring pathway-based activities that can be used for both global cancer-specific and local patient-specific stratification and outcome prediction. We propose personalized cancer-specific integrated network estimation (PRECISE) -- a general graphical model framework, to estimate cancer-specific integrated networks, infer patient-specific networks and elicit interpretable pathway-level signatures, by integrating prior information on existing interaction databases, data-driven de novo causal structures, and upstream molecular profiling data. Using the pan-cancer functional proteomic data on 32 cancer types from The Cancer Genome Atlas, we demonstrate the utility of PRECISE in inferring commonalities and differences in network biology across tumor lineages and in using patient-specific pathway-based signatures for robust tumor stratification and prediction.

Banerjee, Sudipto

On Directed Acyclic Graphical Models for Spatial Disease Mapping

Hierarchical models for regionally aggregated disease incidence data commonly involve region-specific latent random effects which are modelled jointly as having a multivariate Gaussian distribution. The covariance or precision matrix incorporates the spatial dependence between the regions. Common choices for the precision matrix include the widely used intrinsic conditional autoregressive model which is singular, and its nonsingular extension which lacks interpretability. We propose a new parametric model for the precision matrix based on a directed acyclic graph representation of the spatial dependence. Our model guarantees positive definiteness and, hence, in addition to being a valid prior for regional spatially correlated random effects, can also directly model the outcome from dependent data like images and networks. Theoretical and empirical results ensure easy interpretation of model parameters. Our precision matrix is sparse and the model is highly scalable for large datasets. We also derive a novel order-free version which remedies the dependence of directed acyclic graphs on the ordering of the regions by averaging over all possible orderings. The resulting precision matrix is still sparse and available in closed form. We demonstrate the superior performance of our models using simulation experiments and a public health application.

Banks, David

New Ideas in Auctions

Adversarial Risk Analysis allows one to pose a new class of auction problems for n >= 3 players. Instead of the usual common knowledge assumption that enables the classic Bayes Nash equilibrium solution, it is perfectly reasonable to suppose that two opponents have different beliefs about the distribution of a third opponent's bid. This situation has not been solved within the standard game theory framework. This talk describes two ways to address that problem from an ARA perspective.

Barrientos, Andreas

A Bayesian nonparametric procedure for comparing several treatments against a control.

We propose a flexible strategy to test for differences between a control group and several treatment regimes. While the existing tests for this type of comparison are based on the differences between location parameters, our approach identifies differences across the entire distributions. The procedure is developed under a Bayesian nonparametric approach, which avoids strong modeling assumptions over the distributions of each treatment or population. As a by-product, we make inferences on the shift function, which allows us to quantify the differences between two populations considering not only the location but the entire distribution. The procedure's performance is illustrated using simulated and genuine data.

Berry, Don

Whats New and Transformative in a Phase III Clinical Trial?

Bayesian ideas are revolutionizing clinical trials. Some trials are frankly Bayesian. Others benefit from Bayesian contributions to modeling and to using various types of adaptations during the course of the trial. I will focus on Bayesian phase III trials. In the parlance of the U.S. FDA a phase III trial is adequate and well controlled and serves as the basis for marketing approval. I will motivate this focus using a phase II trial called I-SPY 2 that in 2016 was the subject of four articles in a single issue of the New England Journal of Medicine. The following are whats new and transformative in a phase III clinical trials, including having all ten characteristics present in the same trial.

1. Include many treatment arms that enter the randomized trial, are evaluated, and leave the trial (which I call a platform trial)
2. Seamless shift in focus for each arm, from learn stage to confirm stage of the same trial
3. All patients assigned to an arm in both stages count toward registration
4. Common control arm (by patient subtype)
5. Comparisons with all controls via modeling by time machine, even non-concurrently randomized controls
6. Identify arms indications (subsets of responsive patients), if any, including as defined by biomarkers
7. Learning system with continuous updating a la Bayes
8. Interpretation of Type I error
9. Adaptive randomization using Bayesian probabilities in learn stage
10. The trials adaptive aspects are driven by Bayesian predictive probabilities

I will describe some of these characteristics and show how they play out in the trial as a whole.

Berry, Lindsay

Bayesian forecasting of high-dimensional count-valued time series

Modeling and forecasting of high-dimensional time series of non-negative counts is a common interest for many retailers who rely on forecasts for inventory management, production planning, and marketing decisions. Motivated by the field of product demand forecasting, this paper presents the dynamic count mixture model (DCMM) to flexibly and efficiently model time series of counts through a mixture of Bernoulli and Poisson dynamic generalized linear models. The sequential learning and forecasting of the DCMM allows fast, parallel analysis in high-dimensional forecasting frameworks. We extend the DCMM to over-dispersed counts through the addition of a random effect in the conditionally Poisson model. Given the time constraints in product demand forecasting, many forecasters run univariate models independently across products. However, forecasting accuracy may be improved by pooling information across groups of related series when estimating trends or seasonal components. We present an efficient multiscale framework, which incorporates cross series linkages while insulating the parallel estimation of the DCMM. We apply this framework to a case study of supermarket products aimed at forecasting the daily demand 1-14 days in the future.

Best, Nicky

Borrowing Relevant External Data using Bayesian Priors to Inform Regulatory Decision-Making for Small Clinical Trials; Using Prior Elicitation to Support Decision Making in Drug Development

Rare diseases and paediatric populations present several challenges and opportunities for clinical trial design and analysis, due to practical and ethical constraints on sample size. Significant challenges also exist in other settings where recruiting patients is difficult due to ethical, logistical and patient burden issues. In this presentation we consider trials where the feasible sample size that can be recruited has low power to demonstrate efficacy using a conventional design with strong control of type 1 error at the one-sided 2.5\% level. We will discuss the use of Bayesian designs using informative priors based on relevant external data to increase the power and precision of such trials, and propose a range of operating characteristics based on clinical trial simulation that can be useful to evaluate and compare designs. We consider the weighted average probability of an error (type I or type II) as a metric to calibrate trial designs in this setting. We also present various prior and posterior summaries of the available historical, current and total evidence which can be very helpful for sponsors and regulators. Methods will be illustrated in both paediatric settings borrowing efficacy data from adults, and in confirmatory trials in adults borrowing historical controls.

Best, Nicky

Using Prior Elicitation to Support Decision Making in Drug Development

As the use of Bayesian approaches continues to expand in the field of drug development, there is a need for statisticians to have tools to develop robust and defensible informative prior distributions. Whilst relevant empirical data should, where possible, provide the evidential basis for such priors, it is often the case that limitations in data and/or our understanding may preclude direct construction of a data-driven prior. Even in cases where a substantial body of empirical evidence is available, there is very often a translational gap between the setting(s) to which the available data relate and the new setting under consideration. In such cases, statisticians at GlaxoSmithKline (GSK) have been using prior elicitation techniques to enable quantification of existing knowledge in the absence of directly relevant data. In this talk, I will give a flavor of what the prior elicitation process involves and discuss how the elicited priors have been used at GSK, e.g. to help predict probability of success of next study(s) at key milestone decision points, to explore staged development activities and to determine the merits of interim/futility assessments. I will also discuss some of the challenges we have experienced, including the pros and cons of aggregating priors from several experts versus retaining the individual priors, and how to manage the tendency for over-optimism that is inherent in many experts priors. The talk will be illustrated using various case studies covering all stages of clinical drug development.

Bhattacharyya, Anwesha

A Bayesian methodology for high dimensional discrete graphical models

This work introduces a bayesian methodology for fitting large discrete graphical models (Ising and Potts models) that uses the pseudo-likelihood instead of the likelihood. This is a powerful relaxation which allows node-wise parallel computation under separable priors. In high dimensional graphical models we can expect some sparsity in the network. The natural approach to encode this sparsity is to introduce spike-and-slab priors. We propose the use of a weak spike and slab prior in the form of a Gaussian spike for such models. Standard MCMC algorithms can be used for the purpose of evaluating resulting posterior distribution. The use of a bayesian approach allows us to leverage existing prior information and implement variable selection and estimation simultaneously. We illustrate the method through several examples and real data.

Bianchi, Daniele

Large-scale dynamic predictive regressions

We develop a novel "decouple-recouple" dynamic predictive strategy and contribute to the literature on forecasting and economic decision making in a data-rich environment. Under this framework, clusters of predictors generate different latent states in the form of predictive densities that are later synthesized within an implied time-varying latent factor model. As a result, the latent inter-dependencies across predictive densities and biases are sequentially learned and corrected. Unlike sparse modeling and variable selection procedures, we do not assume a priori that there is a given subset of active predictors, which characterize the predictive density of a quantity of interest. We test our procedure by investigating the predictive content of a large set of financial ratios and macroeconomic variables on both the equity premium across different industries and the inflation rate in the U.S., two contexts of topical interest in finance and macroeconomics. We find that our predictive synthesis framework generates both statistically and economically significant out-of-sample benefits while maintaining interpretability of the forecasting variables. In addition, the main empirical results highlight that our proposed framework outperforms both LASSO-type shrinkage regressions, factor based dimension reduction, sequential variable selection, and equal-weighted linear pooling methodologies.

Bierkens, Joris

Zig-Zag Sampling for Doubly Intractable Distributions

In many important models in Bayesian statistics the computation of the likelihood function is intractable. The corresponding posterior distributions are referred to as doubly intractable distributions. For example this situation occurs when inferring the temperature in an Ising model, or the parameters in an exponential random graph model (ERGM). Existing methodology to deal with such models rely on augmented variable algorithms of Moller (2006) and variations thereof (Murray et al, 2006). In order for this class of algorithms to be asymptotically exact it is necessary to draw perfect samples from the forward model using e.g. the Propp & Wilson (1996) methodology. It turns out that, when applying the Zig-Zag Sampler (Bierkens et al., 2016) to problems with (doubly) intractable likelihood, it also suffices to obtain perfect samples from the forward model. This introduces a completely new algorithm to this class of statistical models. In this talk we will introduce the Zig-Zag sampler and discuss its application to doubly intractable distributions, after which we compare the performance of the Zig-Zag Sampler with the existing methodology. This is currently work in progress in collaboration with Antonietta Mira (Lugano) and Gareth Roberts (Warwick).

Bijak, Jakub

Bayesian Agent-Based Population Studies: Uncertainty quantification for complex migration modelling.

International migration is one of the most uncertain and complex demographic processes, and most of the traditional analytical approaches fail to address this uncertainty and complexity in a comprehensive manner. In the talk, we will outline a research programme aimed at developing a simulation model of migration, based on a population of intelligent, cognitive agents, their social networks and institutions, all interacting with one another. The Bayesian statistical principles and uncertainty quantification techniques will be used to design computer experiments, and later to learn about modelling the simulated agents and the way they make decisions. By developing micro-foundations for migration studies, we hope that the proposed approach will help effectively integrate behavioural and social theory with formal modelling. The theoretical and conceptual discussion will be illustrated by a prototype agent-based model of migration, analysed by using Gaussian process emulators.

Bochkina, Natalia

Geometry and robustness of Bayesian inference for misspecified nonregular ill-posed models

We consider a broad class of statistical models that can be misspecified and ill-posed, from a Bayesian perspective. This provides a flexible and interpretable framework for their analysis, but it is important to understand robustness of the chosen Bayesian model and its effect on the resulting solution, especially in the ill-posed case where in the absence of prior information the solution is not unique.Compared to earlier work about the Bernstein-von Mises theorem for nonregular well-posed Bayesian models, we show that non-identifiable part of the likelihood, together with the constraints on the parameter space, introduce a more complex geometric structure of the posterior distribution around the best reconstruction point in the limit, and provide a local approximation of the posterior distribution in this neighbourhood.The results apply to misspecified models which allows, for instance, to evaluate the effect of model approximation on statistical inference.Emission tomography is taken as a canonical example for study, but our results hold for a wider class of generalised linear inverseproblems with constraints.This is joint work with Peter Green (University of Bristol, UK and UTS, Australia).

Bonomolo, Paolo

Detecting Liquidity Traps

In this paper we develop an econometric strategy to compute the probability that an economy is in a liquidity trap, that is a situation where the interest rate is at the effective lower bound and inflation fluctuates around a very low or negative equilibrium. The specification is a generalization of the Steady-State VAR (Villani, 2009), and it is an unobserved component model where the observed variables are decomposed in a short run component, described by a VAR, and a long run component treated as a state space model with Markov Switching regimes. As in Villani (2009), we estimate the VAR using additional information coming from the economic theory about the long run equilibrium values with the additional advantage that we can specify a complete model for the long run, instead of simply using priors for the values of the steady states.With this straightforward generalization we introduce drifting parameters in VARs (the long run equilibrium) avoiding the usual dimensionality problem (see Cogley and Sargent, 2005, and Primiceri, 2005).In our econometric strategy we explicitly consider the presence of a lower bound on interest rates, and we estimate the VAR taking into account this additional source on non-linearity.Combining a model for the long run with the VAR is what allows us to: (i) "detect" the liquidity trap, that is to recognize that an economy has a high probability of being in a liquidity trap quite in advance, before the interest rate becomes bounded; (ii) improve the forecast of inflation and interest rate. We apply this methodology to Japan, US and Euro Area.This is joint with Yildiz Akkaya and Ingvar Strid.

Bouchard-Côté, Alexandre

Piecewise deterministic Markov chain Monte Carlo methods

I will give an overview of some recent work on piecewise deterministic Markov chain Monte Carlo methods, an emerging complementary approach to standard reversible discrete time MCMC methods. In one direction, we are looking at piecewise deterministic processes with non-linear dynamics, which build on the Bouncy Particle Sampler (BPS) [1] but can exploit information about the target such as Normal approximations [2]. In another direction, we investigate the geometric ergodicity of BPS and extensions [3], which in turn motivates methodological innovations. In a third direction,we look at discretizations of the BPS which admit black-box implementations [2]. This is joint work with G. Deligiannidis, A. Doucet, P. Vanetti, S. Vollmer. [1] https://arxiv.org/abs/1510.02451 [2] https://arxiv.org/abs/1707.05296 [3] https://arxiv.org/abs/1705.04579

Bradley, Jonathan

Hierarchical Models with Conditionally Conjugate Full-Conditional Distributions for Dependent Data from the Natural Exponential Family

We introduce a Bayesian approach for analyzing (possibly) high-dimensional dependent data that are distributed according to a member from the natural exponential family of distributions. This problem requires extensive methodological advancements, as jointly modeling high-dimensional dependent data leads to the so-called "big n problem." The computational complexity of the "big n problem" is further exacerbated when allowing for non-Gaussian data models, as is the case here. Thus, we develop new computationally efficient distribution theory for this setting. In particular, we introduce something we call the "conjugate multivariate distribution," which is motivated by the univariate distribution introduced in Diaconis and Ylvisaker (1979). Furthermore, we provide substantial theoretical and methodological development including: results regarding conditional distributions, an asymptotic relationship with the multivariate normal distribution, conjugate prior distributions, and full-conditional distributions for a Gibbs sampler. The results in this manuscript are extremely general, and can be adapted to many different settings. We demonstrate the proposed methodology through simulated examples and several real data analyses,

Broderick, Tamara

Coresets for fast, streaming, distributed Bayesian inference

Bayesian methods are attactive for large-scale data analyses due to their rich modeling properties, uncertainty quantification, and incorporation of prior knowledge. But in practice, many approximate Bayesian inference methods either fail to scale to large data sizes or cannot provide theoretical guarantees on the quality of the approximate posterior in the finite data regime. We leverage the insight that data is often redundant to instead obtain a weighted subset of the data (called a coreset) that is much smaller than the original dataset. We can then input this coreset into existing posterior inference algorithms at much lower computational cost. In this work, we develop an efficient coreset construction algorithm for Bayesian inference. We provide theoretical guarantees on the size and approximation quality of the coreset. The proposed approach further permits efficient construction of the coreset in both streaming and parallel settings, with minimal with minimal additional effort. We demonstrate the efficacy of our approach on a number of synthetic and real-world datasets.

Brynjarsdottir, Jenny

Towards accounting for model discrepancy in CO2 retrievals from the OCO-2 satellite.

The Orbiting Carbon Observatory 2 (OCO-2) collects space-based measurements of atmospheric CO2. The CO2 measurements are indirect as the instrument observes radiances (reflected sunlight) over a range of wavelengths and a physical model is inverted, via Bayes Theorem, to estimate CO2 concentration in the atmosphere. This inference is in fact an estimation of physical parameters, aka calibration, which can be both biased and over-confident when model error is present but not accounted for. The OCO-2 mission addresses this problem in a few different ways, e.g. with a post-inference bias correction procedure based on ground measurements. This talk will discuss methods to account for structured and informative model error directly in the inversion procedure to lessen bias and provide more reliable uncertainty estimates. This is joint work with Jonathan Hobbs and Amy Braverman at JPL.

Cadonna, Annalisia

Spectral Density Estimation for Multiple Time Series

The spectral density function contains information about the second order behavior of a signal and can be interpreted as the distribution of the power over the frequency range. The problem of estimating the spectral density function arises naturally in fields where informationabout frequency behavior is relevant and several signals are recorded concurrently. For example, multichannel electroencephalography (EEG) records measurements of electrical potential fluctuations at multiple locations on the scalp of a subject. I will present a hierarchical Bayesian modeling approach to spectral density estimation for multiple time series, where the log-periodogram of each series is modeled as a mixture of Gaussian distributions with frequency-dependent weights and mean functions. The implied model for each log-spectral density is a mixture of mean functions with frequency-dependent weights. In addition to accommodating flexible spectral density shapes, a practically important feature of the proposed formulation is that it allows for ready posterior simulation through a Gibbs sampler with closed form full conditional distributions for all model parameters. I will show results for multichannel electroencephalographic recordings, which provide the key motivating application for the proposed methodology. I will thenpresent some extensions for non-stationary time series.

Caiado, Camila

Iterating through experiments: Finding tipping points in post-operative cardiac surgery

Joint work with Michael Goldstein, Jordan Oakley and Benjamin Lopez.Over 1000 UK patients die per year of complications after cardiac surgery. Potentially preventable complications, including renal and respiratory failure, often become established after early signs of deterioration are missed. Existing risk prediction scores calculate risk retrospectively by analysing a fraction of the available data and comparing observations with population-derived “normal” ranges. Here we describe a series of studies, with progressively increasing detail, that aim to provide effective risk monitoring. The underlying sequence of Bayesian analyses provides live updates of risk that are tailored appropriately to individuals, and provide a timely decision support tool.Using data from over 3100 cardiac surgery patients at the University Hospital of South Manchester, our Bayesian models: continuously analyse each patient’s data, identify abnormal patterns by comparison with the individual’s own previous values, and build on population values to give a prior assessment exploiting exchangeability arguments. This assessment can be further individualized to the patient by including other proxies for risk such as the EuroSCORE (pre-operative risk for cardiac surgery), as well as the surgeon’s judgements about the patient and operation. Here we focus renal failure, a complication that affects 12% of patients after cardiac surgery. This approach found a tipping point that manifests itself as a critical transition in physiological time series (e.g. urine output, pH, blood pressure)

Calder, Catherine (Kate)

Bayesian Spatial Clustering of Adolescent Activity Patterns Across an Urban Environment

Research on neighborhood effects often focuses on linking features of social contexts or exposures to health, educational, and criminological outcomes. Traditionally, individuals are assigned a specific neighborhood, frequently operationalized by the census tract of residence, which may not contain the locations of routine activities. In order to better characterize the many social contexts to which adolescents are exposed, the Adolescent Health and Development in Context (AHDC) Study collected GPS-based space-time trajectories of approximately 1,400 adolescents in Columbus, OH over two one-week periods. In this presentation, we propose a novel statistical model for clustering individuals in the study based on similarities in their activity trajectories. The approach takes the two-mode network consisting of individuals and stable activity location and identifies communities in this network consisting of individuals whose activity locations are proximate in space and follow similar temporal patterning. Our approach allows us to identify communities of adolescent, which we compare to those traditionally used in neighborhood effects research.

Campbell, Trevor

Bayesian Coreset Construction via Greedy Iterative Geodesic Ascent

Coherent uncertainty quantification is a key strength of Bayesian methods. But modern algorithms for approximate Bayesian posterior inference often sacrifice accurate posterior uncertainty estimation in the pursuit of scalability. This work shows that previous Bayesian coreset construction algorithms---which build a small, weighted subset of the data that approximates the full dataset---are no exception. We demonstrate that these algorithms scale the coreset log-likelihood suboptimally, resulting in underestimated posterior uncertainty. To address this shortcoming, we develop greedy iterative geodesic ascent (GIGA), a novel algorithm for Bayesian coreset construction that scales the coreset log-likelihood optimally. GIGA provides geometric decay in posterior approximation error as a function of coreset size, and maintains the fast running time of its predecessors. We validate GIGA on both synthetic and real datasets, demonstrating that it reduces posterior approximation error by orders of magnitude compared with previous coreset constructions.

Carvalho, Carlos

On the Long Run Volatility of Stocks

In this paper we investigate whether or not the conventional wisdom that the volatility per period of stocks is lower over longer horizons. Taking the perspective of an investor, we evaluate the predictive variance of k-period returns for different models and prior specifications. We adopt the state space modeling framework of Pastor and Stambaugh [2012] as we feel it is the best alternativefor capturing all important expected returns dynamics and accounting for the associated uncertainties. Part of the development includes an extension of the modeling framework to incorporate time-varying volatilities and covariances in a constrained prior information set up. We conclude, that in the U.S. market, there are plausible prior specifications such that stocks are indeed less volatile in the long run. Our results are supported by model assessment exercises that provideevidence to our assumptions. To understand the generality of the results, we extend our analysis to a number of international equity indices. This is joint work with Hedibert Lopes and Robert McCulloch.

Carvalho, Luis

Bayesian Network Regularized Regression for Crime Modeling

We present a new methodology for functional network regression on node attributes using a Laplacian operator based on edge similarities as regularizer. We show how usual regularization penalties can be cast as prior distributions on regression coefficients under a Bayesian setup, and propose a computationally efficient EM fitting procedure. We discuss a specific application to modeling residential burglary in Boston using a hierarchical model with latent indicators for "hot zones" and a conditional zero-inflated negative binomial regression for crime rates.

Castillo, Ismael

Adaptive Polya trees: rates and inference

We consider modifications of the standard Polya tree construction that allow to obtain optimal adaptive posterior convergence rates, in particular in terms of the supremum norm loss, for estimating a density function. We also discuss properties of the corresponding posteriors for inference on functionals of the unknown density. This is joint work with Romain Mismer (Paris).

Chen, Ming-Hui

Bayesian Inference for Network Meta-Regression Using Multivariate Random Effects with Applications to Cholesterol Lowering Drugs

Low-density lipoprotein cholesterol (LDL-C) has been identified as a causative factor for atherosclerosis and related coronary heart disease, and as the main target for cholesterol-lowering and lipid-lowering therapy. Many clinical trials have been carried out on safety and efficacy evaluation of cholesterol lowering drugs. To synthesize the results from different clinical trials, we examine treatment level (aggregate) network meta-data from 29 double-blind, randomized, active or placebo-controlled statins +/- Ezetimibe clinical trials on adult treatment-naÃ¯ve patients with primary hypercholesterolemia. In this paper, we propose a new approach to carry out Bayesian inference for arm-based network meta-regression. Specifically, we group the variances of the random effects based on the clinical nature of treatments, and the determination of the number of groups and group membership is further guided by Bayesian model comparison criteria. The proposed approach is especially useful when some treatment arms are involved in only a single trial. In addition, a new Metropolis-within-Gibbs sampling algorithm is developed to carry out the posterior computations. In particular, the correlation matrix is generated from its full conditional distribution via partial correlations. The proposed methodology is further applied to analyze the network meta-data from 29 trials with 11 treatment arms.

Chen, Qixuan

Use of Bayesian spatial growth mixture models in complex survey data

The use of growth mixture models has gained a lot of favor in psychological research in the past decade. However, the literature on the utility of growth mixture models in complex survey data is sparse. This work was motivated by the Galveston Bay Recovery Study (GBRS), a three-wave panel survey aiming to characterizing trajectories and determinants of post-traumatic stress disorder (PTSD) severity score among residents living in Galveston and Chalmers counties in Texas, USA, after Hurricane Ike stuck the Galveston Bay area in September, 2008. We proposed a Bayesian spatial growth mixture model, which separates the population into multiple sub-groups with different post-disaster trajectories of PTSD and examines risk factors associated with different latent groups. The survey design information was incorporated into the model by allowing the mean PTSD scores in each wave to vary by primary sampling units (PSU) and by including the sampling weight as a covariate in predicting both the PTSD score and the latent class membership. To account for spatial correlation, the model further supplements the independent normal random effects of the PSU with spatial terms in predicting the latent class membership.

Chen, Su

Fast Bayesian Variable Selection for the Linear Model

This paper presents a method for fast Bayesian variable selection in the normal linear regression model with ultrahigh dimensional data. A novel approach is adopted in which an explicit posterior probability for including a covariate is obtained. The method is sequential, one deals with each covariate one by one, and a spike and slab prior is only assigned to the coefficient under investigation. We adopt the well-known spike and slab Gaussian priors with a sample size dependent variance, which achieves strong selection consistency for marginal posterior probabilities even when the number of covariates grows almost exponentially as sample size. Numerical illustrations are used to demonstrate the idea where it is shown that this new approach provides essentially equivalent results to the standard spike and slab priors, i.e. same marginal posterior probability, estimated via Gibbs sampling. Hence, we obtain the same results via the direct calculation of p probabilities, as compared with a stochastic search over a space of 2^p elements. Our procedure only require p probabilities to be calculated, which can be done with exact expressions, thus allows parallel computation when p is large.

Chib, Siddhartha

Bayesian Estimation and Comparison of Moment Condition Models

Bayesian analysis of semiparametric models in which moment conditions supply the only source of information about the unknown distribution of the data has long been a challenge. A fundamental component of the Bayesian architecture, namely the likelihood, appears to be missing. In this work we utilize the exponentially tilted empirical likelihood (ETEL) to unify Bayesian thinking about such problems. Because we allow for over restricted moment conditions, a key concern of the work is the analysis of misspecified moment condition models. We show that even under misspecification, the Bayesian ETEL posterior distribution satisfies the Bernstein - von Mises (BvM) theorem. We also develop a unified theory based on marginal likelihoods (computed by the method of \cite{Chib1995} as extended to Metropolis-Hastings samplers in \cite{ChibJeliazkov2001}) for comparing different moment restricted models. We establish the large-sample model selection consistency of the marginal likelihood and show that the marginal likelihood favors the model with the minimum number of parameters and the maximum number of valid moment restrictions. When the models are misspecified, the marginal likelihood procedure selects the model that is closer to the true data generating process in terms of the Kullback-Leibler divergence. We provide illustrations of some of the many possible applications made possible by this work.

Chiu-Hsing Weng, Ruby

Real-time Bayesian parameter estimation for item response models - with application to Internet ratings data

Bayesian item response models have been used in modelling educational testing and Internet ratings data. Typically, the statistical analysis is carried out using Markov Chain Monte Carlo methods. However, these may not be computationally feasible when real-time data continuously arrive and online parameter estimation is needed. We develop an efficient algorithm based on the Woodroofe-Stein identity to adjust the parameters in real-time. The proposed online algorithm works well for two real datasets, achieving good accuracy but with considerably less computational time.

Choi, Taeryon

On Bayesian semiparametric hierarchical models for longitudinal data analysis

In this work, we propose semiparametric Bayesian hierarchical mixed effects models for analyzing either longitudinal data or clustered data. The hierarchical semiparametric model allows to capture similarities across different groups while each group having distinct smooth curve representation. In the semiparametric mixed effects model structure, we estimate nonparametric smoothing functions of continuous covariates by using a spectral representation of Gaussian processes and the subject-specific random effects by using Dirichlet process mixtures. Also, we incorporate nonparametric modelling on serial correlation with Dirichilet process mixture of AR structure covariance parameters using Ornstein-Uhlenbeck processes. Further, we discuss extending the model structures to multivariate/functional regression models that enable us to make a flexible inference on smooth hyperplane relaxing linearity or additivity assumptions in covariates through tensor product basis functions. We apply our methods to synthetic data and real examples with applications to meta-analysis and stochastic frontier analysis.

Chopin, Nicolas

Convergence of resampling algorithms

This talk covers some new results on the asymptotic behaviour of the error introduced by the resampling step of a particle filter. General conditions are given for the consistency of the resampling scheme. It is also shown that, standard resampling schemes may be modified in a certain way so that the mean square error is o(N^{-1}).Joint work with Mathieu Gerber and Nick Whiteley, Bristol

Clarke, Bertrand

Model Average vs Model Selection

Tri Le and Bertrand Clarke

We compare the performance of five model average predictors -- stacking, Bayes model averaging, bagging, random forests, and boosting -- to the components used to form them. In all five cases we provide conditions under which the model average predictor performs as well or better than any of its components. This is well known empirically, especially for complex problems, although few theoretical results seem to be available. Moreover, all five of the model averages can be regarded as Bayesian. Stacking is the Bayes optimal action in an asymptotic sense under several loss functions. Bayes model averaging is known to

be the Bayes action under squared error. We show that bagging can be regarded as a special case of Bayes model averaging in an asymptotic sense. Random forests are a special case of bagging and hence likewise asymptotically Bayes. Boosted regression is a limit of Bayes optimal boosting classifiers.

Clyde, Merlise

Bayesian inference post selection

In a post p-value era, the Bayesian paradigm provides a coherent framework for full accounting of model uncertainty, however, there are often compelling reasons to select a single model for interpretability or other reasons. While a 0-1 loss function for model selection leads to selection of the highest posterior model, models such as the Median Probability Model are optimal under squared error loss for the joint decision problem of selection and inference/prediction leading to optimal Bayesian estimates of effect sizes under selection. However, reporting intervals conditional on the selected model ignores the act of selection leading to intervals that do not have the correct coverage. Taking a decision theoretic approach we provide adjusted credible intervals for the joint decision problem of selection and interval estimation.

Craigmile, Peter

Locally Stationary Processes and their Application to Climate Modeling

In the analysis of climate it is common to build non-stationary spatio-temporal processes, often based on assuming a random walk behavior over time for the error process. Random walk models may be a poor description for the temporal dynamics, leading to inaccurate uncertainty quantification. Likewise, assuming stationarity in time may also not be a reasonable assumption, especially under climate change. Based on ongoing research, we present a class of time-varying processes that are stationary in space, but locally stationary in time. We demonstrate how to carefully parameterize the time-varying model parameters in terms of a transformation of basis functions. We present some properties of parameter estimates when the process is observed at a finite collection of spatial locations, and apply our methodology to a Bayesian spatio-temporal climate analysis.

Crawford, Lorin

Bayesian Kernel Models for Statistical Genetics and Cancer Genomics

The main contribution of this talk is to examine the utility of nonparametric regression approaches for solving complex problems in statistical genetics and molecular biology. Many of these types of methods have been developed specifically to be applied to solve similar biological problems. For example, kernel regression models have a long history in statistics, applied mathematics, and machine learning. More recently, nonparametric methods have been extensively utilized as tools to broaden understanding of the genetic basis of phenotypic variation. This talk highlights two novel methods which provide innovative solutions to better address the important statistical and computational hurdles faced when analyzing complex biological data sets --- particularly when phenotypic variation is driven by complicated underlying genetic architectures (e.g. the presence of nonlinear effects involving higher order genetic interactions). The first is a Bayesian nonparametric statistical framework that allows for efficient variable selection in nonlinear regression which we refer to as ``Bayesian approximate kernel regression'', or BAKR. The second is an algorithm for identifying genetic variants that are involved in epistasis without the need to identify the exact partners with which the variants interact. We refer to this method as the ``MArginal ePIstasis Test'', or MAPIT. Here, we develop the theory of these two approaches, and demonstrate their power, interpretability, and computational efficiency for analyzing complex phenotypes. We also illustrate their ability to facilitate novel biological discoveries in several real data sets, each of them representing a particular class of analyses: genome-wide association studies (GWASs), molecular trait quantitative trait loci (QTL) mapping studies, and cancer biology association studies.

Cremers, Jolien

A Bayesian approach to circular regression type models

Models for Bayesian multiple regression and mixed-effects regression based on the projected normal distribution have been developed in the embedding approach to circular data. The downside of these models is that the coefficients are defined on two linear bivariate components and that a circular interpretation is not automatically obtained. This problem of interpretation has been solved by a reparameterization of the linear (fixed) regression coefficients into new circular coefficients. A similar reparameterization of the random coefficients results in intercept and slope variances on a circular scale. The advantage of the Bayesian approach is that posterior distributions and uncertainty estimates of these circular coefficients and variances are automatically obtained.
In addition, the Bayesian approach provides several other tools that are useful for interpretation of circular regression type models. Firstly, it allows us to test inequality constrained hypotheses on circular means and other model parameters. Secondly, checks for model assumptions and model fit can be implemented by sampling from the posterior predictive distribution. The new circular coefficients and variances, tests for inequality constrained hypotheses and several other Bayesian tools are implemented in an R package for fitting Bayesian projected normal regression models.

Culpepper, Steven

Using exploratory restricted latent class models to validate cognitive theory

Advances in educational technology provide teachers and schools with a wealth of information about student performance. Restricted latent class models (RLCMs) offer educational researchers, policy-makers, and practitioners with a psychometric framework for designing instructionally relevant assessments and diagnoses about students' skill profiles. Applications of RLCMs require a Q matrix, which catalogues expert knowledge about the underlying cognitive processes and structure required by a collection of educational tasks. The general unavailability of Q for most datasets poses a barrier to widespread applications of RLCMs in education and the social sciences and recent research accordingly developed fully exploratory methods to estimate Q. However, current exploratory methods do not always offer clear interpretations of the uncovered skills and existing exploratory methods do not use expert knowledge to estimate Q. We consider Bayesian estimation of Q using a prior based upon expert knowledge. The developed method can be used to validate which of the underlying attributes are predicted by experts and to identify residual attributes that remain unexplained by expert knowledge. We report Monte Carlo evidence about the accuracy of selecting active expert-predictors and present an application using Tatsuoka's fraction-subtraction dataset.

Dahl, David

Summarizing Distributions of Latent Structure

In a typical Bayesian analysis, consider effort is placed on "fitting the model" (e.g., obtaining samples from the posterior distribution) but this is only half of the inference problem. Meaningful inference usually requires summarizing the posterior distribution of the parameters of interest. Posterior summaries can be especially important in communicating the results and conclusions from a Bayesian analysis to a diverse audience. If the parameters of interest live in R^n, common posterior summaries are means, medians, and modes. Summarizing posterior distributions of parameters with complicated structure is a more difficult problem. For example, the "average" network in the posterior distribution on a network is not easily defined. This paper reviews methods for summarizing distributions of latent structure and then proposes a novel search algorithm for posterior summaries. We apply our method to distributions on variable selection indicators, partitions, feature allocations, and networks. We illustrate our approach in a variety of models for both simulated and real datasets.

Dalla Valle, Luciana

Bayesian non-parametric conditional copula estimation, with application to twin data

Several studies on heritability in twins aim at understanding the different contribution of environmental and genetic factors to specific traits. Considering the national merit twin study, our purpose is to analyse correctly the influence of socio-economic status on the relationship between twins' cognitive abilities. Our methodology is based on conditional copulas, which enable us to model the effect of a covariate driving the strength of dependence between the main variables.We propose a flexible Bayesian non-parametric approach for the estimation of conditional copulas, which can model any conditional copula density. Our methodology extends the work of Wu, Wang and Walker in 2015 by introducing dependence from a covariate in an infinite mixture model. Our results suggest that environmental factors are more influential in families with lower socio-economic position.

Datta, Abhirup

Regularized calibration of verbal autopsy based cause of death predictions in data-scarce settings

In many countries, national or sub-national statistics on cause specific mortality fractions (CSMF) rely on data from verbal autopsy (VA) surveys. The VA data supplemented with `gold standard' physicians' opinions or full autopsy reports for a smaller subset of the deaths are used to determine prediction rules that can ascertain a cause of death (COD) solely based on the VA report of an individual. The VA based COD prediction algorithms like Tariff and InsillicoVA require a substantial training dataset, where both the VA reports and the true gold standard COD are available. In countries where gold standard COD has been ascertained for an extremely small number of cases, these VA prediction algorithms are usually trained on a standard dataset collected in other countries, and are used to predict CSMF for this new country, thereby ignoring the small but important amount of information provided by the handful of full autopsies conducted in the new country. We propose a hierarchical Bayesian calibration algorithm that takes these predictions from any standard VA based algorithm and calibrates the CSMFs using the small gold standard data for the country. Our model involves regularized estimation of a transition matrices of misclassification rates by the VA prediction algorithms for the new country. A novel Gibbs' sampler using augmented data is also proposed. Data analyses reveal how our method outperforms the prevalent practice of simply using VA algorithms trained a standard dataset to predict cause of deaths.

Dawid, Philip

<Bruno de Finetti's Objectivity

While de Finetti famously rejected all objectivist conceptions of Probability, it is less widely understood that he was a strong advocate of objectivity in the assessment of probabilities. In particular, he allowed that one probability forecaster could be better or worse than another, and emphasised the importance of pitting one's probability forecasts against Nature's actual values, advocating a method based on proper scoring rules for doing so. In this talk, I will survey this and other ways of contrasting probabilities and outcomes, and consider whether, and how, it might be possible to make sense of the concept of "the probability" of an unrepeatable event.

Dawkins, Laura

Iterating on surveys, large then small: quantifying commuter behavioural change

Surveys, when conducted in sequence, provide information to update existing knowledge, contributing greater insight than when analysed in isolation. Existing examples of Bayesian analysis for surveys that use a sequential approach often do so to minimise the number of participants required in labour-intensive intervention studies. In these examples, the posterior distributions of the post-intervention survey model parameters are recursively updated using additional participants, until a stopping criterion is met.

Here, we present an alternative, novel approach in which posterior inference from an initial, large survey of the population, characterising the current state of the world, creates strongly informative priors for equivalent parameters in a smaller, post-intervention survey. Consequently, inference focuses on intervention effects, allowing identification of subtle changes in the outcome, given a small sample of participants.

This novel approach was developed to examine the effect of receiving behavioural intervention messages designed to reduce commuter car usage, and hence congestion, in Exeter, UK. It permitted identification of very small (1-2 days/month) reductions in car usage in response to these intervention messages: insight that would require a much larger sample size under existing sequential approaches or a “one-off” analysis.

De Blasi, Pierpaolo

epsilon-Approximations to the Pitman-Yor process

We consider approximations to the Pitman-Yor process obtained by truncating the stick-breaking representation. The truncation is determined by a random stopping rule that achieves an almost sure control on the approximation error in total variation distance. We derive the asymptotic distribution of the random truncation point as the approximation error epsilon goes to zero in terms of a polynomially tilted positive stable distribution. The usefulness of this theoretical result is demonstrated by devising a sampling algorithm to approximate functionals of the epsilon-version of the Pitman-Yor process.

De Iorio, Maria

Bayesian Nonparametric Autoregressive Models via Latent Variable Representation

We propose a probability model for a collection of random distributions indexed by time. The model is based on the dependent Dirichlet process prior and dependence among the random measures is introduced via latent variables. We impose an autoregressive structure on the distribution of the latent variables which allows to introduce time dependence among the random distributions. We propose a Sequential Monte Carlo algorithm to perform posterior inference. Typical applications involve multiple time series in the biomedical context, e.g. repeated measurements or recurrent gap times. Advantages of the proposed approach include wide applicability, ease of computations, interpretability and time dependent clustering of the observation. K-step nonparametric predictive density functions can be derived. The model retains desirable statistical properties for inference, while achieving substantial flexibility.

De Vito, Roberta

Bayesian Multi-study Factor Analysis for High-throughput Biological Data

This paper presents a new modeling strategy for joint unsupervised analysis of multiple high-throughput biological studies. As in Multi-study Factor Analysis, our goals are to identify both common factors shared across studies and study-specific factors. Our approach is motivated by the growing body of high-throughput studies in biomedical research, as captured by the comprehensive set of expression data on breast tumors considered here. We extend Multi-study Factor Analysis to handle sparsity and high dimensional variables, using a Bayesian framework, and general- izing the sparse Bayesian infinite factor model to multiple studies. We devise novel solutions for the identification issue of the loading matrices. We recover the loading matrices of interest ex-post, by adapting the orthogonal Procrustes approach. We propose an efficient and fast computational scheme, based on Gibbs sampling. A comprehensive simulation analysis shows that the proposed approach performs very well in a range of different scenarios and outperforms standard Factor analysis in all the scenarios illustrated. In the breast cancer application, the factors common to all studies appear to have a robust biological interpretation

Deshpande, Sameer

Simultaneous variable and covariance selection with the multivariate spike-and-slab lasso

A Bayesian procedure is proposed for simultaneous variable and covariance selection using continuous spike-and-slab priors in multivariate linear regression models where q possibly correlated responses are regressed onto p predictors. Rather than relying on a stochastic search through the high-dimensional model space, we develop an ECM algorithm similar to the EMVS procedure targeting modal estimates of the matrix of regression coefficients and residual precision matrix. Varying the scale of the continuous spike densities facilitates dynamic posterior exploration and allows us to filter out negligible regression coefficients and partial covariances gradually. Our method is seen to substantially outperform regularization competitors.

Dobra, Adrian

Measuring Human Activity Spaces With Density Ranking Based on GPS Data

Activity spaces are fundamental to the assessment of individuals' dynamic exposure to social and environmental risk factors associated with multiple spatial contexts that are visited during activities of daily living. We develop a novel nonparametric approach which we call density ranking to measure activity spaces from GPS data. We introduce the mass-volume curve, the Betti number curve, and the persistence curve to quantify activity spaces, and to measure the degree of mobility using notions from geometry and topology. We analyze a GPS dataset that comprises the locations visited by 10 individuals over a six months period, and discuss the determination of the corresponding activity spaces based on our new methods.

Elliott, Michael

Novel Methods for Incorporating Sample Designs in Bayesian Inference

Blending the Bayesian paradigm, with its emphasis on complex modeling, with the survey sampling paradigm, with its emphasis on non-parametrics and robustness, has been difficult. Particularly problematic has been incorporating weights into analysis, which, depending on the setting, can be either unnecessary, helpful to avoid magnifying model misspecificiation, or required if sampling is informative. This talk outlines methods to use recently
developed methodology for incorporating complex sample designs in a weighted finite population Bayesian bootstrap procedure (Dong et al. 2014; Zhou et al. 2016) to incorporate design effects into Bayesian analyses via importance weighting. We consider this approach in a few simulation settings, and discuss applications to accounting for complex
sample design in a joint longitudinal data model of mean and variance trajectories; in the setting of small area estimation, using multiple datasets to develop county-level estimates of risky behavior; and in a missing data setting, accommodating both sample design and measurement error when imputing diet and biomarkers.

Engelhardt, Barbara

Combining information across studies and observations for biomedical image analyses

Recent studies have included histological images--those used to diagnose cancer-- alongside large scale gene expression and somatic mutation analyses. We would like to determine which image features are correlated with subsets of the gene expression levels or somatic mutations. The difficulty in applying a naive model to this application is that there are often very few samples, noise in sample labels, or extreme bias in sample classes, and combining information across studies and labels is extremely advantageous. Here, we present work on using multiple studies to identify image features correlated with high dimensional observations of those samples. To do this, we input the images into a sparse group factor analysis model through a generative convolutional network. In this framework, we identify image features correlated with the high dimensional observations while capturing and controlling for study-specific biases. We apply our model across three different studies and show how the multi-study approach improves results in this domain.

Ensor, Kathy

Filtering and Estimation for a Class of Stochastic Volatility Models with Intractable Likelihoods

We introduce a new approach to latent state filtering and parameter estimation for a class of stochastic volatility models (SVMs) for which the likelihood function is unknown. The -stable stochastic volatility model provides a flexible framework for capturing asymmetry and heavy tails, which is useful when modeling financial returns. However, the -stable distribution lacks a closed form for the probability density function, which prevents the direct application of standard Bayesian filtering and estimation techniques such as sequential Monte Carlo and Markov chain Monte Carlo. To obtain filtered volatility estimates, we develop a novel ABC based auxiliary particle filter, which provides improved performance through better proposal distributions. Further, we propose a new particle based MCMC (PMCMC) method for joint estimation of the parameters and latent volatility states. With respect to other extensions of PMCMC, we introduce an efficient single filter particle Metropolis-within-Gibbs algorithm which can be applied for obtaining inference on the parameters of an asymmetric -stable stochastic volatility model. We show the increased efficiency in the estimation process through a simulation study. Finally, we highlight the necessity for modeling asymmetric -stable SVMs through an application to propane weekly spot prices. This is joint work with Emilian R. Vankov and Michele Guindani.

Ferguson, Neil

New approaches to mapping vector-borne disease transmission intensity

I will discuss how a non-parametric approach “the spatial block bootstrap” can be combined with machine-learning to increase the robustness of continental or global maps of dengue and malaria transmission intensity. I will show how a Metropolis Hastings algorithm using a pseudo-likelihood target algorithm can be used to optimise variable selection for out-of-sample predictive accuracy of the random forest models underpinning such maps. Last, I will discuss some of the applications of the resulting maps for the modelling of disease transmission dynamics and intervention strategies.

Fisher, Jared

Monotonic effects of characteristics on returns

Using standard linear regression to describe the cross-section of expected returns implies that the relationships between firm characteristics and returns are both linear and stationary. While this is historically a common practice, finance literature now suggests that these relationships are nonlinear and/or nonstationary. We submit that these effects should also be monotonic. To this end, we present a time-dynamic, additive quadratic spline model that can constrain individual splines to be monotonic. The additive nature of the model allows it to handle many characteristics from many firms. Selection among characteristics and model complexity are controlled through regularization. The model shrinks irrelevant spline coefficients to zero and selects the appropriate number of knots through carefully designed shrinkage priors. This model design admits parsimony and economically meaningful interpretation. As the model maintains interpretability, these relationships can be evaluated for linearity, stationarity, and monotonicity. We present several of these relationships for examination.

Forbes, Catherine

Outlier Robustness and Empirical Likelihood

We provide a review and assessment of Bayesian methods for moment based models when outliers are present. One available approach to dealing with outliers in this context is to specify moments based on the ideas of M-estimation, while another formally augments the available but non-robustified moment conditions with auxiliary variables aimed at detecting and mitigating the potential negative impact arising from outliers. Such methods generally rely on simulation-based methods for the production of the posterior distribution, and we also explore how outliers can impact on this computation. In addition to providing a review of the available approaches, several empirical and simulated examples will be examined to illustrate the techniques available.

Foti, Nick

Bayesian Dynamical Models of Functional Connectivity from Magnetoencephalography Data

Bayesian Dynamical Models of Functional Connectivity from Magnetoencephalography DataA fundamental question in neuroscience is what role brain dynamics play in implementing complex cognitive processes. For example, maintaining a conversation in noisy environments requires listeners to attend to the speaker while filtering out erroneous sounds. Understanding dynamic interactions between brain regions underlying such auditory attention holds the scientific potential to better understand cognitive disorders. This talk will discuss my recent work on Bayesian dynamical models to infer functional connections, i.e. correlated activity between brain regions, from MEG data collected from multiple subjects performing an auditory attention task. In particular, I will discuss work on a Bayesian state-space model with time-varying dynamics that capture directed functional interactions of interest and how they evolve over the course of the attentiontask. An issue with this analysis that makes a Bayesian approach natural is the massive variability between both subjects and trials, which has limited existing analyses of MEG data to considersubjects individually. This variability can arise due to factors including subjects' performances on the task and any actual differences in the underlying dynamics of interest. I will present extensions tothe proposed Bayesian state-space model to account for this variability and perform joint inferences of functional connectivity using recordings from multiple subjects.

Frazier, David

Model Misspecification in ABC: Consequences and Diagnostics

We analyze the behavior of approximate Bayesian computation (ABC) when the model generating the simulated data differs from the actual data generating process; i.e., when the data simulator in ABC is misspecified. We demonstrate both theoretically and in simple, but practically relevant, examples that when the model is misspecified different versions of ABC can lead to substantially different results. Our theoretical results demonstrate that under regularity a version of the ABC accept/reject approach concentrates posterior mass on an appropriately defined pseudo-true parameter value, while the popular linear regression adjustment to ABC concentrates posterior mass on a completely different pseudo-true value. Our results suggest two diagnostic approaches to diagnose model misspecification in ABC. All theoretical results and diagnostics are illustrated in a simple running example. Joint work with Christian Robert and Judith Rousseau.

Fuentes, Montserrat

Bayesian Tensor Regression for Neuroimaging Data

Imaging data with thousands of spatially-correlated data points are common in many fields. In Neurosciences, magnetic resonance imaging (MRI) is a primary modality for studying brain structure and activity. Modeling spatial dependence of MRI data at different scales is one of the main challenges of contemporary neuroimaging, and it could allow for accurate testing for significance in neural activity. The high dimensionality of this type of data (millions of voxels) presents modeling challenges and serious computational constraints. Methods that account for spatial correlation often require very cumbersome matrix evaluations which are prohibitive for data of this size, and thus current methods typically reduce dimensionality by modeling covariance among regions of interest – coarser or larger spatial units – rather than among voxels. However, ignoring spatial dependence at different scales could drastically reduce our ability to detect important activation patterns in the brain and hence produce misleading results. To overcome these problems, we introduce a novel Bayesian Tensor approach, treating the brain image as response and having a vector of predictors. Our method provides estimates of the parameters of interest using a generalized sparsity principle. This method is implemented using a fully Bayesian approach to characterize different sources of uncertainty. We demonstrate posterior consistency and develop a computational efficient algorithm. The effectiveness of our approach is illustrated through simulation studies and the analysis of the effects of cocaine addiction on the brain structure. We implement this method to identify the effects of demographic information and cocaine addiction on the functioning of the brain.

* This work is in collaboration with Hossein Moradi and Gerry Moeller

Gamerman, Dani

Bayesian inference in spatiotemporal Cox processes driven by multivariate Gaussian processes

This talk addresses Bayesian inference for point patterns with unknown form of the intensity function. The intractability of the likelihood provides a barrier for exact inference, ie without having to resort to approximations or to impose functional constraints. We discuss and implement two solutions to this problem in an exact setting: discretization and augmentation. The latter provides a solution for inference in continuously-varying intensity functions. Spatial and spatio-temporal heterogeneity is also contemplated through flexible forms for effects of covariates and their relevance is illustrated with an epidemiological application. Gaussian processes provide a useful prior representation to represent desired qualitative features such as smoothness without imposing specific patterns. Inference for continuously-varying intensity functions is also presented. Illustrations with synthetic and real data are also provided. Extensions to handle discontinuities/non-stationarity are briefly outlined.

Gelfand, Alan

Spatial Statistics and Environmental Challenges

The worlds of spatial statistics and of environmental modeling are both enormous. In a brief one hour lecture it is not possible to cover much of this terrain. So, I will focus on two large problems which connect both of these areas: modeling of species distributions and modeling of environmental exposure. All of the modeling and inference will be done within the Bayesian framework.For the former, modeling of species distributions is a dominant question in ecology, understanding where species are and why. I will discuss modeling of data in the form of presence-absence as well as abundance. This context places us in the geostatistical realm, observing random responses at fixed sampling locations. In this regard, I will consider modeling species individually and jointly. Then, I will turn to presence-only sampling where the data now is more naturally viewed as a spatial point pattern, always degraded by sampling effort.Finally, here, I will try to shed light on the issue of fusion of presence-absence data with presence-only data over a common region. I will argue that much of the previous literature has handled this problem incorrectly and show that a preferential sampling perspective can yield a more satisfying story.For the second, there is continuing strong interest in describing exposure to environmental contaminants, particularly with concerns regarding their connection to adverse health outcomes. I will begin with the basic geostatistical modeling for environmental exposure both in space and in space-time. The simplest versions model pollutants individually or jointly using monitoring station data. However, these days we often supplement such data with data from computer models and/or remotely-sensed from satellites. This takes us to data fusionchallenges with observations at different spatial scales. Finally, with the enormous growth in data collection, I will also discuss fully Bayesian approaches for handling big datasets in space and time.

Gelman, Andrew

Bayesian Aggregation of Average Data

How can we combine datasets collected under different conditions? This is the well-known problem of meta-analysis, for which Bayesian methods have long been used to achieve partial pooling. Here we consider the challenge when one dataset is given as raw data while the second dataset is given as averages only. In such a situation, common meta-analytic methods can only be applied when the model is sufficiently simple for analytic approaches. When the model is too complex, for example nonlinear, an analytic approach is not possible.

The need for meta-analytic methods applied to complex models arises frequently in clinical drug development. Throughout the different phases of a drug development program, randomized trials are used to establish in stages the tolerability, safety, and efficacy of a candidate drug. At each stage one aims to optimize the design of future studies by extrapolation from the available evidence at the time. This includes collected trial data and relevant external data. How- ever, relevant external data are typically available as averages only, for example from trials on alternative treatments reported in the literature. Moreover, realistic models suitable for the desired extrapolation are often complex (longitudinal and nonlinear). We provide a Bayesian solution by using simulation to approximately reconstruct the likelihood of the external summary and allowing the parameters in the model to vary under the different conditions. We first evaluate our approach using fake-data simulations and then demonstrate its application to the problem in drug development that motivated this research, a hierarchical nonlinear model in pharmacometrics, implementing the computation in Stan.

George, Edward I.

Bayesian Hospital Mortality Rate Estimation: Calibration and Standardization for Public Reporting

Bayesian models are increasingly fit to large administrative data sets and then used to make individualized recommendations. In particular, Medicare's Hospital Compare webpage provides information to patients about specific hospital mortality rates for a heart attack or Acute Myocardial Infarction (AMI). Hospital Compare's current recommendations are based on a random-effects logit model with a random hospital indicator and patient risk factors. Except for the largest hospitals, these recommendations or predictions are not individually checkable against data, because data from smaller hospitals are too limited. Before individualized Bayesian recommendations, people derived general advice from empirical studies of many hospitals; e.g., prefer hospitals of type 1 to type 2 because the observed mortality rate is lower at type 1 hospitals. Here we calibrate these Bayesian recommendation systems by checking, out of sample, whether their predictions aggregate to give correct general advice derived from another sample. This process of calibrating individualized predictions against general empirical advice leads to substantial revisions in the Hospital Compare model for AMI mortality, revisions that hierarchically incorporate information about hospital volume, nursing staff, medical residents, and the hospital's ability to perform cardiovascular procedures. And for the ultimate purpose of meaningful public reporting, predicted mortality rates must then be standardized to adjust for patient-mix variation across hospitals. Such standardization can be accomplished with counterfactual mortality predictions for any patient at any hospital. It is seen that indirect standardization, as currently used by Hospital Compare, fails to adequately control for differences in patient risk factors and systematically underestimates mortality rates at the low volume hospitals. As a viable alternative, we propose a full population direct standardization which yields correctly calibrated mortality rates devoid of patient-mix variation. (This is joint research with Veronika Rockova, Paul Rosenbaum, Ville Satopaa and Jeffrey Silber).

Glickman, Mark

Paired Comparison Models with Tie Probabilities and Order Effects as a Function of Strength

Paired Comparison Models with Tie Probabilities and Order Effects as a Function of Strength: Paired comparison models, such as the Bradley-Terry model and its variants, are commonly used to measure competitor strength in games and sports. Extensions have been proposed to account for order effects (e.g., home-field advantage) as well as the possibility of a tie as a separate outcome, but such models are rarely adopted in practice due to poor fit with actual data. We propose a novel paired comparison model that accounts not only for ties and order effects, but recognizes that the probability of a tie may be greater for stronger pairs of competitors, and that order effects may be more pronounced for stronger competitors. The model and several variants are evaluated on games outcomes from all US Chess Open tournaments from 2006 through 2016, events with players of wide-ranging abilities. The models are fit in a Bayesian framework with informative prior distributions based on pre-tournament ratings. Model comparisons are performed by leave-one-out cross-validation.

Glynn, Chris

Dynamics of Homelessness in Urban America

The relationship between housing costs and homelessness has important implications for the way that city and county governments respond to increasing homeless populations. Though many analyses in the public policy literature have examined inter-community variation in homelessness rates to identify causal mechanisms of homelessness, few studies have examined time-varying homeless counts within the same community. To examine trends in homeless population counts in the 25 largest U.S. metropolitan areas, we develop a dynamic Bayesian hierarchical model for time-varying homeless count data. Particular care is given to modeling uncertainty in the homeless count generating and measurement processes, and a critical distinction is made between the counted number of homeless and the true size of the homeless population. For each metro under study, we investigate the relationship between increases in the Zillow Rent Index and increases in the homeless population. Sensitivity of inference to potential improvements in the accuracy of point-in-time counts is explored, and evidence is presented that the inferred increase in the rate of homelessness from 2011-2016 depends on prior beliefs about the accuracy of homeless counts.

Gorham, Jackson

Measuring Sample Quality with Stein's Method

As the size of datasets has grown, classical methods like Markov chain Monte Carlo have become increasingly burdensome from a computational perspective. Practitioners have been turning to biased Markov chain Monte Carlo procedures that are able to trade off asymptotic exactness for computational speed. Unfortunately, previously used diagnostics to aid with these methods are insufficient for assessing the asymptotic bias incurred.We will introduce a new computable quality measure based on Stein’s method that quantifies the maximum discrepancy between sample and target expectations over a large class of test functions. Our first major theoretical contribution will be showing that our measure converges to zero only if a sample weakly converges to its target distribution. Empirically we will show this discrepancy avoids the problems faced by previous diagnostics, e.g., effective sample size.By considering Ito diffusions with fast-mixing rates, we will be able to extend the purview of acceptable target distributions from distributions with strongly log-concave densities with bounded third and fourth derivatives to a much larger class that includes multimodal and heavy-tailed distributions. Finally, we will study a variation of our previous measures that is computationally feasible to obtain for much larger samples. This variation will combine our ideas with those from reproducing kernel Hilbert spaces to define a closed-form expression for our measure. We will show that many commonly used kernel functions provably do not control convergence, while other kernels will be posited that do control weak convergence.

Grazian, Clara

Modelling Ranking Data with Wallenius Distribution

Ranking data sets is useful when statements on the order of observations are more important than the magnitude of their differences and little is known about the underlying distribution of the data. Human beings naturally tend to rank objects in the everyday life such as, for instance, food, shops, singers, football teams, according to their preferences. More generally, to rank a set of objects means to arrange them in order with respect to some characteristic. In some cases, the goal of an analysis is studying the importance of particular categories of elements, given the observed ranking of such elements. Ranked data are often employed in contexts where objective and precise measurements can be impossible or unreliable and the observer collects ordinal information about preferences, judgements, relative or absolute ranking among competitors, called items. We develop inference for ranked data based on the Wallenius distribution, a generalization of the hypergeometric distribution, which consider elements in different sizes but also with different weights. The deriving likelihood is not analytically available, therefore we propose a method to manage the model based on approximate Bayesian computation (ABC).

Griffin, Jim

Exploiting conjugacy to build time dependent completely random measures

A flexible approach to build stationary time-dependent processes exploits the concept of conjugacy in a Bayesian framework: in this case, the transition law of the process is defined as the predictive distribution of an underlying Bayesian model. Then, if the model is conjugate, the transition kernel can be analytically derived, making the approach particularly appealing. We aim at achieving such a convenient mathematical tractability in the context of completely random measures (CRMs), i.e. when the variables exhibiting a time dependence are CRMs. In order to take advantage of the conjugacy, we consider the wide family of exponential completely random measures. This leads to a simple description of the process which has a autoregressive structure of order 1. The proposed process can be straightforwardly employed to extend CRM-based Bayesian nonparametric models such as feature allocation models to time-dependent data. These processes can be applied to problems from modern real life applications in very different fields, from computer science to biology. In particular, we develop a dependent latent feature model for the identification of features in images and a dynamic Poisson factor analysis for topic modelling, which are fitted to synthetic and real data.

Gruber, Lutz

Estimating Open Populations from Multiple Structurally Different Data Sets

We present a novel statistical model for open populations that can integrate multiple data sources such ones based on field interviews, point counts, and time lapse photography. Our statistical model is built around a differential equation for intraday arrivals and departures (immigration and emigration) to which various aspects of each data set connect. The differential equation core of our model enables inference on highly relevant, but unobserved quantities of interest such as total number hunting trips during a day, number of unique hunters accessing a site during a day, person-hour usage, or length of stay.

A real data example based on the Nebraska hunter survey, which is one of the most comprehensive studies of hunters in the US, illustrates the viability of our proposed modeling approach and provides unique insights into hunter turnout, site utilization as well as seasonal effects. Our study integrates field interview data, point count surveys, and intraday time lapse photography. Improved understanding of key factors impacting hunter turnout can help managers of open access programs with their land portfolio investment and policy decisions to improve ecological conditions, hunter satisfaction as well as increase revenue from license sales.

Gu, Mengyang

Scaled Gaussian stochastic process for mathematical model calibration and prediction

We consider the problem of calibrating imperfect mathematical models using experimental data. To compensate for the misspecification of the mathematical model, a discrepancy function is usually included and modeled via a Gaussian stochastic process (GaSP), leading to better results of prediction. The calibrated mathematical model itself, however, sometimes fits the experimental data poorly, as the calibration parameters become unidentifiable. In this work, we propose the scaled Gaussian stochastic process (S-GaSP), a novel stochastic process for calibration and prediction. This new approach bridges the gap for two predominant methods in calibrating misspecified models, namely the L_2 calibration and GaSP calibration. A computationally feasible approach is introduced for this new model under the Bayesian paradigm. The theoretical properties will also be studied with numerical illustrations for both simulated and real data.

Guglielmi, Alessandra

Bayesian Nonparametric Covariate-Driven Clustering: An Application to Blood Donors Data

Blood is an important resource in global health care and therefore an efficient blood supply chain is required. Predicting arrivals of blood donors is fundamental since it allows for better planning of donations sessions. With the goal of characterizing behaviors of donors, we analyze gap times between consecutive blood donations. In order to take into account population heterogeneity we adopt a Bayesian model for clustering. In such a context, defining the model boils down to assign the prior for the random partition itself and to flexibly assign the cluster-specific distribution, since, conditionally on the partition, data are assumed iid within each cluster and independent between clusters. In particular, we drive the prior knowledge on the random partition by increasing the probability that two donors with similar covariates belong to the same cluster. The resulting model is a covariate-dependent nonparametric prior, thus departing from the standard exchangeable assumption. Specifically, we modify the prior on the partition prescribed by the class of normalized completely random measures by including in the prior a term that takes into account the distance between covariates. After a short discussion on the theoretical implications of this prior, we fit our model to a large dataset provided by AVIS (Italian Volunteer Blood-donors Association), which is the largest provider of blood donations in Italy. (Joint work with Raffaele Argiento and Ilaria Bianchini).

Guindani, Michele

Bayesian Nonparametric Methods in Biomedical Imaging

In this talk, we will discuss the use of flexible Bayesian nonparametric approaches in biomedical imaging. We will first discuss an application to the analysis of task-related fMRI data from multi-subject experiments, where the aim is to account for the between-subjects heterogeneity in neuronal activity and cluster subjects into subgroups characterized by similar brain responses. We will then discuss an application to cancer radiomics, an emerging discipline that promises to elucidate lesion phenotypes and tumor heterogeneity through the analysis of large amounts of quantitative imaging features that can be derived from medical images. We will show how a fully Bayesian probabilistic framework may help characterizing the heterogeneity of adrenal lesions images obtained from CT scans more precisely than a class of machine-learning approaches currently used in cancer radiomics. We further assess whether the subtypes resultant from our analysis are clinically oriented by investigating their correspondence with pathological diagnoses.</blockquote></div></blockquote>

Gutierrez, Luis

A Bayesian nonparametric multiple testing procedure for comparing several treatments against a control

We propose a Bayesian nonparametric strategy to test for differences between a control group and several treatment regimes. Most of the existing tests for this type of comparison are based on the differences between location parameters. In contrast, our approach identifies differences across the entire distribution, avoids strong modeling assumptions over the distributions for each treatment, and accounts for multiple testing through the prior distribution on the space of hypotheses. The proposal is compared to other commonly used hypothesis testing procedures under simulated scenarios. Two real applications are also analyzed with the proposed methodology.

Gutiérrez-Peña, Eduardo

Bayesian non-parametric inference for the overlap of daily animal activity patterns

* Joint work with G. Nuñez-Antonio, A. Contreras-Cristán and M. Mendoza

The study of the interaction among species is an active area of research in Ecology. In particular, it is of interest to evaluate the overlap of their ecological niches. Temporal activity is commonly used to explore ecological segregation among animal species, and many contributions focus on the overlap between the activity of two species. There exist different methods to estimate this overlap. One approach is to estimate the density of the temporal activity of each species and then evaluate the overlap between these densities. This leads to the analysis of circular data. Most of the procedures currently in use impose some restrictive assumptions on the probabilistic models used to describe temporal activity and only provide approximate measures of the uncertainty involved in the process. We propose a Bayesian non-parametric approach based on the projected normal distribution. We also briefly discuss a simple -but more flexible- measure of niche overlap. The procedure is applied to the analysis of real camera-trap data concerning three mammalian species from the 'El Triunfo' biosphere reserve (Chiapas, Mexico).

Hahn, Richard

Bayesian Regression Tree Models for Causal Inference: Regularization, Confounding and Heterogenous Effects

We develop a semi-parametric Bayesian regression model for estimating heterogeneous treatment effects from observational data. Standard nonlinear regression models, which may work quite well for prediction, can yield badly biased estimates of treatment effects when fit to data with strong confounding. Our Bayesian causal forests model avoids this problem by directly incorporating an estimate of the propensity function in the specification of the response model, implicitly inducing a covariate-dependent prior on the regression function. This new parametrization also allows treatment heterogeneity to be regularized separately from the prognostic effect of control variables, making it possible to informatively shrink to homogeneity, in contrast to existing Bayesian non- and semi-parametric approaches.

Hammerling, Dorit

A Bayesian hierarchical model for climate-change detection and attribution

Regression-based detection and attribution methods continue to take a central role in the study of climate change and its causes. We propose a Bayesian hierarchical approach to this problem, which allows us to address several open methodological questions. Specifically, we take into account the uncertainties in the true temperature change due to imperfect measurements, the uncertainty in the true climate signal under different forcing scenarios due to the availability of only a small number of climate model simulations, and the uncertainty associated with estimating the climate-variability covariance matrix, including the truncation of the number of empirical orthogonal functions (EOFs) in this covariance matrix. We apply Bayesian model averaging to assign optimal probabilistic weights to different possible truncations, and incorporate all uncertainties into the inference on the regression coefficients. We provide an efficient implementation of our method in a software package and illustrate its use with a realistic application.

Hannart, Alexis

Probabilities of Causation of Climate Change

Multiple changes in Earth's climate system have been observed in past decades. Determining how likely each of these changes are to have been caused by human influence is important for decision making and policy. Here we describe an approach for deriving the probability that anthropogenic forcings have caused a given observed change. The proposed approach is anchored in causal counterfactual theory (Pearl 2000) which has already been used in the context of weather and climate-related events attributions. These concepts can also be extended to the context of climate change attribution. For this purpose, and in agreement with the principle of fingerprinting in the conventional D&A framework, a trajectory of change is converted into an event occurrence defined by maximizing the causal evidence associated with the forcing under scrutiny. Bayesian concepts are central to this derivation. Other key assumptions in D&A, in particular those related to numerical models error, can also be adapted conveniently to this approach. An illustration suggests that our approach may yield a significantly higher estimate of the probability that anthropogenic forcings have caused the observed temperature change, thus supporting more assertive causal claims.

Haran, Murali

Predicting the Impact of Vaccination on Disease Dynamics

Models for the transmission dynamics of infectious diseases are critical for estimating the current burden of the disease from imperfect surveillance. They are also central for assessing potential effects of vaccine intervention strategies. We study a variety of Susceptible-Infectious-Recovered (SIR) compartmental models for rotavirus disease dynamics in southern Niger. The SIR models are embedded in a hierarchical modeling framework to include multiple data sources such as hospital surveillance data and cluster surveys of households. Model-specific parameters are estimated using Markov chain Monte Carlo. We then use Bayesian model averaging to generate ensemble estimates of the current dynamics, the burden of infection in the region, as well as the impact of vaccination on both the short-term dynamics and the long-term reduction of rotavirus incidence under varying levels of coverage. This is joint work with JaewooPark (Penn State Statistics) and Matthew Ferrari (Penn State Center for Infectious Disease Dynamics).

Hegerl, Gabi

Detection and attribution of climate change and estimating climate sensitivity: why we need a Bayesian perspective

This talk discusses detection and attribution methods based on a least square fit, and discusses the key uncertainties that come into play: uncertainties in observations (gaps are excluded from analysis, but observations are uncertain in their trends as well), in climate model fingerprints in response to individual forcings, and of course, natural variability which will create uncertainty in any estimate of climate responses from observations and, even more so, in estimates of responses to multiple external influences on climate simultaneously (e.g., greenhouse gases, aerosols, solar and volcanic forcing). It introduces methods, including Bayesian methods, to deal with this uncertainty and a simple attempt to consider them all simultaneously, which highlights the need to fully account for uncertainties in a transparent way.

Held, Leonhard

Power Priors Based on Multiple Historical Studies

Incorporating historical information into the design and analysis of a new clinical trial has been the subject of much recent discussion. For example, in the context of clinical trials of antibiotics for drug resistant infections, where patients with specific infections can be difficult to recruit, there is often only limited and heterogeneous information available from the historical trials. To make the best use of the combined information at hand, we consider an approach based on the multiple power prior which allows the prior weight of each historical study to be chosen adaptively by empirical Bayes. This estimate has advantages in that it varies commensurably with differences in the historical and current data and can choose weights near 1 if the data from the corresponding historical study are similar enough to the data from the current study. Fully Bayesian approaches are also considered. An analysis of the operating characteristics in a binomial setting shows that the proposed empirical Bayes adaptive method works well, compared to several alternative approaches, including the meta-analytic prior.

Helena Loschi, Rosangela

Modeling Underreported Data with a Random Censoring Poisson Model

A major challenge when monitoring risks in socially deprived areas, is that economic, epidemiological and social data are typically underreported. To deal with this problem, counts in suspected areas are usually approached as censored information. The censored Poisson model (CPM), used to model underreported counts, assumes that all censored areas must be precisely known a prior, which is not a reasonable assumption in most practical situations. We develop a new model that describes the joint behavior of the observed data and the data reporting process. We propose a random mechanism to specify the censored (underreported) regions thus introducing a random censoring Poisson model (RCPM). Assuming the proposed model, we are able to estimate both the censoring probability and the relative risk for the event of interest in each region. We elicit different prior distributions for the censoring probabilities. The main focus is to infer about the relative risks and about the censoring indicator latent variables. Despite obtaining full conditional distributions with closed form, the Gibbs sampler shows to be an inefficient algorithm to sample from the posteriors. We develop an MCMC algorithm to sample from the posterior distribution. The algorithm relies on the data augmentation strategy (Tanner&Wong,1987, Chib1992) simplifying substantially the posterior sampling. We run a simulation study evaluating the proposed algorithm and comparing the proposed model and the CPM in different scenarios. We perform a sensitivity analysis by eliciting different prior distributions for the censoring probabilities in RCPM, and by fixing different censoring criteria in CPM. We apply the proposed model to account for potential underreporting in the early neonatal mortality counts in regions of Minas Gerais State, Brazil, in which data quality is known to be truly poor.

Herring, Amy

Bayesian Multi-study Factor Analysis for High-throughput Biological Data

In modern biomedical studies, it has become commonplace to collect high-dimensional data, and hence dimensionality reduction tools are of critical importance and are routinely used. Some of the most common include clustering and factor analysis. The basic tenet behind dimensionality reduction is that we can replace a high dimensional set of variables by some low-dimensional summary. This is certainly necessary to make sense of complex data and also overcome problems with high-dimensional, low sample size data. However, data-driven dimensionality reduction for a particular group of individuals may fundamentally lack generalizability to other groups of individuals. This creates major problems in interpretation of results. Motivated in particular by environmental epidemiology studies collecting exposome data and by nutritional epidemiology, we develop fundamentally new methods for improving robustness and generalizability of dimensionality reduction.

Higdon, Dave

Combining federal and local data sources to better understand the state of a local community

A variety of local and federal data sources give information about where individuals and households are, their demographics, their economic and social condition, what they do, how they move, and how they interact - socially, behaviorally, and economically. This research focuses on a couple of small case-studies that combine diverse data sources with probabilistic models to gain better understanding of a local community. One example has grown from a collaboration with Department of Housing and Urban Development where a static model of households for Arlington County was developed to assess privacy/identifiability issues with different data releases for the American Housing Survey. Another focuses on using local EMS service calls to assess safety and health incidents within Arlington County. While the individual questions vary, depending on the investigation, much of the empirical model structure and estimation methodology can be reused for different investigations.

Hill, Jennifer

Nonparametric Bayesian Causal Inference with Heterogeneous Treatment Effects

Causal inference in the absence of a randomized experiment typically requires modeling of the expectation of the potential outcomes conditional on treatment and a (potentially high-dimensional) set of covariates.This modeling is complicated when this response surfaces for the potential outcomes are non-linear and further complicated when they are not parallel to each other, leading to heterogeneous treatment effects.Grouped data structures create additional challenges to estimation.We will present nonparametric Bayesian approaches to treatment effect estimation involving Gaussian Processes and treed algorithms that can accommodate non-linear, non-parallel response surfaces and grouped data structures.We will also discuss approaches to lack of common support in these scenarios.

Hilton, Jason

The Futures of UK Population.

A great number of policy and planning decisions depend on predictions about the future population sizes. In the United Kingdom, migration is the contributor to future population change about which we know the least: the data are patchy and the migration processes are very uncertain and volatile. On top of this general uncertainty, migration policies in the UK will likely be revised once an eventual settlement is reached on Brexit. When planning for the future, we need to know not only how many people we think there will be, but also how confident we can be in saying so. In this talk, we will discuss the range of possible effects of different migration regimes on the future structure of the UK population, and the potential consequences. To do so, we will present a fully Bayesian model of population change, with the two other components of demographic dynamics - mortality and fertility - described by fully probabilistic parametric and semi-parametric models, and migration incorporated through a range of stochastic scenarios. Sensitivity of the results to those scenarios will enable assessing the importance of migration for shaping the future demographic dynamics of the UK.

Holan, Scott

Computationally Efficient Bayesian Hierarchical Models for High-Dimensional Dependent Multinomial Data

We introduce a Bayesian approach for multivariate spatio-temporal prediction for high-dimensional multinomial data.Our primary interest is when there are possibly millions of observation referenced over different variables, geographic regions, and times. This problem is methodologically and computationally challenging, as jointly modeling correlated data of this size leads to the so-called ``big n problem” within a non-Gaussian setting.We develop a new computationally efficient distribution theory for this setting and provide substantial theoretical development including: results regarding conditional distributions, marginal distributions, and full-conditional distributions for a Gibbs sampler. To incorporate dependence between variables, regions, and times, a multivariate spatio-temporal mixed effects model is used. The results in this manuscript are extremely general, and can be used for data that exhibit fewer sources of dependency than what we consider (e.g., multivariate, spatial-only, or spatio-temporal-only). We show the effectiveness of our approach through a simulation study and demonstrate our proposed methodology with an important application analyzing data obtained from the Longitudinal Employer-Household Dynamics program, which is administered by the U.S. Census Bureau.The methods presented here are of independent interest and apply to many diverse areas of statistics, including industrial statistics (e.g., multinomial control charts) and ecological modeling, among others.

Holmes, Chris

A Bayesian Nonparametric Approach to Testing for Dependence Between Random Variables and the Two-sample problem

Here we describe Bayesian nonparametric procedures that lead to tractable, explicit and analytic quantification of the relative evidence for dependence vs independence and for the two-sample hypothesis testing. Our approach uses PÃ³lya tree priors on the space of probability measures which can then be embedded within a decision theoretic test. The PÃ³lya tree prior is centered either subjectively or using an empirical procedure.

Huerta, Gabriel

Spatio-temporal modeling of heavy tailed data via non-Gaussian latent processes

We introduce a new spatio-temporal model that allows for Generalized Pareto marginal distributions where time and space dependence is incorporated via the use of latent variables in a hierarchical fashion. Furthermore, the model relies on a conjugate, non-Gaussian structure that considers a Gamma-Poisson process. We study some of the properties of this new model and in particular, tail dependence, a characteristic of key relevance in extreme value analysis. We follow a Bayesian approach to estimate the model to produce posterior inference on quantities of interest such as high quantiles or predictions. We compare our approach with others existing in the literature that also consider a peak-over-threshold marginal distribution.

Huggins, Jonathan

Scaling Bayesian inference using exponential family approximations

Modern statistical applications present many challenges for standard approximate posterior inference algorithms: the latent parameter is often high-dimensional, the models can be complex, and there are large amounts of data that may only be available as a stream or distributed across many computers. Existing algorithms have so far remained unsatisfactory because they either (1) fail to scale to large data sets, (2) provide limited approximation quality, or (3) fail to provide theoretical guarantees on the quality of inference. To simultaneously overcome these three possible limitations, I leverage the critical insight that in the large-scale setting, much of the data is redundant. Therefore, it is possible to compress data into a form that admits more efficient inference. I propose a method called PASS-GLM, which constructs an exponential family model that approximates the original model. The data is compressed by calculating the finite-dimensional sufficient statistics of the data under the exponential family. I also obtain a priori guarantees on the approximate posteriors produced by PASS-GLM.

Ickstadt, Katja

Shrinkage estimation methods for subgroup analyses

Subgroup analyses are performed in confirmatory clinical trials, where the treatment effect is estimated in subgroups of the overall trial population defined by certain patient characteristics. The appropriateness of subgroup-specific treatment effects is controversial because of multiplicity and small sample sizes within the subgroups. A useful alternative are estimators of subgroup effects which take the overall effect estimate into account. Shrinkage estimators belong to this class. They combine the overall effect estimate with the estimate within a given subgroup by using a Bayesian framework including a prior distribution for the interaction effect. We briefly introduce the shrinkage estimation approaches by Dixon & Simon (1991) and by Simon (2002), one with a non-informative and one with an informative prior. These methods have been defined for subgroup factors with two categories. We extend them to factors with more than two categories and provide solutions for computational issues. Moreover, we compare naïve and shrinkage approaches in a simulation study for different models and scenarios. The aim is to understand in which cases the shrinkage estimators are superior to the common naïve estimator.Joint work with Julian Riehl.

Inacio, Vanda

Bayesian nonparametric inference for the covariate-adjusted ROC curve

Accurate diagnosis of disease is of fundamental importance in clinical practice and medical research. Before a medical diagnostic test is routinely used in practice, its ability to distinguish between diseased and non-diseased states must be rigorously assessed through statistical analysis. The receiver operating characteristic (ROC) curve is the most popular used tool for evaluating the discriminatory ability of continuous-outcome diagnostic tests. It has been acknowledged that several factors (e.g., subject-specific characteristics, such as age and/or gender) can affect the test's accuracy beyond disease status. Recently, the covariate-adjusted ROC curve has been proposed and successfully applied as a global summary measure of diagnostic accuracy that takes covariate information into account. We develop a highly flexible nonparametric model for the covariate-adjusted ROC curve, based on a combination of a B-splines dependent Dirichlet process mixture model and the Bayesian bootstrap, that can respond to unanticipated features of the data (e.g., skewness, multimodality, and/or excess of variability). Multiple simulation studies demonstrate the ability of our model to successfully recover the true covariate-adjusted ROC curve and to produce valid inferences in a variety of complex scenarios. Our methods are motivated by and applied to an endocrine dataset where the main goal is to assess the accuracy of the body mass index, adjusted for age and gender, for predicting clusters of cardiovascular disease risk factors.

Jara, Alejandro

Bayesian nonparametric approaches for the analysis of compositional data based on Bernstein polynomials

We will discuss Bayesian nonparametric procedures for density estimation and fully nonparametric regression for compositional data, that is, data supported in a multidimensional simplex. The procedures are based on modified classes of Bernstein polynomials. We show that the modified classes retain the well known approximation properties of the classical versions defined on an hypercube and on a multidimensional simplex. Based on these classes, we define prior distributions, and study the support and

asymptotic behavior of the posterior distribution under iid sampling. Finally, novel classes of probability models for sets of time-dependent probability distributions are proposed. Appealing theoretical properties such as support, continuity, marginal distribution, correlation structure, and consistency of the posterior distribution are studied.

Jauch, Michael

Bayesian Analysis with Orthogonal Matrix Parameters

Models for multivariate data are often naturally parametrized in terms of orthogonal matrix parameters. Bayesian analysis with orthogonal matrix parameters presents two major challenges: posterior sampling on the constrained parameter space and incorporation of prior information, such as sparsity. We propose methodology to address both of these challenges. To sample from posterior distributions defined on the set of orthogonal matrices, we introduce a parameter expansion scheme based on the polar decomposition. To incorporate sparsity information, we construct prior distributions having element-wise marginal distributions

approximately matching (up to rescaling) conventional sparsity-inducing priors. We illustrate these techniques in simulation studies and applications to data.

Jensen, Mark

Real exchange rates and unit roots: Nonparametric Bayesian learning about the distribution

<p>The goal of this paper is to learn about the distribution of the autoregression coefficient for real exchange rates, including the probability of a unit-root. The paper is an exercise in Bayesian statistics. The approach we take allows us to learn not only about the distribution for each specific case for which we have data, but also the generic case for which we have no data as yet. The posterior distribution for the generic case constitutes a well-informed prior distribution for a new case when such data becomes available. The estimation of the distribution for the generic case amounts to indirect density estimation for a latent variable. With this in mind, we adopt a nonparametric Bayesian prior that embodies great flexibility and puts positive probability on a unit-root as a special case.</p></blockquote></blockquote>

Jensen, Shane

Bayesian Modeling of Crime in Philadelphia

Urban data analysis has been recently improved through publicly available high resolution data, allowing us to empirically investigate urban design principles of the past half century. Philadelphia is an interesting case study for this work with its recent population growth and substantial urban development. We are primarily interested in measures of the health and safety of a neighborhood and how they relate to local neighborhood features, including land use zoning, business locations and activity as well as economic measures and population density. I will focus this talk on current efforts to spatially and temporally model crime in the city of Philadelphia over the past decade

Ji, Yuan

Triple-adaptive Bayesian designs for the identification of optimal dose combinations in dual-agent dose-finding trials

We propose a flexible design for the identification of optimal dose combinations in dual-agent dose-finding clinical trials. The design is called AAA, standing for three adaptations: adaptive model selection, adaptive dose insertion, and adaptive cohort division. The adaptations highlight the need and opportunity for innovation for dual-agent dose finding, and are supported by the numerical results presented in the proposed simulation studies. To our knowledge, this is the first design that allows for all three adaptations at the same time. We find that AAA improves the statistical inference, enhances the chance of finding the optimal dose combinations, and shortens the trial duration. A clinical trial is being planned to apply the AAA design.

Jones, Beatrix

Inferring common inverse-covariance elements across conditions using decoupled shrinkage and selection

Scientists are accustomed to comparing means between case and control conditions, however interesting differences may also occur in the covariance structure. For moderate-to-high dimensional data, simply concluding that covariance matrices are different is of little utility; differences need to be localized to particular variables and relationships. While previous approaches infer common sparse structure in the inverse covariance matrices, we address the case where the individual matrices may not be sparse, but the matrix of differences is. We employ the decoupled shrinkage and selection approach to infer common elements across inverse covariance matrices for multiple conditions (e.g. case and control). Our inputs to this process are samples from independent posteriors over the inverse covariance for each condition, leading to substantial flexibility in model choice. Decoupled shrinkage and selection then allows us to visualize the tradeoff between model fit and common matrix elements across conditions, providing a principled comparison of selection strategies. The fit of the Bayes estimates of each inverse covariance to samples from its posterior predictive distribution provides a distribution of fit; this distribution is used as a reference to judge an acceptable fit-sparsity trade off and decide on a final model. The method is illustrated with examples from metabolomics.

Jordan, Michael

On Gradient-Based Optimization: Accelerated, Stochastic and Nonconvex

It has proved challenging to develop algorithms for Bayesian inference that scale effectively to high-dimensional parameter spaces and complex posterior distributions. The asymptotic guarantees provided by criteria such as ergodicity and reversibility of Markov chains are often silent, or pessimistic, when it comes to convergence rates, to the dimension dependence of those rates and the dependence of rates on the geometry of posterior distributions. We will overview some of the ways that similar challenges have been met in the optimization literature in recent years, where new theory and algorithms have provided non-asymptotic rates, sharp dimension dependence, elegant ties to geometry and practical relevance. In particular, I discuss several recent results: (1) a new framework for understanding Nesterov acceleration, obtained by taking a continuous-time, Lagrangian/Hamiltonian/symplectic perspective, (2) a discussion of how to escape saddle points efficiently in nonconvex optimization, and (3) the acceleration of Langevin diffusion.

Kaplan, Andee

Counting Casualties in the Syrian Civil War with Bayesian Record Linkage

How does one quantify the number of identifiable deaths in the ongoing Syrian conflict? What might seem an easy question to answer belies a complex problem in the statistical field of record linkage, which seeks to combine multiple sources of data in cases where records cannot be easily matched. Simply put, war zone records are messy and large-they are full of inaccuracies and duplicates, are comprised of multiple data sets from multiple sources, and contain hundreds of thousands of records. Further, existing data sets are far from complete, and represent only a fraction of the overall number of fatalities. In this talk we will use Bayesian record linkage methods to address the problem of entity resolution for four databases of casualties from the Syrian civil war and provide multiple evaluations of the validity of the methods.

Karabatsos, George

Bayes calculations from quantile implied likelihood

A Bayesian model can have a likelihood function that is analytically or computationally intractable, perhaps due to large data sample size or high parameter dimensionality. For such a model, this article introduces a likelihood function that approximates the exact likelihood through its quantile function, and is defined by an asymptotic chi-square distribution based on confidence distribution theory. This Quantile Implied Likelihood (QIL) gives rise to an approximate posterior distribution, which can be estimated either by maximizing the penalized log-likelihood, or by a standard adaptive Metropolis or importance sampling algorithm. The QIL approach to Bayesian Computation is illustrated through the Bayesian modeling and analysis of simulated and real data sets having sample sizes that reach the millions. Models include the Student's t, g-and-h, and g-and-k distributions; the Bayesian logit regression model; and a novel high-dimensional Bayesian nonparametric model for distributions under unknown stochastic precedence order-constraints. Note: This will be a presentation of the paper appearing in: https://arxiv.org/abs/1708.05341

Karagiannis, Georgios

Bayesian calibration of expensive computer models with input dependent parameters

Computer models, aiming at simulating a real physical procedures, often involve complex parametrisations whose optimal parameter values may change with respect to the model inputs. In such scenarios, traditional model calibration methods assume that these optimal values are constant to the inputs, and hence they may produce misleading results. We present a fully Bayesian methodology for the calibration of expensive computer models. Our procedure is able to produce optimal values for the uncertain model parameters as functions of the model inputs, as well as it can characterize the associated uncertainties via posterior distributions. Moreover, it is particularly suitable to address problems where the computer model requires the selection of a sub-model from a set of competing ones, but the choice of the “optimal” sub-model may change with the input values. Central to methodology is the idea of formulating the unknown parameter as a step function whose uncertain structure is modeled via a binary treed process. Our method is compared with the traditional Bayesian model calibration against a benchmark example, and it is implemented in a climate model application.

Kavelaars, Xynthia

Stopping rules in Bayesian adaptive clinical trials with multiple endpoints.

Randomized controlled trials are considered the gold standard to investigate the effectivity of medical treatments. While clinical trials often study multiple endpoints, current trials mainly apply univariate analysis that fail to take correlations between endpoints into account to each of these outcomes. Analyzing multiple endpoints multivariately allows for borrowing strength across outcomes and detection of multivariate treatment effects. Multivariate analysis complicates the trial design however, and planning sample sizes adequately is challenging when multiple parameters must be taken into account. The flexibility of adaptive designs to react to uncertainty in the design stage and correct for unanticipated effects in the data is of particular value in trials with multiple outcomes. Adaptive designs require decision criteria, and it is unclear how to formulate these in a multivariate context. We extend the adaptive trial design framework with a multivariate approach to statistical analysis and decision-making with binary variables. First, we propose to analyze multiple binary outcomes using a multivariate beta prior distribution with a multivariate Bernoulli likelihood. Second, we present stopping rules for the resulting multivariate beta posterior distribution that take the multivariate nature of the data into account.

Keller, Colleen

Bayesian Search for Missing Aircraft

In recent years, there have been a number of highly publicized searches for missing aircraft such as the ones for Air France flight AF 447 and Malaysia Airlines flight MH 370. Bayesian search theory provides an established approach for planning searches for missing aircraft. The theory has been applied to searches for the wreck of AF 447, MH370, and many other missing aircraft, often successfully concluding the search after formal search efforts have failed. This talk presents the basic elements of the theory and how it was proposed to search authorities after official efforts failed. It uses real world examples such as the AF447 search, a successful search for a missing Cessna 182 in Arizona, US (2008), and the search for Steve Fossettâs missing plane in California (2007). In all these examples, the difficulties in explaining the Bayesian approach and obtaining data necessary to implement it will be described."

Kohn, Robert

Efficiently Combining Pseudo Marginal and Particle Gibbs Sampling

Particle Markov Chain Monte Carlo methods are used to carry out inference in non-linear and non-Gaussian state space models, where the posterior density of the states is approximated using particles. Deligiannidis (2017) introduce the correlated pseudo marginal sampler and show that it can be much more efficient than the standard pseudo marginal approach. Mendes (2018) propose a particle MCMC sampler that generates parameters that are highly correlated with the states using a pseudo marginal method that integrates out the states, while all other parameters are generated using particle Gibbs. Our article shows how to combine these two approaches to particle MCMC to obtain a flexible sampler with a superior performance to each of these two approaches. We illustrate the new sampler using a multivariate factor stochastic volatility model with leverage. Joint work with David Gunawan and Christopher K. Carter

Kon Kam King, Guillaume

Bayesian inference for hidden Markov models via duality and approximate filtering distributions

Filtering hidden markov models, which can be seen as performing sequential bayesian inference on the hidden state of a latent signal, is an analytically tractable problem only for a handful of models. Among those feature models with finite- dimensional state space and linear gaussian systems, which give rise to the celebrated Baum-Welch and Kalman filters. Recently, Papaspiliopoulos and Ruggiero (2014) and Papaspiliopoulos et al. (2016) proposed a principled approach for extending the realm of analytically tractable models, exploiting a duality relation between the hidden process of interest and an auxiliary process, called dual and related to the time reversal of the former. When such a dual process is available and has certain characteristics, the solution of the filtering problem is available analytically and takes the form of a finite mixture of distributions, which can be evaluated by means of a recursion similar to the Baum-Welch filter. Here, we study the computational effort required to implements the above strategy in the case of two hidden Markov models given respectively by the Cox-Ingersoll-Ross process with Poisson observations and the K-dimensional Wright-Fisher process with multinomial observations. In both cases, the number of components involved in the filtering distributions increases polynomially with the number of observations, yielding a so-called computable filter. This behaviour could render the algorithm impractical for large dimensional hidden spaces or very long observation sequences and undermine its practical relevance. However, the mathematical form of the filtering distributions suggest that, in certain regimes of separation between observation times and speed of the underlying signal, the number of components which contribute most of the mixture mass remains small. This in turn suggests several natural and very efficient approximation strategies. In this contribution, we assess the performance of these strategies in terms of accuracy, speed and prediction, which we can all benchmark against the exact solution. Omiros Papaspiliopoulos, Matteo Ruggiero.

Kovalchik, Stephanie

Deciphering the First Strike - A Bayesian Hierarchical Count Model for Measuring the Serve Advantage in Professional Tennis

The serve advantage is a well-recognised feature of elite tennis and a professional player almost universally wins a higher percentage of points when serving than when receiving. Despite the importance of the concept in tennis, a statistical model for the serve advantage has yet to be proposed. In this paper, I present a hierarchical logistic Bayesian model that estimates a servers conditional probability of winning a point given the number of shots played. The conditional log odds of winning the point is a non-linear decay function of the rally length. When the model was applied to tens of thousands of points from professional matches, it was found that the decay model best described outcomes on first serve. Out-of- sample testing of the first serve advantage model provided 71% predictive accuracy for mens point outcomes and 63% for womens point outcomes, both significant improvements over a naive model. Evaluation of the posterior distribution of the first serve advantage model revealed a serve advantage of 45% for average male and female servers, with the womens curve saturating at the 4th shot in a rally, at a rally ability of 50%, and the mens saturating between the 4th and 5th shot, at a rally ability of 53%. The proposed model is a useful tool for estimating the first serve advantage and for forecasting how service outcomes are influenced by rally length.

Latham, Ben

Bayesian Optimization with Field Experiments

Randomized experiments, commonly referred to as A/B tests, are the gold standard for evaluating the effects of changes to Internet services.Data in these tests may be time-consuming to collect and outcomes may have high variance, resulting in large measurement error. There are also typically constraints on noisy outcomes that must be respected. Bayesian optimization is a promising technique for optimizing multiple continuous parameters via field experiments, but performance can degrade with high noise levels. I will discuss a new approach for constrained Bayesian optimization that is tailored to the high noise levelsof A/B tests. I will then describe two applications of Bayesian optimization to field experiments at Facebook: optimizing a production ranking system, and optimizing JIT compiler settings to improve web server performance.

Lee, Herbie

Modeling an Augmented Langranian for Bayesian Blackbox Constraned Optimization

Constrained blackbox optimization is a difficult problem, with most approaches coming from the mathematical programming literature. The statistical literature is sparse, especially in addressing problems with nontrivial constraints. This situation is unfortunate because statistical methods have many attractive properties: global scope, handling noisy objectives, sensitivity analysis, and so forth. To narrow that gap, we propose a combination of Bayesian response surface modeling, expected improvement, and the augmented Lagrangian numerical optimization framework. This hybrid approach allows the statistical model to think globally and the augmented Lagrangian to act locally. We focus on problems where the constraints are the primary bottleneck, requiring expensive simulation to evaluate and substantial modeling effort to map out. In that context, our hybridization presents a simple yet effective solution that allows existing objective-oriented statistical approaches, like those based on Gaussian process surrogates and expected improvement heuristics, to be applied to the constrained setting with minor modification. This work is motivated by a challenging, real-data benchmark problem from hydrology where, even with a simple linear objective function, learning a nontrivial valid region complicates the search for a global minimum.

Lee, Juhee

Bayesian nonparametric methods for joint analysis of recurrent events and survival times

Recurrent events are oftentimes observed with a terminal event, such as death. If the recurrent events and the risk of death are related, then death informatively censors the recurrent event process. We present Bayesian nonparametric methods for joint modeling of a recurrent event process and survival time. The models take the dependence between the recurrent events and death into account through latent frailties and provide enhanced inference for the underlying processes. More specifically, assuming independent subject latent frailties, marginal stochastic models are defined for the recurrent event process and survival time as functions of baseline covariates and the frailties. We exploit Bayesian nonparametric approaches to model the intensity function of the recurrent event process and the hazard function of the survival process for robust inference. We also investigate more general stochastic models to achieve a wider range of dependence structures between the processes.

Lewin, Alex

Bayesian inference on high-dimensional Seemingly Unrelated Regressions, applied to metabolomics data.

Increasingly, epidemiologists collect multiple high-dimensional molecular data sets on large cohorts of people, in order to find associations between genetic variants, biomarkers and other phenotypes. In order to do this effectively these multivariate data sets should be modelled jointly, taking into account correlations in the data. Sparse solutions are usually required, and performing variable selection in this setting is critical. We present a Bayesian Seemingly Unrelated Regressions (SUR) model for high-dimensional data, allowing for both sparse variable selection and correlation between the outcomes. This model can be fit using a Gibbs sampler, but this quickly becomes computationally unfeasible as the dimensions of the problem grow, due to lack of conjugacy in the model. Previous work in this area has made use of either the assumption of independence between the outcomes or selected predictors jointly for all the outcomes, both of which allow for conjugate priors. In order to overcome some of the computational difficulty with the general SUR model, we use a reparametrisation of the model in which the likelihood factorises completely into a product of conditional distributions, and build a MCMC sampler capable of handling real molecular biology data involving 100's or 1000's of responses.

Li, Meng

Gaussian process on the unit circle with an application to image boundary detection

Gaussian processes (GPs) indexed by the unit circle provide a flexible prior for closed smooth curves. We show that a squared exponential GP indexed by the unit circle exhibits fundamental differences with a squared exponential GP on the unit interval. We provide analytical eigen decomposition of the covariance kernel of a squared exponential GP on the unit circle, leading to highly scalable implmentation due to efficient inverse of large covariance matrices. We further conduct an extensive study of the reproducing kernel Hilbert space of a squared exponential GP on the unit circle. We apply the circle-indexed GP prior to an image boundary detection problem, and show that the posterior contraction rate is (nearly) minimax optimal in recovering the boundary. An R package and interactive shiny application are available to implement the proposed method.

Li, Wentao

Validating Approximate Bayesian Computation via Posterior Convergence

The likelihood function is key to many statistical methods, including maximum likelihood estimation and Markov chain Monte Carlo. But for many realistic models in modern applications, it is difficult to evaluate the likelihood functions due to high model complexity, e.g. models with nonlinear dynamics and latent structure. Approximate Bayesian computation (ABC) method implements Bayesian inference without evaluating the likelihood function, only requiring the ability to simulate pseudo datasets from the model. More specifically, the inference is based on simulated datasets that are close to the observed data. ABC has been popular within population genetics and ecology over a decade, and has recently found wide applications in other areas involving financial time series and stochastic differential equation. This talk will discuss the limitation of standard implementation of ABC, in the sense that it can perform well in terms of point estimation, but will over-estimate the uncertainty about the parameters. If we use the regression correction of Beaumont et al.(2002), then ABC can also accurately quantify this uncertainty, hence achieve the same asymptotic accuracy as the likelihood-based methods.

Lijoi, Antonio

Full Bayesian inference with multiple-samples data

Data from multiple samples typically display a dependence structure that is not consistent with the exchangeability assumption. In the talk we consider classes of random measures based models that are particularly suited to address density estimation and survival analysis problems in this setting. We display some novel conditional representations that are the key for devising simulation algorithms for a full Bayesian inference. These may be further used to test hypotheses on the distributional homogeneity across different samples. Illustrative examples with real and simulated data will be finally discussed.

Lindgren, Finn

Large Spatio-temporal Modelling and Computing for Past Weather and Climate

Combining multiple and large data sources of historical temperatures into unified spatio-temporal analyses is challenging from both modelling and computational points of view. This is not only due to the size of the problem, but also due to the highly heterogeneous data coverage and the latent heterogeneous physical processes. The EUSTACE project will give publicly available daily estimates of surface air temperature since 1850 across the globe for the first time by combining surface and satellite data using novel statistical techniques. Of particular importance is to obtain realistic uncertainty estimates, due to both observation uncertainty and lack of spatio-temporal coverage. To this end, a spatio-temporal multiscale statistical Gaussian random field model is constructed, with a hierarchy of spatially non-stationary spatio-temporal dependence structures, ranging from weather on a daily timescale to climate on a multidecadal timescale. Connections between SPDEs and Markov random fields are used to obtain sparse matrices for the practical computations. The extreme size of the problem necessitates the use of iterative solvers, which requires using the multiscale structure of the model to design an effective preconditioner. We will also discuss some new computational developments towards full Bayesian estimation of the dependence parameters, and fast computation of posterior predictive variances.

Lobato, Jose

The Bayesian Optimization for Accelerated Exploration of Chemical Space

Chemical space is so large as to make a brute force search for molecules with improved properties infeasible. Bayesian optimization methods can accelerate the discovery process by sequentially identifying the most useful experiments to be performed next. However, existing methods have shortcomings that limit their applicability to the molecule search problem. First, they lack scalability to the large amounts of data that are required to successfully navigate chemical space. Second, they are unable to learn feature representations for the data, which reduces their statistical efficiency in the large data scenario. Third, they often fail when the search space is discrete as is the case of chemical space. In this talk I will give a brief introduction to Bayesian optimization methods and then I will present different contributions that aim to solve or at least alleviate the aforementioned problems.

Lopes, Hedibert

Dynamic sparsity on dynamic regression models

We consider variable selection and shrinkage for the Gaussian dynamic linear regression within a Bayesian framework. In particular, we propose a novel method that allows for time-varying sparsity, based on an extension of spike-and-slab priors for dynamic models. This is done by assigning appropriate Markov switching priors for the time-varying coefficients' variances, extending the previous work of Ishwaran and Rao (2005). Furthermore, we investigate different priors, including the common Inverted gamma prior for the process variances, and other mixture prior distributions such as Gamma priors for both the spike and the slab, which leads to a mixture of Normal-Gammas priors (Griffin et al., 2010) for the coefficients. In this sense, our prior can be view as a dynamic variable selection prior which induces either smoothness (through the slab) or shrinkage towards zero (through the spike) at each time point. The MCMC method used for posterior computation uses Markov latent variables that can assume binary regimes at each time point to generate the coefficients' variances. In that way, our model is a dynamic mixture model, thus, we could use the algorithm of Gerlach et al. (2000) to generate the latent processes without conditioning on the states. Finally, our approach is exemplified through simulated examples and a real data application. This is joint work with Paloma Vaissman Uribe.

Love, Tanzy
A Topic Model in Social Space
Social media outlets have greatly increased the availability, prevalence, and relevance of networking data. Societal norms are characterized by network connections, with people communicating and interacting frequently with other individuals from around the world. As networks become ever more common, there grows a large source of data relating to inter-personal relationships and connections. Much of this data naturally links to additional sources of information that can be used to better understand network models. We present a unified Bayesian approach for hierarchically soft clustering discrete data linked to members of a network. Focusing on text documents authored by members of a social network as our source of discrete data, we introduce a model that simultaneously represents a network and clusters text in a single social space, providing a unified representation of social relationships and topic membership. This method is evaluated through simulation studies, and extensions are introduced to allow for the more general inclusion
of network data into discrete-data soft-clustering methods.

Lu, Bo
Bayesian Model-based Sensitivity Analysis for Heterogeneous Treatment Effects
Unmeasured confounding is a major threat to the validity of analyzing observational data. Sensitivity analyses have been proposed to evaluate the impact of some unmeasured covariate on the findings based on the observed data. Both parametric and nonparametric sensitivity analyses have been developed for evaluating a homogeneous treatment effect. In the presence of heterogeneous treatment effects, however, these methods become fairly complicated to implement and interpret. We propose a model-based sensitivity analysis that models the difference between true and observed treatment effect as a function of the propensity score. Using the propensity score simplifies the modeling substantially, thanks to the dimension reduction property. Bayesian approach is used to infer the confidence bound of the true treatment effect as there is a component of missing confounders. We illustrate our method by examining the potential impact of college attendance on later-in-life fertility using the National Longitudinal Survey of Youth (NLSY).

Ma, Li
A Bayesian hierarchical model for related densities using Polya trees
Bayesian hierarchical models are used to share information between related samples and obtain more accurate estimates of sample-level parameters, common structure, and variation between samples. When the parameter of interest is the distribution or density of a continuous variable, a hierarchical model for continuous distributions is required. A number of such models have been described in the literature using extensions of the Dirichlet process and related processes, typically as a distribution on the parameters of a mixing kernel. We propose a new hierarchical model based on the Polya tree, which allows direct modeling of densities and enjoys some computational advantages over the Dirichlet process. The Polya tree also allows more flexible modeling of the variation between samples, providing more informed shrinkage and permitting posterior inference on the dispersion function, which quantifies the variation among sample densities. We also show how the model can be extended to cluster samples in situations where the observed samples are believed to have been drawn from several latent populations.

MacEachern, Steve
TBA
TBA

Mackey, Lester

Measuring Sample Quality with Stein's Method

To improve the efficiency of Monte Carlo estimation, practitioners are turning to biased Markov chain Monte Carlo procedures that trade off asymptotic exactness for computational speed. The reasoning is sound: a reduction in variance due to more rapid sampling can outweigh the bias introduced. However, the inexactness creates new challenges for sampler and parameter selection, since standard measures of sample quality like effective sample size do not account for asymptotic bias. To address these challenges, we introduce new computable quality measures based on Stein's method that quantify the maximum discrepancy between sample and target expectations over a large class of test functions. We use our tools to compare exact, biased, and deterministic sample sequences and illustrate applications to hyperparameter selection, biased sampler selection, one-sample hypothesis testing, and sample quality improvement.

Martin, Ryan

Empirical priors for wranglin' with structured high-dimensional problems

The modern sparse Bayesian cowboys, with their one-grouphorseshoe-inspired shrinkage priors, have all but run the traditionaltwo-groups priors out of Dodge. But there are new buckaroos in townwith two-group priors that use the data to inform their center, different from classical empirical Bayes. This extra giddy up from the data-driven centering permits these new empirical priors to take convenient conjugate forms, making computation relatively simple, without sacrificing on theoretical posterior convergence rate properties. In this talk, I will present this new empirical Bayes approach for structured high-dimensional problems for a variety of sparsity settings.

Maruotti, Antonello

Circular state-space models for time dependent data

Motivated by ecological data, we develop a state-space model in discrete time for circular time-series and longitudinal data. In detail an autoregressive circular multi-state regression model is introduced. The time-dynamics is split into true and apparent contagion. In the former case, actual and future outcomes are directly influenced by past values, which cause a substantial change over time in the corresponding distribution due to serial dependence. The latter case arises when observations are drawn from heterogeneous populations, each population having a time-varying different preferred direction. We propose to estimate the parameters in a non-parametric Bayesian framework, relying on the sticky hierarchical Dirichlet process hidden (semi-)Markov model (sHDP-HMM). This allows to jointly estimate model parameters and the latent process structure without fixing it a priori. We show how to estimate the parameters using a Markov chain Monte Carlo (MCMC) algorithm.

Mastrantonio, Gianluca

Introducing dependenceamong compositonal vectors: parametric and non parametric approaches

A huge literature about clustering spatial time data exists. The problem has been studied both in a parametric and in a nonparametric setting. There are several problems in defining a proper clustering procedure, depending on the type of relationship between the clusters.

From a parametric point of view, a classic approach is to introduce mixture models and studying the posterior distribution of the mixture weights. We propose a mixture model where the mixing probabilities are time specific and are assumed to follow a Logistic-Normal distribution. We introduce dependence betweens the vectors of mixing probabilities by means of a Gaussian processes representation.

In a nonparametric setting, Dirichlet processes are often used, however how to introduce dependence between processes is still an open problem. The Dirichlet process may be extended in a hierarchical version, so that several processes share the same set of atoms with process-dependent weights, however the original construction of the hierarchical Dirichlet process considers weights that are independent for different processes. In this work we also propose a way to introduce dependence in the marginal distributions of the vectors of weights, by imposing a Gaussian copula whose correlation matrix has a given dependence structure (for instance, implying spatial-temporal dependence).

Mathews, Heather

Computationally efficient hybrid methods for network analysis

Modern methods for the analysis of network data rely on complex latent space representations to evaluate the effects of covariate information on edge formation. These methods are extremely flexible but become computationally prohibitive as networks grow in size. To overcome this computational hurdle, we propose hybrid methods that can capture part of the latent space very quickly while spending most of the computational effort on evaluating the effects of the covariates of interest. Specifically, we consider the flexible, but computationally costly additive and multiplicative effects latent space model where one of the features captured by the latent space is group-like behavior that is observed in real world networks. There exist simpler algorithms (such as spectral methods) that can capture some of these more complicated structures at a fraction of the computational cost. We employ a Bayesian framework for fitting the more complex models while conditioning on the output of the simpler algorithms; thus trading off some statistical risk for a decrease in computational cost. This type of hybrid method leads to efficient approximate recovery of group membership information as well as to capturing covariate information. We demonstrate this methodology on several network datasets.

Mattei, Alessandra

Assessing causal effects in the presence of treatment switching through principal stratification

In clinical trials focusing on survival outcomes for patients suffering from Acquired Immune Deficiency Syndrome (AIDS)-related illnesses and particularly painful cancers in advanced stages, patients in the control arm are often allowed to switch to the treatment arm if their physical conditions are worse than certain tolerance levels. The Intention-To-Treat analysis, comparing groups formed by randomization regardless of the treatment actually received, is often used; although it provides valid causal estimates of the effect of assignment, it does not give information about the effect of the actual receipt of the treatment and ignores the information of treatment switching in the control group. Other existing methods in the literature propose to reconstruct the outcome a unit would have had if s/he had not switched but they are usually based on strong assumptions, like that there exist no relation between patient’s prognosis and switching behavior or the effect is constant in same scale. Clearly, the switching status of the units in the control group contains important post-treatment information and it is useful to characterize the heterogeneity of the treatment effect. We propose to re-define the problem of treatment switching using principal stratification and introduce new causal estimands, principal causal effects for patients belonging to subpopulations defined by the switching behavior under the control treatment, which appropriately adjust for the post-treatment information and characterize treatment effect heterogeneity. For inference, we use a Bayesian approach, which allows us to properly take into account that (i) switching happens in continuous time generating a continuum of principal strata; (ii) switching time is not defined for units who never switch in a particular experiment; and (iii) both survival time, the outcome of primary interest, and switching time are subject to censoring. We illustrate our framework using simulated data based on the Concorde study, a randomized controlled trial aimed to assess causal effects on time-to- disease progression or death of immediate versus deferred treatment with zidovudine among patients with asymptomatic HIV infection.

Matthews, Robert

Principles and Practice of the Analysis of Credibility

The inferential inadequacies of statistical significance testing are now widely recognised. There is, however, no consensus on how to move research into a post p < 0.05 era. We present a potential route forward via the Analysis of Credibility, a novel methodology that allows researchers to go beyond the simplistic dichotomy of significance testing and extract more insight from new findings. Using standard summary statistics, AnCred assesses the credibility of significant and non-significant findings on the basis of their evidential weight, and in the context of existing knowledge. The outcome is expressed in quantitative terms of direct relevance to the substantive research question, providing greater protection against misinterpretation. Worked examples are given to illustrate how AnCred extracts additional insight from the outcome of typical research study designs. Its ability to cast light on the use of p-values, the interpretation of non-significant findings and the so-called "replication crisis" is also discussed.

McAlinn, Kenichiro
Dynamic sparse factor models
Factor analysis has proven to be a core and essential tool for multivariate analysis, especially in the fields of macroeconomics and finance. However, due to the dynamic nature of the data, combined with the large number of relevant series available, developments in dynamic factor analysis for high-dimensional multivariate time series have been limited. To address this critical issue, we develop a novel methodology for sparse dynamic factor analysis using the recently proposed class of Dynamic Spike-and-Slab (DSS) priors. Under this class of priors, sparsity is dynamically induced through the factor loadings via time-varying mixing weights between a spike process for the irrelevant coefficients and a slab autoregressive process for the active coefficients. Efficient Bayesian estimation is proposed for MAP estimates using the EM algorithm, and full posterior samples using traditional MCMC methods. We apply the proposed methodology on topical high-dimensional macroeconomic and financial datasets, and demonstrate the efficacy in terms of inference and forecasts.

McElreath, Richard
Teaching Bayesian Statistics Without Frequentist Language
Most students encounter Bayesian statistics after learning classical, or Frequentist, statistics. As a result, Bayesian concepts and models are nearly always introduced using Frequentist language. This can result in lasting confusion about the Bayesian approach, even among those who use it routinely. To advance this argument, I examine two cases of Frequentist language in widespread use in Bayesian statistics and reexplain the underlying concepts using new terms. The first case is the replacement of Frequentist "parameters" and "data" with Bayesian "variables", both latent and observed. The second case is the replacement of both "likelihood" and "prior" with "distribution". I demonstrate with examples how relaxing these Frequentist distinctions makes it easier to introduce solutions to problems like missing values and measurement error. I do not advocate changing statistical terminology. I advocate only appreciation of the friction created by using Frequentist terms in Bayesian contexts and strategizing in our teaching to avoid this friction.

Mena, Ramses
Exchangeable mixed renewal processes
The tractability of exchangeable sequences devised from Bayesian considerations is exploited to define mixed renewal-type processes. After a relative small variation of classical results, much of the available theory remains valid for the more general setting of exchangeability. Here, we present some of these results and exemplify them with parametric and non-parametric sequences. We also provide some simulated and real data estimation examples to highlight the importance of the more general view we consider.

Mengersen, Kerrie
Transferability: a Bayesian Perspective
Transferability is a term used in a range of disciplines to indicate the transfer of a model developed in one context to a different context. This can be considered as a form of extrapolation, where the degree of extrapolation is potentially unknown. Whereas traditional modelling approaches struggle with transferability, a Bayesian framework offers an opportunity to accommodate this additional uncertainty through a reformulation of the loss function and the model space. In this poster, we explore these ideas and evaluate their implementation in a substantive case study that focuses on conservation of coral reefs. Here, the question is whether, and how, we can transfer a model developed in a data-rich setting on the Great Barrier Reef in eastern Australia to a data-poor setting on Ningaloo Reef in Western Australia.

Mira, Antonietta

Approximate Bayesian Computation for Inference in Mechanistic Network Models

Many systems of scientific interest can be investigated as networks, where network nodes correspond to the elements of the system and network edges to interactions between the elements. Increasing availability of large-scale data, and steady improvements in computational capacity, are continuing to fuel the growth of this field. Network models are now used commonly to investigate complexity at the systemic level. There are are least two prominent paradigms to the modeling of networks, which might be called the statistical approach and the mechanistic approach. In the statistical approach, one describes a model that specifies the likelihood of observing a given network, i.e., these are probabilistic models of data that happen to arrive in the shape of a network. In the mechanistic approach, one specifies a set of domain-specific mechanistic rules that are used to grow or evolve the network over time. Both modeling approaches provide distinct angles and advantages to our understanding of complex systems. One of the main strengths of the statistical approach is the ability to perform parameter inference and model selection in a statistically principled manner, but a downside is that these models usually do not scale well, and incorporation of data generating mechanisms in the likelihood function can be difficult. Mechanistic models, on the other hand, make it easy to incorporate aspects of network generation and these models are generally scalable, but there are few principled tools available for them for parameter inference and model selection. I will discuss an approximate Bayesian computation (ABC) framework, also implemented and parallelized in python, that may be used to carry out both parameter inference and model selection for mechanistic network models. In the former, a mechanistic network model might incorporate some parameter theta whose value is not fixed and must therefore be learned from data. The proposed approach is able to infer the point estimate and the associated uncertainty, such that given a single observed network, it is possible to learn about the value of the parameters that was used to generate the network. In a related problem of model selection, consider a setting where there are two or more mechanistic network models, one observed network, and the goal is to learn our degree of belief that any of the candidate models was used to generate the network. The proposed approach is able to quantify the probability that any given model in the collection generated the data, and this approach can therefore be used to select the most likely model from among this collection of candidate models. The advantage of the parameter inference and model selection framework presented is that it is, in principle, applicable to any mechanistic network model. Examples of applications of ABC for inference in mechanistic network models will be presented. Joint work with Jukka-Pekka Onnela, Department of Biostatistics, Harvard University.

Moews, Ben

Massively parallel Bayesian nonparametrics for cosmological parameter estimation

Current efforts in cosmological parameter estimation often suffer from both the computational costs of approximating distributions in high-dimensional parameter spaces and the wide-spread need for model tuning. Specifically, calculating the likelihoods of parameter proposals through simulations imposes high computational costs, leading to excessive time requirements per experiment. We propose and implement an iterative approach that incorporates Bayesian nonparametrics, allowing for the more accurate drawing of samples from an approximated posterior distribution. Exploiting both parallelisation and recent advances in statistical methodology, we provide the astronomy community with an easy-to-use and self-optimizing tool for accelerated cosmological parameter estimation.

Mohamed, Shakir

Bayesian Deep Learning for Reasoning and Acting

Our recent recent research is motivated by the goal of equipping artificial agents with the abilities needed to realistically reason and act in complex environments. These agent-abilities are varied, and include scene understanding, concept formation, future-thinking, abduction, memory, and planning, amongst many others. To be effective, each of these functions must enable plausible reasoning: the use of probabilistically consistent inferences and decision-making based on experience. This talk will explore the plausible reasoning that is possible with Bayesian deep learning and develop several agent-reasoning components: generative models with the ability for one-shot generalisation; models that can form temporally-extended future predictions; temporal models with external memory; approaches for visual concept learning; density-estimators that aid exploration; and policy search methods for hierarchical planning.

Moores, Matthew
Sequential Monte Carlo for multivariate calibration in chemometrics
Raman spectroscopy can be used to identify molecules by the characteristic scattering of light from a laser. Each Raman-active dye label has a unique spectral signature, comprised by the locations and amplitudes of the peaks. The presence of a large, non-uniform background presents a major challenge to analysis of these spectra. We introduce a sequential Monte Carlo (SMC) algorithm to separate the observed spectrum into a series of peaks plus a smoothly-varying baseline, corrupted by additive white noise. The peaks are modelled as Lorentzian, Gaussian, or pseudo-Voigt functions, while the baseline is estimated using a penalised cubic spline. Our model-based approach accounts for differences in resolution and experimental conditions. We incorporate prior information to improve identifiability and regularise the solution. By utilising this representation in a Bayesian functional regression, we can quantify the relationship between molecular concentration and peak intensity, resulting in an improved estimate of the limit of detection. The posterior distribution can be incrementally updated as more data becomes available, resulting in a scalable algorithm that is robust to local maxima. These methods have been implemented as an R package, using RcppEigen and OpenMP.

Mortera, Julia
Inference about relationships from DNA mixtures
We present methods for inference about relationships between contributors to a DNA mixture and other individuals of known genotype: a basic example would be testing whether a contributor to a mixture is the father of a child of known genotype. The evidence for such a relationship is evaluated as the likelihood ratio for the specified relationship versus the alternative that there is no relationship. Our methods are based on a statistical model for DNA mixtures that uses continuous gamma distributed data, such as electropherogram peak heights. We use Bayesian network (BN) as a computational device. Since exact probability propagation methods of Bayes nets do not work with continuous distributions, we introduce auxiliary nodes in a Bayesian network so as to compute the exact evaluation of the likelihood function and other quantities by probability propagation. We introduce auxiliary likelihood nodes to also represent family relationships.
Furthermore, the methods can be readily extended to analyse different scenarios as our methods are not limited to the particular genotyping kits used in the examples, to the allele frequency databases used, to the numbers of contributors assumed, to the number of traces analysed simultaneously, nor to the specific hypotheses tested.
We use Bayesian network (BN) as a computational device. Since exact probability propagation methods of Bayes nets do not work with continuous distributions, we introduce auxiliary nodes in a Bayesian network so as to compute the exact evaluation of the likelihood function and other quantities by probability propagation. We introduce auxiliary likelihood nodes to also represent family relationships.
Furthermore, the methods can be readily extended to analyse different scenarios as our methods are not limited to the particular genotyping kits used in the examples, to the allele frequency databases used, to the numbers of contributors assumed, to the number of traces analysed simultaneously, nor to the specific hypotheses tested.

Mueller, Peter
Prior Choices for Reciprocal Graphical Models
Constructing gene regulatory networks is a fundamental task in systems biology. We introduce a Gaussian reciprocal graphical model for inference about gene regulatory relationships by integrating mRNA gene expression and DNA level information including copy number and methylation. Data integration allows for inference on the directionality of certain regulatory relationships, which would be otherwise indistinguishable. We use a mapping of the graph to simultaneous equation models to implement inference and discuss alternative default prior choices and Bayesian model selection techniques. We illustrate our approach by simulations and two applications in ZODIAC pairwise gene interaction analysis and colon adenocarcinoma pathway analysis.

Mukherjee, Bhramar

Can Bayesian Methods Offer a Practical Alternative to P-values in Genetic Association Studies?

In this talk I will discuss the historic context of defining the genomewide threshold of 5.0 x 10^{-8} and whether modifications/updates are necessary dependent on the type of studies. Bayesian proposals for multiple testing correction will be discussed and new evaluation metrics for assessing average performance of methods across a large number of tests will be proposed. I will argue that empirical Bayes approaches that trade off between bias and variance in a data adaptive way for both estimation and ranking hold promise for such agnostic large-scale analyses.

Mulder, Kees

Bayesian estimation and hypothesis tests for a circular Generalized Linear Model

Motivated by a study from cognitive psychology, we develop a Generalized Linear Model for circular data using the von Mises distribution. Although circular data arise in a wide variety of scientific fields, the number of methods for their analysis is limited. Our model allows inclusion of both continuous and categorical covariates. In a frequentist setting, this type of model is plagued by the likelihood surface of its regression coefficients, which is not logarithmically concave. In a Bayesian context, a weakly informative prior solves this issue, while for other parameters noninformative priors are available. In addition to an MCMC sampling algorithm, we develop Bayesian hypothesis tests based on the Bayes factor for both equality and inequality constrained hypotheses. In a simulation study, it can be seen that our method performs well. The analyses are available in the package CircGLMBayes. Finally, we apply this model to a dataset from experimental psychology, and show that it provides valuable insight for applied researchers. Extensions to dependent observations are within reach by means of the multivariate von Mises distribution.

Murray, Jared

Interpreting complex models: Efficient, valid posterior inference for meaningful quantities

We propose a conceptually simple framework for making Bayesian inferences about interpretable models that summarize complex posterior distributions. This provides a vehicle for understanding large, complicated, and often nonparametric models. Our approach is able to map the output of state-of-the-art predictive tools onto scientifically meaningful quantities while maintaining valid posterior inference. It also provides a bridge between Bayesian methods and recently popular frequentist methods for post-selection inference. We illustrate the general approach in two important special cases: Summarizing high-dimensional linear regression models with lower-dimensional alternatives, and interrogating the fit of a nonparametric regression model (Bayesian additive regression trees).

Neumann, Cedric

Approximate Bayesian Computational and Kernel Based methods for extending the Bayesian paradigm for forensic evidence to high dimensional settings

Since Lindley’s 1977 paper on forensic science, the Bayes Factor (BF) has been advocated to quantify the weight of forensic evidence; however due to the lack of tractable likelihood structures in the natural high-dimensional space where complex evidence is observed, it cannot be directly evaluated for most settings. In this presentation, we will review two methods for extending the Bayesian paradigm to forensic model selection problems in high dimensional settings. The first approach is based on a modification to the standard ABC model selection method that sidesteps the need to select a threshold when comparing summary statistics. The second approach is based on the construction of a map from the product sample space, where the evidence naturally lives, to the real line (usually referred to as a kernel). This mapping allows us to transform the problem from a high dimensional model selection problem into an inference about the dependency structure of the pairwise scores on the real line that is analogous to Bayesian methods for relational data. We will discuss these methods in the context of fingerprint and dust evidence. The methods will be contrasted with score based methods that were developed in the late 2000's and early teens.

Ni, Yang

Heterogeneous Reciprocal Graphical Models

We develop novel hierarchical reciprocal graphical models to infer gene networks from heterogeneous data. In the case of data that can be naturally divided into known groups, we propose to connect graphs by introducing a hierarchical prior across group-specific graphs, including a correlation on edge strengths across graphs. Thresholding priors are applied to induce sparsity of the estimated networks. In the case of unknown groups, we cluster subjects into subpopulations and jointly estimate cluster-specific gene networks, again using similar hierarchical priors across clusters. We illustrate the proposed approach by simulation studies and two applications with multiplatform genomic data for multiple cancers. (Joint work with P. Mueller)

Nicholls, Geoff

Calibrating biased credible intervals: a wrapper validating HPD intervals for partitions clustering the levels of a categorical variable

The computational demands of exact Bayesian inference are a limiting factor in routine Bayesian analysis. The basic statistical instruments we use are commonly expressed as high dimensional integrals of complex intractable functions. Credible intervals are particularly valuable. We have available a number of approximation schemes, which may be used inside asymptotically exact Monte Carlo schemes, which lead to ``double asymptotics''. In these settings the estimated interval has exact Bayes coverage, for the prior under consideration, in the two limits (A) that the approximation is exact, and (B) that the Monte Carlo method converges in its sampling variation. We treat problems in which the log-likelihood in an MCMC algorithm is replaced with a deterministic or randomly distributed approximation. While some of these methods have a bias with provably good double asymptotics, it is unclear, at any finite computational cost, what harm is done to the Bayes coverage. The ideal validation would show that approximation bias is small compared to the Monte Carlo error. Some recent work, based on an idea due to Geweke, have suggested that a generic ``wrapper'' validating approximations, is available. It focuses on controlling the bias in point estimates of posterior expectations of parameters, and it is not clear to us to what extent it tests the reliability of posterior credible intervals. We are motivated by a question addressed to two meat-digestion datasets. The question of interest concerns the rate at which different amino acids in meat are digested. Earlier work suggested that the amino acids fall into a small number of groups. In a natural regression model the different amino acids enter as categorical variables explaining measured digestion response. The problem is cast as one of clustering the levels of a categorical variable in a mixed effects model. We use a Dirichlet process to model the distribution over possible groupings. We frame this as a marginal distribution with a CRP prior over the unknown groupings of the labels of the amino acids in a hierarchical model. A great deal of software is available to estimate marginal likelihoods in this setting, but the impact, on key credible intervals, of replacing marginal likelihoods with estimates in an MCMC algorithm is hard to quantify. We show how the same set of ideas may be adapted to address posterior coverage. We have two novel contributions: first, we identify a natural family of models expressing the grouping as a clustering on random effects, with a simple Bayesian computational scheme to estimate the distribution over partitions, and second we give a generic approximation-validation scheme which validates the credible intervals we provide. The former is a special case of a great deal of earlier work, but certain elements of the model family seem to us natural and of wider interest.

Nieto-Barajas, Luis

Exploring pairwise dependence via Dirichlet Process Mixtures

We propose Bayesian nonparametric methods using Dirichlet Process Mixture (DPM) models for detecting pairwise dependence between random variables while accounting for uncertainty in the form of the underlying distributions. A key criteria is that the procedures should scale to large data sets. In this regard we find that the formal calculation of the Bayes factor for a dependent-vs.-independent DPM joint probability measure is not feasible computationally. To address this we present Bayesian diagnostic measures for characterising evidence against a null model of pairwise independence. We illustrate our approach in simulation studies as well as for a real data analysis.

Nishimura, Aki

Methods of applied mathematics for Bayesian computation

Many techniques in applied mathematics and scientific computing have not made their ways into statistical computation, but they have much to offer in advancing the field. In this talk, I will present an application of non-smooth Hamiltonian dynamics to address the inability of Hamiltonian Monte Carlo (HMC) to handle discrete parameters — one of the most prominent weaknesses of the widely-used algorithm. The proposed extension, discontinuous HMC, remains suitable as a general-purpose algorithm for probabilistic programming languages (such as Stan) and demonstrates competitive performances compared to other alternatives. When properly-tuned, discontinuous HMC is guaranteed to outperform a variable-at-a-time Metropolis approach as the two algorithms coincide under a specific (and sub-optimal) implementation of discontinuous HMC. I will also discuss a more recent work on scaling Bayesian variable selection to “large n, large p” problems (e.g. n >= 10^5, p >= 10^4). For approaches based on global-local shrinkage priors, a major computational bottleneck is sampling the regression coefficients from its conditional distribution given in the form of a high-dimensional multivariate Gaussian. We achieve an order of magnitude speed-up in carrying out this computation using a novel algorithm based on the conjugate gradient method for solving a large linear system.

Ogburn, Elizabeth

Social networks, causal inference, and chain graphs

Most methods for causal inference using non-experimental data are explicitly based on or implicitly equivalent to Bayes network models. In social network settings, these models require longitudinal data on exposures and outcomes, resulting in two challenges. First, it is often impossible to collect the kind of data required; the time intervals must be short enough to capture every potential transmission event. Second, even if the full longitudinal data are available, the resulting model will generally be high-dimensional and often too big to fit to the available data. Most researchers deal with reduced data, comprised of observations collected at one or a few time points. We propose and justify a parsimonious parameterization for social network data with interference and contagion, corresponding to a particular family of graphical models known as chain graphs. We demonstrate that chain graph models approximate the projection of the full longitudinal Bayes network model onto the observed data, which is missing most time points from the full data. We illustrate the use of chain graphs for approximate causal inference about contagion, interference, and collective decision making in social networks when the longitudinal evolution of treatments/outcomes is not fully observed with data on Supreme Court decisions.

O'Hagan, Anthony

In Praise of Subjectivity?

Bayesian analysis requires that probabilities are subjective. Attempts to escape this apparently unwelcome fact are numerous, and they are ultimately misguided because science itself is necessarily subjective. Instead, we should embrace the opportunity to incorporate additional knowledge into the analysis through the prior distribution. But that doesn't make subjectivity admirable in itself. Science strives towards objectivity and so should we - all subjective input should be based on careful, rigorous judgement. In this talk, I will enlarge on these ideas with various examples.

Ommen, Danica

"How strong is the evidence? And how can you tell?"

Formal Bayesian methods lead to the quantification of the weight of forensic evidence (WOE). In practice, however, there are many practical hurdles in the implementation of this theoretical framework for interpreting and evaluating evidence. This has led to the development of methods that can ideally be used in a similar manner to the formal Bayes Factor(s) to update prior beliefs in the hypotheses of prosecution and defence. Large scale settings, such as automated fingerprint identification systems used to screen databases of fingerprints for matches to a fingermark found at the scene of a crime, are of special interest. Decisions made using these approximations to the WOE will be inherently ad-hoc, and will be justified based on their performance in large scale testing scenarios. In this presentation we will focus on the strategies for developing and testing approximations for formal WOE that can be deployed in operational forensic settings. The methods used for developing an approximate WOE are typically based on statistical approximation theory as well as numerical approximations. The approaches to validating an approximate WOE are typically based on entropy and related ideas which will be discussed.

Paci, Lucia

Dynamic model-based clustering for spatio-temporal data

We propose a finite space-time mixture model to identify level-based clusters in spatio-temporal data and study their temporal evolution along the observational time frame. Spatio-temporally varying mixing weights are introduced to allocate observations at nearby locations and consecutive time points with similar cluster's membership probabilities. As a result, a clustering varying over time and space is accomplished. Conditionally on the cluster's membership, a state-space model is deployed to describe the temporal evolution of the sites belonging to each group. Fully posterior inference is provided under a Bayesian framework through Monte Carlo Markov chain algorithms. We illustrate our approach using real data examples.

Paciorek, Christopher

Beyond the black box: using, programming, and sharing hierarchical modeling algorithms such as MCMC and Sequential MC using NIMBLE

While algorithmic and software advances have enabled widespread use of MCMC, sequential Monte Carlo, and related methods, many new and old algorithms remain inaccessible for practical use without re-writing them for each model. Moreover, having different packages for different algorithms makes it difficult to combine methods or try different methods on the same problem. And it is generally difficult to modify the algorithms in existing packages to improve performance for specific applications. NIMBLE provides a flexible system for using, programming, and sharing hierarchical modeling algorithms that operate on models defined using BUGS syntax. To illustrate the benefits of the system, we show how different MCMC samplers, including samplers developed by a user, can be mixed and matched for specific models. We illustrate NIMBLE's new ability to fit Bayesian nonparametric models, showing how BUGS syntax is extended for Dirichlet process and related models and how various samplers for such models are defined within the system and used within the overall NIMBLE MCMC engine. Finally we illustrate NIMBLE's ability to compose algorithms by discussing particle MCMC, making use of NIMBLE's sequential MCMC capabilities within an MCMC sampler. We close by discussing how NIMBLE enables sharing of new methods and reproducibility of research.

Page, Garritt
Spatio-Temporal Product Partition Models
Data that are spatially referenced typically represent an instantaneous point in time at which the spatial process is measured. Because of this, it is becoming more common to monitor spatial processes over time with interest in being able to capture interactions between spatial and temporal structures. Approaches to carrying this out require employing non-separable and non-stationary covariance functions, which are not very realistic for many of the processes being modeled, or adopting more realistic but much more complicated and computationally intense methods. We propose studying the temporal evolution of spatial structure by way of a dependent random partition model developed for partitions that are spatially and temporally referenced. The method accommodates space and time by introducing spatiotemporal information in the clustering mechanism. The result is the ability to monitor how spatially referenced partitions evolve over time and a procedure that produces non-stationary and non-separable spatio-temporal dependence structures at a reasonable computational cost.

Peruzzi, Michele
Bayesian Modular and Multiscale Regression
We tackle the problem of multiscale regression for predictors that are spatially or temporally indexed with a Bayesian modular approach. The regression function at the finest scale is expressed as an additive expansion of coarse to fine step functions. Our Modular and Multiscale (M\&M) methodology can be applied to high-dimensional data that arise through very fine measurements. Unlike more complex methods for functional predictors, our approach provides easy interpretation of the results. Additionally, it provides a quantification of uncertainty on the data resolution, solving a common problem researchers encounter with simple models on down-sampled data. We show that our modular and multiscale posterior has an empirical Bayes interpretation, with a simple limiting distribution in large samples. An efficient sampling algorithm is developed for posterior computation, and the methods are illustrated through simulations studies and an application to brain image classification.

Petris, Giovanni
Bayesian temporal and spatial analysis for forest inventory applications
The National Forest Inventory (NFI) conducted by the USDA Forest Service, Forest Inventory and Analysis (FIA) program, provides crucial information about forest conditions. Although the NFI was originally designed to address strategic-level questions about forest resources across large geographical areas, there is an increasing interest in using NFI information to assess forest conditions over smaller areas and their changes over time. While the coverage of the NFI is rather sparse, both in time and space, auxiliary variables collected at much finer spatial and temporal resolutions, such as those obtained from remote sensors, can bring dramatic gains in the precision of forest condition estimates. In the present study, we propose a Bayesian approach to combine satellite imagery data, collected at a fine spatial and temporal resolution, with NFI data, more detailed but collected at a much sparser resolution, with the ultimate goal to predict forest conditions at unsampled times and locations.

Petrone, Sonia
Fast approximate Bayesian procedures for mixture models.
Bayesian nonparametric methods have become popular tools with wide application in many fields. However, fast computation remains an open issue. M. Newton et al. (see Newton and Zhang, 1999, Biometrika) proposed a recursive predictive updating rule as a simple and fast approximation of the Bayesian estimate of the mixing distribution in Dirichlet process mixture models. A special case of Newton's algorithm is the quasi-Bayes sequential procedure proposed by Smith and Makov (JRSS,B, 1978). Convergence results have proven the validity of Newton's recursive scheme as a consistent frequentist estimator. However, the original motivation, whether it provides an approximation of a Bayesian procedure, remains open. In this work we address this question. Using the notions of asymptotic exchangeability and conditionally identically distributed sequences, we show that the recursive algorithm does provide an asymptotic approximation of a Bayesian procedure obtained under exchangeability. Beyond the - important - case of mixture models, our study provides a rigorous framework to formalize the idea that, with nowadays abundance of data and pressure for fast computations, a slightly miss-specified but computationally more tractable model may provide an attractive compromise in the trade-off between information and computations efficiency.

Pirathiban, Ramethaa

Iterating on variable selection: Variable choice for better variable selection in ecology

Continual learning is essential in ecology when species distribution models (SDM) are used to project shifts in the geographic range of species under climate change. Variable selection plays an important role in such models, since variables measuring different aspects of habitat can be very informative about the underlying physiology and ecology of a species. Ecologists often have easy access to automated variable selection methods, but these neither facilitate nor incorporate ecological understanding, so are not guaranteed to select a best set of variables. We demonstrate a Bayesian approach to improve variable selection over a sequence of studies. This starts with eliciting variable choice (from hundreds of candidates), whilst also acknowledging and capturing an ecological hierarchy of relationships amongst competing variables. Elicitation may use direct ranking, rating on multiple criteria or pairwise comparisons. In a second phase, a Bayesian model analyses this expert `data’ to reveal an underlying expert mental model, regarding the relative importance of variables. In a third phase of Bayesian hierarchical modelling, this forms informative priors for variable selection within ecological themes. Updating knowledge over multiple phases resulted in more useful, accurate and meaningful models, here to explain or predict the distribution of forest trees under climate change.

Polson, Nic

Lasso Meets Horseshoe

The goal of our paper is to survey and contrast the major advances in two of the most commonly used high-dimensional techniques, namely, the Lasso and horseshoe regularization methodologies. Lasso is a gold standard for best subset selection of predictors while the horseshoe is a state-of-the-art Bayesian estimator for sparse signals. Lasso is scalable and fast using convex optimization whilst the horseshoe is a non-convex penalty. Our novel perspective focuses on three aspects, (i) efficiency and scalability of computation and (ii) methodological development and performance and (iii) theoretical optimality in high dimensional inference for the Gaussian sparse model and beyond.

Poynor, Valerie

Bayesian nonparametric modeling for mean residual life regression

The mean residual life function describes the expected remaining lifetime of a subject given survival up to a particular time point. In addition to its practically useful interpretation, it also characterizes the survival distribution. However, this functional has received limited attention in terms of inference methods under a probabilistic modeling framework. We seek to provide general inference methodology for mean residual life regression. We develop a Dirichlet process mixture model for the joint stochastic mechanism of the covariates and survival responses. This approach implies a flexible model structure for the mean residual life of the conditional response distribution, allowing general shapes for mean residual life as a function of covariates given a specific time point, as well as a function of time given particular values of the covariate vector. We extend the mixture model to incorporate dependence across experimental groups. This extension is built from a dependent Dirichlet process prior for the group-specific mixing distributions with common locations and weights that vary across groups. We study properties of the regression models, and discuss methods for prior specification and posterior inference. The methodology is illustrated with simulated data examples, and the model is also applied to a data set comprising right censored survival times.

Pratola, Mathew

Some Thoughts on Model Discrepancy After Hanging Out with Physicists

Computer model calibration experiments enable scientists to combine simulators of real-world processes with observational data to form predictions and solve inverse problems. The typical statistical approach to modeling such experiments is to model observational data as the sum of simulator output given a parameter vector of interest, plus a systemic model discrepancy. The model discrepancy is captured by a statistical model and is supposed to inform one of the real-world process behavior missing from the theoretical model encoded in the simulator. While a powerful idea that ultimately improves predictions, the approach is also fraught with challenges such as the well-known issue of identifiability. In this talk I will relate some new perspectives on model discrepancy arising from recent conversations with physics collaborators and in particular the notion of effective field theories.

Puelz, David

Regret-Based Selection for Sparse Dynamic Portfolios

This paper considers portfolio construction in a dynamic setting. We specify a loss function comprised of utility and complexity components with an unknown tradeoff parameter. We develop a novel regret-based criterion for selecting the tradeoff parameter to construct optimal sparse portfolios over time.

Quiroz, Matias

Exact subsampling MCMC

Speeding up Markov Chain Monte Carlo (MCMC) for datasets with many observations by data subsampling has recently received considerable attention in the literature. Most of the proposed methods are approximate, and the exact solutions proposed so far have either been documented to be highly inefficient or to be limited to small dimensional models. We propose an MCMC method for estimating expectations of any function of the parameters by combining pseudo-marginal MCMC with data subsampling and an importance sampling correction for occasionally negative likelihood estimates. Our algorithm is based on a product of Poisson estimators, which incorporates computationally cheap control variates to lower variability and provides an unbiased but not necessarily positive estimate of the likelihood. In each iteration, we update the random observations used for estimation in a subset of products, thereby inducing a controllable positive correlation between the log of the absolute value of the estimates at the current and proposed draw in the Metropolis-Hastings ratio. This allows highly variable estimators of the log of the estimated likelihood without adversely affecting the sampling efficiency of the Markov chain. We provide analytically derived heuristic guidelines to select the tuning parameters for our estimator by optimizing a computational time that takes into account the MCMC sampling inefficiency, the computational cost, and the probability of an estimator being positive. We illustrate the usefulness of the method in a logistic regression and find a dramatic improvement compared to non-subsampling MCMC and, in particular, a popular exact subsampling approach recently proposed in the literature.

Raghunathan, Trivellore

Integration of Information from Multiple Data Sources on Correlated Variables at Different Levels of Aggregation for Small Area Estimation

Demands for small area estimates are ever increasing and are useful for the local policy evaluation and implementation. Increasing concerns about privacy and confidentiality is preventing agencies from providing data at the desired level of geography. This paper develops procedures for combining information from multiple data sources that provide data at different levels of aggregation on correlated variables. The levels of aggregation may be nested (such as counties within States and States) or non-nested (Standard Error Computing Units, Census Tracts, Counties, Zip code etc.). The data sources may be from sample survey or Non-survey. The procedures are motivated from a Bayesian perspective and using the missing data framework. Several examples are used as illustration and the sampling properties of the small area estimates are evaluated using simulated data sets. The bias and loss of efficiency are compared to the estimates derived from the ideal and the same level of aggregation from every data source.

Rainforth, Tom

Inference Trees: Adaptive Inference with Exploration

We introduce inference trees (ITs), a new class of adaptive inference methods that incorporate ideas from Monte Carlo tree search to allocate computation in a manner that balances exploration and exploitation, ensures consistency, and alleviates common pathologies in existing adaptation strategies. ITs adaptively sample from hierarchical partitions of the parameter space, while simultaneously learning these partitions in an online manner. This enables ITs to not only identify regions of high posterior mass, but also maintain uncertainty estimates to track regions where significant posterior mass may have been missed. ITs can be combined with any inference technique that provides a consistent estimate of the marginal likelihood, but are particularly effective when combined with sequential Monte Carlo. This allows long-range dependencies to be captured, leading to improvements beyond those from proposal adaptation alone.

Ranganath, Rajesh

Implicit Models and Posterior Approximations

Probabilistic generative models tell stories about how data were generated. These stories uncover hidden patterns (latent states) and form the basis for predictions. Traditionally, probabilistic generative models provide a score for generated samples via a tractable likelihood function. The requirement of the score limits the flexibility of these models. For example, in many physical models we can generate samples, but not compute their likelihood --- such models defined only by their sampling process are called implicit models. In this talk, I will present a family of implicit models that combine hierarchical Bayesian models with deep models. I will detail a posterior approximation algorithm for this class of models based on variational inference. The accuracy of variational inference hinges on the expressivity of the approximating family --- our variational approximations make use of implicit distributions to improve the quality of the posterior approximations. Along the way, I will explore models for symbol generation and regression.

Regazzini, Eugenio

On Gini's Empirical Bayes Methodology

The talk introduces novel extensions of an empirical Bayes method (EBM) proposed in a pioneering work by Gini, in 1911, for a sequence of events conditionally iid, given a parameter with a partially unknown prior distribution of the Beta type. Pompilj (1951) extended this method to an analogous sequence of {0,1,...,d}-valued vectors with initial distribution of the Dirichlet type. Later, Forcina (1981) brought the Gini paper out of obscurity, by simply mentioning a remote affinity with the more familiar EBM proposed by Robbins in the 1950s. As a matter of fact, the two approaches are significantly different, which justifies the effort to extend the Gini EBM to more general statistical situations, while retaining its original distinguishing features. First, an extension, of a parametric nature, to conditionally iid observations, distributed according to a general exponential model, will be illustrated. Secondly, we will show that the representations used by both Gini and Pompilj, as bases for their EBMs, hold true for any prior distribution and, then, rely on a suitable approximation procedure in order to achieve extensions to nonparametric settings.

Renfro, Lindsay

Bayesian Adaptive Trial Design to Address Clinical Endpoint and Predictive Biomarker Uncertainty

Randomized clinical trials are the cornerstone of evidence-based medicine and the gold standard for establishing causal relationships between new treatments and improved patient outcomes. However, as diseases like cancer are increasingly understood on a molecular level, clinical trials that are designed for too-general patient populations are often unable to reveal subpopulations where a therapy is more or less effective. Uncertainty regarding the best clinical endpoint(s) compounds the challenge to reaching early and precise conclusions regarding treatment benefit, particularly when different patient subgroups may respond to treatment via different mechanisms of action, thereby “washing out” a treatment effect measured from one endpoint and across all patients. We develop a novel randomized clinical trial design capable of detecting efficacy through multiple endpoints (e.g., binary and time-to-event) and in the presence of possible patient subpopulations (e.g., biomarker-defined). This design allows patients from different subgroups to respond to treatment via different mechanisms of action (i.e., through different endpoints), and further allows for early stopping for efficacy or futility, either overall or within a marker-based subgroup. We derive the operating characteristics of our design framework and evaluate its performance compared to traditional group-sequential designs assuming a single primary endpoint and homogeneous patient population.

Rice, Ken

Knowing the signs: decision-based arguments for significance tests

The Bayes/non-Bayes controversy about “reject/accept” hypothesis testing is well-known. But “reject/say nothing” significance testing, is much less controversial. In this paper we show how significance tests are in fact the only acceptable actions, for a specified class of decision problems – and the breadth of this class makes this a compelling foundation for many statistical testing developments. We initially present the basic (and normative) decision-theoretic justification for one-sided significance tests, including direct motivations for significance thresholds “alpha”, but also show how the same ideas lead naturally to two-sided tests, more finely-graded decisions, Bayes Factors, credible intervals, adjustments for multiple tests, and risk-assessment measures including Bayarri et al’s rejection ratio (J Math Psych, 2016). Closely-related dual ideas also lead to a novel interpretation of p-values/tail areas. All of this work provides methods with closely-aligned Bayes/non-Bayes interpretations.

Riebler, Andrea Ingeborg

Distributing the variance: Priors for hierarchical models

In Bayesian hierarchical models it is common to define priors for each model component separately. However, the more components the model includes and the deeper the hierarchy goes, the more difficult it gets for the applied scientist to provide prior information about the relative effect of each model component. We propose to use that in most applied studies a natural intuition/prior guess about the scale of the linear predictor exists based on expert knowledge. The penalised complexity framework is used to control the overall variance of the linear predictor and to control how each model component contributes toit. In this way the global graphical structure of the underlying model is accounted for when building priors. For the linear model this approach is related to putting a prior on the coefficient of determination R^2 and the residual variance, which generalises for more complex models. We will assess the new priors using different models and compare their performance and sensitivity to commonly used priors.

Rios Insua, David

Adversarial Statistical Decision Theory

In the talk I shall describe issues in relation with Adversarial Statistical Decision Theory (ASDT). Statistical Decision Theory (SDT) provides foundations to much of Bayesian Analysis. ASDT adds an adversary which perturbates the standard STD to its benefit and provides foundations for Adversarial Risk Analysis. I shall frame first various ASDT variants and then outline issues in relation with point estimation, hypothesis testing and classification.

Risser, Mark

Spatially-Dependent Multiple Testing Under Model Misspecification

The Weather Risk Attribution Forecast (WRAF) is a forecasting tool that uses output from global climate models to make simultaneous attribution statements about whether and how greenhouse gas emissions have contributed to extreme weather across the globe. However, in conducting a large number of simultaneous hypothesis tests, the WRAF is prone to identifying false “discoveries.” A common technique for addressing this multiple testing problem is to adjust the procedure in a way that controls the proportion of true null hypotheses that are incorrectly rejected, or the false discovery rate (FDR). Unfortunately, generic FDR procedures suffer from low power when the hypotheses are dependent, and techniques designed to account for dependence are sensitive to misspecification of the underlying statistical model. In this paper, we develop a Bayesian decision theoretic approach for dependent multiple testing and a nonparametric hierarchical statistical model that flexibly controls false discovery and is robust to model misspecification. We illustrate the robustness of our procedure to model error with a simulation study, using a framework that accounts for generic spatial dependence and allows the practitioner to flexibly specify the decision criteria. Finally, we apply our procedure to several seasonal forecasts and discuss implementation for the WRAF workflow.

Rockova, Veronika

Theory for BART

The remarkable empirical success of Bayesian additive regression trees (BART) has raised considerable interest in understanding why and when this method produces good results. Since its inception nearly 20 years ago, BART has become widely used in practice and yet, theoretical justifications have been unavailable. To narrow this yawning gap, we study estimation properties of Bayesian trees and tree ensembles in nonparametric regression (such as the speed of posterior concentration, reluctance to overfit, variable selection and adaptation in high-dimensional settings).Our approach rests upon a careful analysis of recursive partitioning schemes and associated sieves of approximating step functions. We develop several useful tools for analyzing additive regressiontrees, showing their optimal performance in both additive and non-additive regression. Our results constitute a missing piece of the broader theoretical puzzle as to why Bayesian machine learningmethods like BART have been so successful in practice.

Rodriguez, Abel

Dynamic models for network models on time-varying populations using

Network data, in which observations correspond to dyads capturing interactions across a set of nodes, is common in social and biological sciences, as well as in engineering. A specially popular class of models for network data is the class of Stochastic Blockmodels (SBMs). SBMs, which can be considered extension of classical mixture models to the context of network data, partition the set of nodes into structurally equivalent clusters with common interactionprobabilities. In this talk we consider nonparametric versions of the classical SBM that can be used to model a time series of network observations in which not only the network structure, but also theset of nodes, evolve slowly over time. The models we discuss are motivated by applications to financial market transactions, in which traders (the nodes of the network) might enter or leave themarket at any time. The models we consider use fragmentation and coagulation processes to model the evolution of the community structure over time, and an autoregressive process to model theevolution of the interaction probabilities. Furthermore, the network structure and the birth/death of nodes are jointly modeled through the community indicators.The models are illustrated using both simulated and real datasets.

Rosenthal, Jeff

Bayesian for the General Public

Jeff is author of "Struck by Lightning: The Curious World of Probabilities", a collection of stories that explain probability in an engaging way, accessible to a wide range of readers. It is described as a "book about probability and randomness in everyday life, intended to inform and entertain readers without requiring any mathematics background". Since Bayesian thinking is founded on an understanding of probability, which is not often well taught in elementary courses on Data Analysis, this book fills an important gap. Here, Jeff tells Bayesians about some of the rhetorical techniques he has used to explain statistics in a way that is entertaining, and even, fun. He often relies on amusing anecdotes and stories, but also used a drama, via a "film noir comedic spoof in which a 'Probabilistic Private Investigator' named Ace Spade solves crimes using probability theory". He will also talk about the balance of fun and learning, by providing readers with a mock final exam to help readers ""show off"" their knowledge of probability.

Rosner, Gary

Bayesian Methods in Regulatory Science

Perhaps nowhere is the need for agreement on a routine and robust method of evaluating statistical significance greater than in regulatory science. Statististicians in charge of drug and device approvals require methods that are straightforward to explain and implement, so that applicants feel they are being treated equitably. This talk presents several aspects of the development process for new therapies where Bayesian methods have found acceptance, and others where resistance persists. We look at a vision of the drug development framework and the way Bayesian inference fits naturally within it. We also advocate greater application of Bayesian decision theory in regulatory science.

Rossell, David

High-dimensional posterior model probabilities

An important property for Bayesian model selection is that the posterior probability of the data-generating model (or that amongst the considered models closest to it) converges to 1, and that the corresponding convergence rate is fast. This property suffices to guarantee model selection consistency as typically proven for frequentist methods, but is in fact a much stronger result. While in finite dimensions such consistency and rates have been established for a long time and are well-understood, the results for high-dimensional cases (where the number of parameters p grows with the sample size n) are less complete. We present a novel theoretical framework to establish consistency and rates. It is based on first principles, intuitively straightforward and its basic arguments hold in complete generality and potentially under model misspecification. We apply the framework to canonical variable selection in linear regression and compare common prior specifications, obtaining interesting insights regarding sparsity/sensitivity tradeoffs, e.g. the effects of sparse model space priors and the prior variance.

Rousseau, Judith

Asymptotic behaviour of credible regions

The reknown theorem of Bernstein von Mises in regular finite dimensional models has numerous interesting consequences, in particular it implies that a large class of credible regions are also asymptotically confidence regions, which in turns imply that different priors lead to the same credible regions to first order. Unfortunately the Bernstein von Mises theorem does not necessarily hold in high or infinite dimensional models and understanding the asymptotic behaviour of credible regions is much more involved. In this talk I will describe what are the new advances that have been obtained over the last 6 years or so in this area and I will in particular discuss some interesting phenomena which have been exhibited in high dimensional models. In particular I will discuss the behaviour of model choice types of priors, encountered for instance in mixture models with unknown number of components, in regression models with a large number of covariates etc... where we can show that in a significant number of cases these priors tend to over penalize (or over smooth), leading to non robust confidence statements.

Roy, Daniel

Computing Nonvacuous Generalization Bounds for Deep (Stochastic) Neural Networks with Many More Parameters than Training Data

Gintare K Dziugaite (University of Cambridge)

One of the defining properties of deep learning is that models are chosen to have many more parameters than available training data. In light of this capacity for overfitting, it is remarkable that simple algorithms like SGD reliably return solutions with low test error. One roadblock to explaining these phenomena in terms of implicit regularization, structural properties of the solution, and/or easiness of the data is that many learning bounds are quantitatively vacuous in this "deep learning" regime. In order to explain generalization, we need nonvacuous bounds. We return to an idea by Langford and Caruana (2001), who used PAC-Bayes bounds to compute nonvacuous numerical bounds on generalization error for stochastic two-layer two-hidden-unit neural networks via a sensitivity analysis. By optimizing the PAC-Bayes bound directly, we are able to extend their approach and obtain nonvacuous generalization bounds for deep stochastic neural network classifiers with millions of parameters trained on only tens of thousands of examples. We connect our findings to recent and old work on flat minima and MDL-based explanations of generalization.

Ruggiero, Matteo

Accelerating Bayesian inference for hidden Markov models via duality

Filtering hidden Markov models can be seen as performing sequential Bayesian inference on the hidden state of a latent dynamic signal, which can be of parametric or nonparametric form, given data collected at discrete times. Besides a handful of model, the posterior distribution is not generally available in closed form due to the infinite series representation of the transition kernel of the signal. Some recent results exploit duality for Markov processes to obtain an analytic recursion which computes the filtering distribution in closed form for certain classes of signals related to gamma and Dirichlet finite and infinite dimensional structures. However, this entails computable filters whose nominal cost grows polynomially in the number of observations, which can make the associated algorithms impractical for large dimensional hidden spaces or very long observation sequences.
Here we exploit the above results aiming at attaining fast inference for signals driven by Cox-Ingersoll-Ross and Wright-Fisher processes and for their nonparametric counterpart. We use the mixture representation of the filtering distribution obtained via duality to approximate the true posterior via different pruning techniques, which select the most relevant components for inference, do not rely on stochastic truncation and are fast to implement. We assess the performance of these strategies in terms of accuracy, speed and prediction, which can be benchmarked against the exact solution.

Sanso, Bruno

Bayesian Non-parametric methods for Renewal Processes

We propose a flexible BNP method for renewal processes that is built by assuming that the inter-arrival times correspond to a mixture of Erlang densities. The components of the mixture have a common scale parameter, and have weights that are defined through an underlying distribution function that corresponds to a Dirichlet process prior. The resulting model can accommodate clustering and declustering processes and provides a flexible way of estimating the K function that uses analytically available expression for the Laplace transform of relevant functions. The model is computationally attractive, as it allows a seamless treatment of a likelihood that includes a complex normalizing constant. This is achieved using a data augmentation approach that provides an efficient MCMC sampling scheme. We illustrate our proposed model with synthetic data as well as data that that have previously appeared in the literature.

Savitsky, Terrance

Fully Bayesian Estimation Under Informative Sampling

Bayesian estimation is increasingly popular for performing model based inference to support policymaking. These data are often collected from surveys under in-formative sampling designs where subject inclusion probabilities are designed to be correlated with the response variable of interest. Sampling weights constructed from marginal inclusion probabilities are typically used to form an exponentiated pseudo likelihood that adjusts the population likelihood for estimation on the sample. We propose an alternative adjustment based on a Bayes rule construction that simulta-neously performs weight smoothing and estimates the population model parameters in a fully Bayesian construction. We formulate conditions on known marginal and pairwise inclusion probabilities that define a class of sampling designs where L1 consistency of the joint posterior is guaranteed. We compare performances between the two approaches on synthetic data. We demonstrate our method on an application concerning the National Health and Nutrition Examination Survey exploring.

Schmidt, Alexandra

Spatial confounding in hierarchical models

In spatial generalized linear mixed models, covariates that are spatially smooth are often collinear with spatial random effects. This collinearity, which affects the bias and precision of the regression coefficients, is known in the spatial statistics literature as spatial confounding. A potential remedy is to assume a restricted spatial regression model wherein the spatial random effects are constrained to be orthogonal to the fixed effects of the model. We discuss the problem of spatial confounding in hierarchical models and investigate its effects in an analysis of student performance in the 2013 Brazilian Mathematical Olympics for Public Schools. This is joint work with Widemberg S. Nobre and Joao B. M. Pereira, from the Federal University of Rio de Janeiro, Brazil.

Schmidt, Mikkel

Bayesian modelling of surface-enhanced Raman spectroscopy

Surface-enhanced Raman spectroscopy (SERS) is a sensitive technique for detection and estimation of molecules at low concentration. Recent advances in nanotechnology allow large-area mapping of SERS data. With an appropriate statistical analysis, this can potentially significantly improve detection and estimation at very low concentrations. We present a Bayesian statistical model for fitting large SERS maps based on a parametric spectral peak model, a flexible baseline model, and a spatial model for hotspot formation. We derive an efficient inference procedure using variational approximation to allow fast statistical analysis of large-scale SERS data. The proposed model and algorithm is a step towards improving the sensitivity of SERS in detection of trace amounts of molecules.

Scott, James

Bayes-optimal post-selection inference under sparsity.

A large body of recent Bayesian work has focused on the question of how to find sparse signals. Less work, however, has been done on thenatural follow-up question: how to make sensible inferences for themagnitude of those signals once they've been found. Ordinary Bayesian credible intervals are not necessarily appropriate for this task: in many circumstances, they suffer from selection bias, owing to the fact that the target of inference is chosen adaptively. There are manypurely frequentist proposals for addressing this problem. But thesetypically require sacrificing the benefits of "information borrowing"inherent to Bayesian modeling, resulting in confidence intervals thatare needlessly wide. On the flip side, there are also Bayesianproposals for addressing this problem, most notably that of Yekutieli(2012) and George and Yekutieli (2012), who construct selection-adjusted posterior distributions. However, the resulting credible intervals have poor frequentist performance: for nearly all values of the underlying parameter, they fail to exhibit the correct nominal coverage. Thus there is an unmet need for approaches to inference that: 1) correctly adjust for selection, 2) incorporate the benefits of a priordistribution, and 3) maintain exact frequentist coverage. We address this gap by proposing an approach for constructing Bayes-optimal post-selection confidence sets that maintain correct coverage, uniformly across the whole parameter space. Our approach is general, but we illustrate it in three commonly encountered settings: a simple random-effects model, the sparse normal-means problem, and the problem of smoothing a noisy signal over a discrete lattice.

Shaby, Benjamin

Hierarchical Scale Mixtures for Flexible Spatial Modeling

Scale mixtures of Gaussian processes have emerged as desirable candidates for modeling extremal phenomena in space. They are intuitive, simple to describe constructively, and flexible in the types of extremal dependence that they they can represent. Inferences for these models using censored likelihoods has been limited to very small datasets due to the presence of a high-dimensional Gaussian integral that must be evaluated numerically. Rather than integrating over a latent Gaussian process, we condition on it, expressing the model hierarchically. This way, we allow Markov chain Monte Carlo to do the hard integration, and open the door to inference on much larger datasets than were previously possible.

Shand, Lyndsay

Spatially Varying Autoregressive Models for Prediction of New HIV Diagnosis

In demand of predicting new HIV diagnosis rates based on publicly available HIV data that is abundant in space but has few points in time, we propose a class of spatially varying autoregressive (SVAR) models compounded with conditional autoregressive (CAR) spatial correlation structures. We then propose to use the copula approach and a flexible CAR formulation to model the dependence between adjacent counties. These models allow for spatial and temporal correlation as well as space-time interactions and are naturally suitable for predicting HIV cases and other spatio-temporal disease data that feature a similar data structure. We apply the proposed models to HIV data over Florida, California and New England states and compare them to a range of linear mixed models that have been recently popular for modeling spatio-temporal disease data. The results show that for such data our proposed models outperform the others in terms of prediction.

Shinota, Shinichiro

Data Fusion Approaches for Species Distribution

Species distribution models are often constructed for presence-absence data that are composed of binary responses well designed locations in planned surveys and lend themselves to logistic and probit regression models for statistical inference. On the other hand, presence-only records may provide further information on the species distribution but often suffer from unknown biases due to haphazard collection schemes. In this talk, we propose a data fusion approach of presence-absence and presence-only data then discuss modeling for both data from preferential sampling perspectives.

Si, Yajuan

Bayesian hierarchical weighting adjustment and survey inference

We combine Bayesian prediction and weighted inference as a unified approach to survey inference. The general principles of Bayesian analysis imply that models for survey outcomes should be conditional on all variables that affect the probability of inclusion. We incorporate the weighting variables under the framework of multilevel regression and poststratification, as a byproduct generating model-based weights after smoothing.

We investigate deep interactions and introduce structured prior distributions for smoothing and stability of estimates. The computation is done via Stan and implemented in the open source R package "rstanarm" ready for public use. Simulation studies illustrate that model-based prediction and weighting inference outperform classical weighting.

We apply the proposal to the New York Longitudinal Study of Wellbeing. The new approach generates robust weights and increases efficiency for finite population inference, especially for subsets of the population.

Simoni, Anna

Bayesian Estimation and Comparison of Conditional Moment Models

In this paper we consider an extension of the Bayesian exponentially titled empirical likelihood framework developed in Chib, Shin and Simoni (2016) for moment condition models to conditional moment models. The basic idea is to transform the conditional moments in a sequence of unconditional moments by using a vector of approximating functions with increasing dimension. We consider both correctly and incorrectly specified conditional moment models and demonstrate that in both cases the Bayesian posterior distribution satisfies the Bernstein-von Mises theorem, subject to a rate condition on the number of approximating functions. We also develop an approach based on marginal likelihoods for comparing different conditional moment restricted models and establish the model selection consistency of the marginal likelihood. Several examples are used to illustrate the framework and results.

Sinha, Debajyoti

Latent Variable Based Semiparametric Regression Model and Bayesian Analysis of Skewed Multivariate Data

For many real-life studies with skewed multivariate responses, the level of skewness and association structure assumptions are essential for evaluating the covariate effects on the response and its predictive distribution. We present a novel semiparametricmultivaria<wbr>te model leading to a theoretically justifiable semiparametric <wbr>Bayesian analysis of multivariate skewed responses. Similar to multivariate Gaussian densities, this multivariate model is closed under marginalization, allows a wide class of multivariate associations, and has meaningful physical interpretations of skewness levels and covariate effects on the marginal density. Compared to existing models, our model enjoys several desirable practical properties, including Bayesian computing via available software, and assurance of consistent Bayesian estimates of parameters and the nonparametric error density under a set of plausible prior assumptions. We illustrate the practical advantages of our methods over existing parametric alternatives via application to a clinical study to assess periodontal disease and through a simulation study.

Smith, Richard

Influence of climate change on extreme wheather events

The exceptionally active 2017 hurricane season has led to considerable speculation over the role of human-induced climate change on these and other extreme weather events. Here, we address these questions with specific references to the extreme precipitations due to Hurricane Harvey in Houston and surrounding areas. We construct a new dataset consisting of extreme precipitation events over the entire Gulf of Mexico region from 1949-2017; Harvey is by far the most extreme event over this time frame but an analysis of other high precipitation events shows a clear association with increases in both sea surface temperature and atmospheric carbon dioxide. Models are developed for the relative risk of a Harvey-type event in a world subject to anthropogenic climate forcings compared with one that is not, and for the projected increases in probabilities of such events in the future. The statistical methods are based on extreme value theory, using Bayesian methods to assess uncertainties, and we address various challenges in the combination of observational and climate model data. The topic of this talk is closely associated with the invited paper session, “Bayesian Methods for Detection and Attribution of Climate Change.”

Söhl, Jakob

Nonparametric Bernstein-von Mises theorems for discretely observed compound Poisson processes

We show a Bernstein-von Mises theorem for discretely observed compound Poisson processes. The compound Poisson processes are observed at low frequency, i.e., the distance between the observations is fixed while more and more observations are accrued as the observation time increases. The Lévy density is modeled by a wavelet series prior. We show contraction rates for the posterior distribution of the Lévy density that are optimal up to log factors in sup-norm. In addition, we prove a Bernstein-von Mises theorem for the distribution function of the Lévy density. The results are tied close together as we first derive suboptimal convergence rates to prove the Bernstein-von Mises theorem, which in turn is used to show optimal convergence rates. This is joint work with Richard Nickl.

Soyer, Refik

Adversarial Issues in Reliability: An ARA Perspective

There are problems in reliability analysis and life testing that may involve two or more actors with competing interests. These problems with adversarial components can be set-up as games and can be solved using game theory methods that are based on “common knowledge” assumption. In this paper, we present an alternative approach based on adversarial risk analysis framework to deal with such problems. We illustrate the framework through acceptance sampling and life testing problems.

Spencer, Simon

Bayesian inference for protein secondary structure from circular dichroism data

Circular dichroism spectroscopy is an important tool for determining the structural characteristics of biomolecules, particularly the secondary structure of proteins. We demonstrate the importance of correctly capturing the covariance structure within a measured spectrum. We propose a Bayesian model that estimates this covariance structure and quantifies the uncertainty associated with the inferred secondary structures and characteristic spectra associated with each secondary structure type. Furthermore, using tools from Bayesian model selection, we determine the best secondary structure classification scheme and illustrate a technique for comparing whether or not two or more measured protein spectra share the same secondary structure. These advances would not have been possible using existing approaches that do not account for correlation and uncertainty appropriately. Finally, our approach can be easily extended to combine information from different spectroscopic techniques.

Stangl, Dalene

How Teaching Statistics from a Bayesian Perspective Keeps Education Transformative

David Hansen, a Professor at Columbia University, has a wonderful speech available on YouTube titled “Is Education Possible Today?” The entire talk discusses why the answer to that question is not obvious and explains why in education the question is more important than any answer or response. For transformative education, it is the necessity of the question and the importance of keeping the question open---one must follow the course why the answers are not “the" answer. Education is about “keeping questions open.” In this talk I propose that statistics education from the very first course, while being transitioned away from a “math-heavy” subject to a “computer science-heavy” subject, should also be transitioning to a “philosophy-heavy” subject, a Bayesian philosophy-heavy subject. A subject of how is it via quantitative information and thinking, that we can come to know, to assess the uncertainty in our knowing, and to make decisions based on what we know accounting for our uncertainty.

Starling, Jennifer

Functional response regression with funBART

Bayesian Additive Regression Trees (BART) has been shown to be an effective framework for modeling nonlinear regression functions, with strong predictive performance in a variety of contexts. The BART prior over a regression function is defined by independent prior distributions on tree structure and leaf or end-node parameters. Leaf parameters are typically modeled as independent and identically distributed from a carefully-scaled Gaussian distribution. While highly effective in practice, functions sampled from this prior are not smooth in any of the covariates. In many contexts it is desirable to enforce or encourage a degree of smoothness in “special” covariates such as a time index.

We develop a new version of the BART prior for modeling a function that evolves smoothly over time given other covariates. We demonstrate the utility of this approach by applying our model to a timely women’s health problem, where the outcome of interest is measured at different gestational ages for each subject, conditional on covariates. We discuss the benefits of this approach in a variety of women’s health and obstetrics modeling problems where gestational age is a typical covariate, as well as potential applications in a range of other disciplines.

Startz, Dick

Bayesian Projection of Carbon Emissions

The recently published Intergovernmental Panel on Climate Change (IPCC) projections to 2100 give likely ranges of global temperature increase in four scenarios for population, economic growth and carbon use. However, these projections are not based on a fully statistical approach. Here we develop a statistically based probabilistic forecast of CO2 emissions and temperature change to 2100. Using data for 1960-2010, including the UN's probabilistic population projections for all countries, we develop a joint Bayesian hierarchical model for Gross Domestic Product (GDP) per capita and carbon intensity. We find that the 90% interval for cumulative CO2 emissions includes the IPCC's two middle scenarios but not the extreme ones. The likely range of global temperature increase is 2.0-4.9°C, with median 3.2°C and a 5% (1%) chance that it will be less than 2°C (1.5°C). Population growth is not a major contributing factor. Our model is not a 'business as usual' scenario, but rather is based on data which already show the effect of emission mitigation policies. Achieving the goal of less than 1.5°C warming will require carbon intensity to decline much faster than in the recent past.

Steorts, Rebecca

Bayesian Entity Resolution for Society Impacts in Machine Learning

Very often information about social entities is scattered across multiple databases. Combining that information into one database can result in enormous benefits for analysis, resulting in richer and more reliable conclusions. Among the types of questions that have been, and can be, addressed by combining information include: How accurate are census enumerations for minority groups? How many of the elderly are at high risk for sepsis in different parts of the country? How many people were victims of war crimes in recent conflicts in Syria? In most practical applications, however, analysts cannot simply link records across databases based on unique identifiers, such as social security numbers, either because they are not a part of some databases or are not available due to privacy concerns. In such cases, analysts need to use methods from statistical and computational science known as entity resolution (record linkage or de-duplication) to proceed with analysis. Entity resolution is not only a crucial task for social science and industrial applications, but is a challenging statistical and computational problem itself. In this talk, we describe the past and present challenges with entity resolution, using novel Bayesian methodological developments under a microclustering framework, with sound theoretical guarantees and computational speeds ups. This work touches on the interdisciplinary research that is crucial to problems with societal impacts that are at the forefront of both national and international news.

Stephenson, Briana

Joint Model-based Clustering of Multivariate and Grouped Data: An Application to Dietary Exposures on Cardiometabolic Markers in the Hispanic Community Health Study/Study of Latinos

Cardiometabolic markers have been found to vary by Hispanic/Latino backgrounds. These differences appear to correspond to diet and the different nutrient and food group intakes that are associated with various backgrounds. Robust Profile Clustering (RPC) provides a dual flexible clustering model, where participants may cluster at two levels: (1) globally, where participants are clustered according to behaviors shared across an overall population, and (2) locally, where individual behaviors can deviate and cluster according to their identified subgroup. We link this clustering model generated from the RPC with a health outcome using a predictive clustering model known as a supervised RPC joint model. We generate the utility of this model for the application of deriving dietary patterns in participants from the Hispanic Community Health Study/Study of Latinos and determining its association to cardiovascular disease.

Stevenson, Oliver

Modelling career trajectories of cricket players using Gaussian processes

Generally speaking, a sportsperson's career follows a typical trajectory. They begin at a young age with some raw but undeveloped ability, which improves over time as the athlete gains experience and participates in specialised training and coaching programmes. Eventually an athlete reaches the peak of their career, after which ability tends to decline until the athlete opts to retire from their chosen discipline. Additionally, fluctuations in ability are common due to factors such as form, injury and other external circumstances. This is particularly relevant in cricket, where many players continually experience peaks and troughs in terms of individual performance, due to the nature of the sport. We fit a Bayesian model that employs Gaussian processes to model how the abilities of New Zealand cricketers change across a playing career. Given the high dimensionality of the model, and for ease of model comparison, nested sampling is used to fit the processes. The model allows for more precise quantification of a player's ability at any given point of their career, than traditional methods. The results provide coaches with more information at their disposal when determining whether or not a player is improving, which has practical implications when comparing players and in team selection policy.

Stingo, Francesco

Bayesian Graphical Modeling of Heterogeneous Data

Heterogeneous data have emerged across many scientific areas such as genomics and imaging studies, one example of interest being observed data collected under various conditions. In such situations, using the pooled data as the basis for inference of a single network may lead to the identification of spurious relationships. We introduce a novel methodology that directly models the heterogeneity in the data and results in multiple networks, as supported by the data. We investigate several approaches under a few types of prior distribution reflecting different degrees of substantive prior knowledge.

Stoehr, Julien

Pre-stored likelihood-free inference

In settings where the likelihood function involves an intractable normalising "constant" (albeit function of the parameter), such as Markov random fields, MCMC solutions that overcome this difficulty are often quite costly and sometimes too costly to run on realistic datasets. We propose in this novel approach a simulation-based resolution relying on a pre-stored reference table that is simulated once for all from the prior predictive distribution as in ABC methods. While this simulation step is costly, the size of the table is much smaller than the size of the intended final sample. We then implement a valid MCMC algorithm iteratively extracting starting values from the ground table that considerably reduces the number of steps towards the (costly) production of pseudo-observations, based on the same substitute to the normalising constant as the exchange algorithm (Murray et al., 2006). Our proposal is illustrated on various Markov random fields.

Streftaris, George

Model assessment and comparison in infectious epidemics

The dynamics of disease spread in epidemics, or system change in ecology, are often described through spatio-temporal compartmental models where important characteristics, such as the distribution of sojourn time in a particular infection state, or the spatial transmission kernel, can be represented using a range of different settings. This leads to questions of model assessment within a certain class of models, which typically suffers from problems related to lack of replication, the incomplete nature of observed data, lack of prior knowledge on parameters and computational challenges. This talk presents approaches to model criticism and evaluation using methodology based on the properties of the so-called Bayesian latent residuals, which become available through data augmentation within Markov chain Monte Carlo estimation schemes. Model diagnostics leading to targeted assessment of model mis-specification are considered and comparison between candidate models is also presented, using a latent likelihood ratio type test that avoids common problems associated with Bayesian model choice.

Stroud, Jonathan

A Bayesian Adaptive Ensemble Kalman Filter for Sequential State and Parameter Estimation

The ensemble Kalman filter (EnKF) is a computational technique for approximate inference on the state vector in spatio-temporal state-space models. It has been successfully used in many real-world nonlinear data-assimilation problems with very high dimensions, such as weather forecasting. However, the EnKF is most appropriate for additive Gaussian state-space models with linear observation equation and without unknown parameters. Here, we consider a broader class of hierarchical state-space models, which includes two additional layers: The parameter layer allows handling of unknown variables that cannot be easily included in the state vector, while the transformation layer can be used to model non-Gaussian observations. For Bayesian inference in such hierarchical state-space models, we propose a general class of extended EnKFs, which approximate inference on the state vector in suitable existing Bayesian inference techniques (e.g., Gibbs sampler or particle filter) using the EnKF or the related ensemble Kalman smoother. Extended EnKFs enable approximate, computationally feasible filtering and smoothing in many high-dimensional, nonlinear, and non-Gaussian spatio-temporal models with unknown parameters. We highlight several interesting examples, including assimilation of heavy-tailed and discrete data, and filtering and smoothing inference on model parameters.

Sussman, Daniel

Decision-Theoretic Aspects of Causal Inference under Network Interference

In fields such as marketing, social science, and public health, designed and natural experiments will frequently occur in a networked environment. In this setting, it is reasonable to assume that the treatment of one individual may impact nearby individuals in a network. We explore a series of assumptions on the set of potential outcomes which parsimoniously allow for this interference in a setting where the network is observed. Our focus is on estimation and we explore Bayesian estimates, optimal design-unbiased estimates, and minimax estimates of direct, interference, and total effects. We compare these estimates via simulations on random graph models and real-world graphs and consider the tradeoffs between them.

Swartz, Tim

Analysis of Substitution Times in Soccer

Analysis of Substitution Times in Soccer: In high level soccer, teams are afforded a maximum of three substitutions during a match of 90 minutes. This talk considers the problem of determining optimal substitution times in soccer. Goals scored (by the trailing/leading team) are modeled as a function of covariates that describe team strength, home team advantage, match time and substitution status. Priors are elicited which provide smoothing on time-varying parameters. For these time-varying parameters, a first-order order autoregressive process is introduced. Other hyperparameters are elicited through the use of gambling odds. The dataset is based on 2,989 goals scored in top soccer leagues. We find that with evenly matched teams, there is a goal scoring advantage to the trailing team during the second half of a match. This is an important finding in terms of defending and attacking tactics. We also observe that there are no discernible times during the second half when there is a benefit due to substitution. The second finding contradicts previous results from the literature.

Szabo, Botond

Uncertainty quantification and model selection for sparse Bayesian methods

We investigate the performance of the two, arguably most popular Bayesian priors for sparsity (the horseshoe prior and the spike-and-slab;prior) for uncertainty quantification and model selection in the sparse multivariate normal means model. We do so in an adaptive setting;without assuming knowledge of the sparsity level (number of signals).;First we derive posterior contraction rates for the horseshoe considering both the hierarchical Bayes method of putting a prior on the;unknown sparsity level and the empirical Bayes method with the sparsity;level estimated by maximum marginal likelihood. Then we investigate the;frequentist coverage properties of the marginal and l2-Bayesian credible;sets and the performance of various model selection methods considered;in the literature. By general theory honest confidence sets cannot adapt in size to an unknown sparsity level. Accordingly the hierarchical and;empirical Bayes credible sets based on the horseshoe prior are not;honest over the full parameter space. We characterise the signals for;which credible balls and marginal credible intervals do give correct;uncertainty quantification. Finally we consider the spike-and-slab prior;for uncertainty quantification and derive similar results to the;horseshoe.The talk is based on joint works with Ismael Castillo, Stéphanie van der Pas and Aad van der Vaart.

Thomas, Owen

Likelihood-free inference by ratio estimation

We consider the problem of parametric statistical inference when likelihood computations are prohibitively expensive but sampling from the model is possible. Several likelihood-free methods have been developed to perform inference in the absence of a likelihood function. The popular synthetic likelihood approach infers the parameters by modelling summary statistics of the data by a Gaussian probability distribution. In another popular approach called approximate Bayesian computation, the inference is performed by identifying parameter values for which the summary statistics of the simulated data are close to those of the observed data. We here present an alternative inference approach that is as easy to use as synthetic likelihood but not as restricted in its assumptions, and that enables automatic selection of relevant summary statistic from a large set of candidates. The basic idea is to frame the problem of estimating the posterior as a problem of estimating the ratio between the data generating distribution and the marginal distribution. This problem can be solved by logistic regression, and including regularising penalty terms enables automatic selection of the summary statistics relevant to the inference task. We illustrate the general theory on toy problems and use it to perform inference for stochastic nonlinear dynamical systems.

Thorarinsdottir, Thordis

Bayesian modelling of cluster point process models

Point process data arises naturally in various fields of science such as biology, ecology, epidemiology, and environmental sciences. However, the point process modelling framework is very involved and inference can often only be performed approximately and with great care. At the same time, a great number of different models are available where the subtle differences between the individual models can be hard to detect. In this talk, we discuss to which extend Bayesian modelling approaches may be applied to the class of cluster process models. Cluster point processes have the following general structure. There is a point process of cluster centres and to each cluster centre is associated a random number of points forming a subsidiary process, where the points in the subsidiary process are distributed about the cluster centre in some specific way. A model for a cluster point process thus consists of three components; a component describing the cluster centre process, a component describing the cluster sizes, and a component describing the distribution of the subsidiary points around the cluster centre, the dispersion process. We consider how Bayesian approaches may be used to perform inferences for all three components, a feat which is often not possible using other inference methods. Recent advances in simulation methods have made these approaches feasible for a number of point process models for which MCMC methods have traditionally been burdensome.

Toulis, Panos

Propensity score methodology in the presence of network entanglement between treatments

In several modern applications of experimental design and causal inference the treatment is not defined on individual units but on pairs or, more generally, on groups of units. For example, teachers may choose pairs of students to teach a new curriculum. We call such treatments entangled because the individual treatment generally depends on a common population quantity. In this paper, we consider that population quantity to be a changing network between units, and the individual treatments to be functions of the network change. Our focus here is on estimating the causal effect of entangled treatments that are endogenous and cannot be directly manipulated. When treatment cannot be manipulated, be it entangled or not, it is necessary to avoid selection bias, commonly through a propensity score methodology. We show that the classical propensity score methodology ignores treatment entanglement, and leads to causal estimates that may be severely biased. To characterize such bias we introduce and analyze a novel similarity function between propensity score models, which we use to quantify misspecication of propensity scores due to entanglement. To avoid the bias we model the temporal network change, and calculate an individual unit's propensity score over all possible network evolutions.

Tran, Minh-Ngoc

Deep Learning: A Statistician's Perspective

We look at Deep Learning (DL) from a statistician's perspective and build a close link between Computer Science and Statistics. This helps to bring the established statistical theory and methodologies into training DL models, and opens up many new flexible models in statistics. In particular, we describe new variants of the conventional generalized linear models for independent data and generalized linear mixed models for panel data, that enjoy a greater prediction power. We develop a high-dimensional Gaussian variational approximation method for training these DL models that uses a parsimonious factor parametrization of the covariance structure. We demonstrate how to implement efficient natural gradient approaches for the variational optimization in this context by leveraging the factor structure and using iterative conjugate gradient methods for solving large linear systems. The methods are illustrated in challenging DL applications. The use of the natural gradient in these applications allows faster and more stable convergence of the variational algorithm.
This is joint work with David Nott and Robert Kohn.

Trippa, Lorenzo

Bayesian Uncertainty Directed Trial Designs

Most Bayesian response-adaptive designs unbalance randomization rates towards the most promising arms with the goal of increasing the number of positive treatment outcomes during the study, even though the primary aim of the trial is different. We discuss Bayesian uncertainty directed designs (BUD), a class of Bayesian designs in which the investigator specifies an information measure tailored to the experiment. All decisions during the trial are selected to optimize the available information at the end of the study. The approach can be applied to several designs, ranging from early stage multi-arm trials to biomarker-driven and multi-endpoint trials. We discuss the asymptotic limit of the patient allocation proportion to treatments, and illustrate the finite-sample operating characteristics of BUD designs through examples, including multi-arm trials, biomarker-stratified trials, and trials with multiple co-primary endpoints.

Turbatu, Laura

Accuracy and validity of posterior distributions using the Cressie-Read empirical likelihoods

The class of Cressie-Read empirical likelihoods are constructed with weights derived at a minimum distance from the empirical distribution in the Cressie-Read family of divergences indexed by γ under the constraint of an unbiased set of M-estimating equations. At first order, they provide valid posterior probability statements for any given prior, but the bias in coverage of the resulting empirical quantile is inversely proportional to the asymptotic efficiency of the corresponding M-estimator. The Cressie-Read empirical likelihoods based on the maximum likelihood estimating equations bring about quantiles covering with O(n^{1}) accuracy at the underlying posterior distribution. The choice of γ has an impact on the variance in small samples of the posterior quantile function. Examples are given for the M-type estimating equations of location and for the quasi-likelihood functions in the generalized linear models.

van Zanten, Harry

On distributed Bayesian computation

Botond Szabo (Leiden University)

Due to the rapidly increasing amount of available information computer scientists and statisticians are facing new challenges to deal with big data problems. One of the most popular and frequently applied approaches to solve this problem are the distributed methods where the data is split into multiple local servers and computations are done locally parallel to each other. This reduces computational time and memory requirement. Then the local machines transmit the outcome of their computations to a global server which aggregates the local results into a global one. In the Bayesian literature various methods were proposed for distributed computational methods with seemingly good practical performance, but with limited theoretical underpinning. In our work we investigate the existing distributed methods in a standard nonparametric setting (the Gaussian white noise model) and compare their theoretical performance, i.e. posterior contraction rate and coverage of credible sets. Then we turn our attention to adaptive techniques, which do not use any information about the regularity of the true functional, but estimate it from the data. We show that the proposed standard methods provide sub-optimal contraction rate and coverage tending to zero for certain underlying functional parameters. We illustrate our theoretical findings with simulations.

Vasishth, Shravan

Type M error in practice: A case study

Treating a result as newsworthy, i.e., publishable, because the p-value is less than 0.05 leads to overoptimistic expectations of replicability. The underlying cause of these overoptimistic expectations is Type M(agnitude) error (Gelman & Carlin, 2014): when underpowered studies yield significant results, the effect size estimates are invariably exaggerated and noisy. Here, we demonstrate the adverse consequences of this statistical significance filter. We do this by carrying out direct replication attempts of published results from a recent paper. Six experiments (self-paced reading and eyetracking, 168 participants in total) show that the published (statistically significant) claims are so noisy that even non-significant results are fully compatible with them. We also demonstrate the stark contrast between these small-sample studies and a larger-sample study (100 participants); the latter yields much less noisy estimates but also a much smaller magnitude of the effect of interest. The small magnitude looks less compelling but is more realistic. We suggest that researchers (i) move their focus away from statistical significance, (ii) attend instead to the precision of their estimates, and (iii) carry out direct replications in order to demonstrate the existence of an effect.</td></tr></tbody></table>

Vidakovic, Brani

Bayes in Wavelet Domains: Anything New?

Wavelets in Statistics are used mostly in two tasks: (i) wavelet shrinkage (signal estimation/denoising/data compression), and (ii) assessment of scaling (fractal indices, multifractals, long memory, self-similarity).
Bayes rules, as shrinkage rules, well formalize what we know about both the signal and noise and perform superbly in the tasks of wavelet shrinkage. In addition, many non-Bayesian shrinkage strategies have a Bayesian interpretation. In the assessment of scaling, often there is strong theoretical and historical evidence of what scaling indices should be. For example, turbulent signals have the Hurst exponent in the neighborhood of 1/3, efficient market time series should scale as Brownian motions, and many more. This information can (and should) be incorporated into the inference. When fractal/multifractal descriptors are used in tasks of machine learning, the instances where the existing prior information is incorporated typically lead to procedure improvements. In this talk we overview the current status of the Bayesian models in the wavelet domain and present several contributions by the speaker and his team on the tasks mentioned in (i) and (ii).

Villa, Cristiano

On a Global Objective Prior from Scoring Rules

We use proper scoring rules to introduce a class of objective prior which provides a constant value for the score at each parameter value within the parameter space. We can also derive the class of prior as solving the Euler--Lagrange equations when seeking to minimise an amount of information in the prior density. The thinking here is that we are <span>interested in the information contained in the prior density rather than <span>the information the prior density holds for a parameter

Vitelli, Valeria

The Bayesian Mallows model for analysing ranks and preference data: from recommendation systems to genomics

Ranking and comparing items is crucial for collecting information about preferences in many areas, from marketing to politics. The Mallows rank model is among the most successful approaches to analyse rank data, but its computational complexity has limited its use in many relevant applications. We develop new computationally tractable methods for Bayesian inference in Mallows models, capable of performing inference on the consensus ranking of the items, also when based on partial rankings, such as top-k items or pairwise comparisons. When assessors are many or heterogeneous, we propose a mixture model for clustering them in homogeneous subgroups, with cluster-specific consensus rankings. We make probabilistic predictions on the class membership of assessors based on their ranking of just some items, and predict missing individual preferences, as needed in recommendation systems. We also show that the use of ranks can effectively overcome the computational complexities related to combining genomic datasets, with various purposes.

Volfovsky, Alexander

Designs for estimating the treatment effect in possibly uncertain networks with interference

We introduce new, easily implementable designs for drawing causal inference from randomized experiments on networks with interference. Inspired by the idea of matching in observational studies, we introduce the notion of considering a treatment assignment as a ``quasi-coloring" on a graph. Our idea of a perfect quasi-coloring strives to match every treated unit on a given network with a distinct control unit that has identical number of treated and control neighbors. We demonstrate the good performance of naive estimators under this design for a wide range of interference functions as well as in the presence of homophily. In settings where the network is only partially observed or is observed with uncertainty, we leverage the uncertainty capture in a Bayesian model of the network to improve our designs.

Wagner, Helga

Bayesian Effect Fusion in Regression Models

Sparse modelling and variable selection is one of the most important issues in regression type models, as in applications often a large number of covariates on comparably few subjects are available. Many methods have been developed, which allow to identify regressors with a non-negligible effect. These are however not appropriate for a categorical covariate, where the effect is captured by a group of level effects and sparsity cannot only be achieved by excluding single irrelevant effects or the whole group of level effects but also by fusing levels which have essentially the same effect on the response. We present two different specifications of the prior on level effects that encourage sparsity by selection and effect fusion:The first is a multivariate Normal distribution with a precision matrix that allows for either almost perfect or almost zero dependence of level effects. As an alternative we consider a sparse finite mixture prior where a spike at zero is combined with a location mixture of spiky components to allow clustering of level effects. For both priors Bayesian inference is feasible by MCMC methods.

Wakefield, Jon

Bayesian time series SIR models with application

The modeling of infectious disease count data over time and space is an important endeavor as it allows an understanding of disease transmission and in particular demographic contributions to spatial heterogeneity in the force of infection. The time series susceptible-infectious-removed (TSIR) model has received a great deal of attention and has been used with a wide variety of infectious diseases including measles and hand, foot and mouth disease. In this talk I will present a spatial version of the TSIR model and examine various aspects of the model including the use of different neighborhood structures (including a so-called gravity model), how to model under-reporting and the use of random effects to model spatial heterogeneity. The models will be applied to measles data.

Wang, Xiaojing

Learning Semiparametric Regression with Missing Covariates Using Gaussian Processes Models

Missing data often appear as a practical problem while applying classical models in the statistical analysis. In this paper, we consider a semiparametric regression model in the presence of missing covariates for nonparametric components under a Bayesian framework. As it is known that Gaussian processes are a popular tool in nonparametric regression because of their flexibility and the fact that much of the ensuing computation is parametric Gaussian computation. However, in the absence of covariates, the most frequently used covariance functions of a Gaussian process will not be well defined. We propose an imputation method to solve this issue and perform our analysis using Bayesian inference, where we specify the objective priors on the parameters of Gaussian process models. Several simulations are conducted to illustrate effectiveness of our proposed method and further, our method is exemplified via two real datasets, one through Langmuir equation, commonly used in pharmacokinetic models, and another through Auto-mpg data taken from the StatLib library.

Wang, Yixin

Frequentist Consistency of Variational Bayes

David Blei (Columbia University) A key challenge for modern Bayesian statistics is how to perform scalable inference of posterior distributions. To address this challenge, variational Bayes (VB) methods have emerged as a popular alternative to the classical Markov chain Monte Carlo (MCMC) methods. VB methods tend to be faster while achieving comparable predictive performance. However, there are few theoretical results around VB. In this paper, we establish frequentist consistency and asymptotic normality of VB methods. Specifically, we connect VB methods to point estimates based on variational approximations, called frequentist variational approximations, and we use the connection to prove a variational Bernstein–Von Mises theorem. The theorem leverages the theoretical characterizations of frequentist variational approximations to understand asymptotic properties of VB. In summary, we prove that (1) the VB posterior converges to the Kullback-Leibler (KL) minimizer of a normal distribution, centered at the truth and (2) the corresponding variational expectation of the parameter is consistent and asymptotically normal. As applications of the theorem, we derive asymptotic properties of VB posteriors in Bayesian mixture models, Bayesian generalized linear mixed models, and Bayesian stochastic block models. We conduct a simulation study to illustrate these theoretical results.

Watanabe, Toshiaki

Bayesian Analysis of Time-Varying Heterogeneous Autoregressive Models

Realized volatility (RV), which is the sum of the squared intraday returns, has attracted the attention of financial econometricians as an accurate estimator of volatility. One of the most widely used models to describe the dynamics of RV is the heterogeneous autoregressive (HAR) model where RV is modelled as a linear function of the past daily, weekly and monthly RVs. This model is known to perform well in volatility forecasts.

This article extends the HAR model to the time-varying HAR model where the values of the constant term, the coefficients and the error variance may change over time. It is assumed that the constant term and the coefficients follow random-walks and the log of the error variance follows a stationary autoregressive process. A Bayesian method using Markov chain Monte Carlo for the analysis of this model is developed. Using the RVs of S&P500 and Nikkei 225, it is examined whether this extension improves the predictive ability of volatility.

Wehrhahn, Claudia

A dynamic point process model for human movement

Human mobility has been a widely studied topic in the social sciences (particularly in sociology and anthropology), and more recently, in computer science and engineering. Because of the paucity of individual level data, traditional models of human mobility have focused on coarse-scale movement patterns. However, as high resolution data from Global Positioning Systems (GPS) has become more widely available over the last 10 years, interest in individual-level models of human mobility has grown. In this talk we present a nonparametric Bayesian model that can be used to study human mobility. The positions of individuals at each point in time are treated as a realization of a (non-homogeneous) Matern type III repulsive point process. The use of a non-homogenous Matern process allows us to mimic two key features of real human mobility traces: on one hand, individuals tend to congregate in specific places and move preferentially along certain routes, on the other, individuals tend to maintain a certain distance between them (the so-called personal space in the Proxemics literature). Mobility is captured in our model by linking the parameters of the point processes (the intensity function of the underlying Poisson process, as well as the repulsive ratio). For this purpose, we introduce a novel prior on collections of non-negative functions that uses mixtures of Bernstein polynomials. The model is illustrated using both simulated and real datasets.

Wehrhahn Cortes, Claudia

Restricted Nonparametric Mixtures models for Disease Clustering

Identifying disease clusters (areas with an unusually high incidence of a particular disease) is a common problem in epidemiology and public health. We describe a Bayesian nonparametric mixture model for disease clustering that constrains clusters to be made of contiguous areal units. This is achieved by modifying the exchangeable partition probability function associated with the Ewen's sampling distribution. The model is illustrated using data on cancer rates.

Weiss, Robert

Prior specification for logistic regression coefficients using meta-analytic and compatible priors

Bayesian methods can help in the analysis of clinical trial data, by adding proper prior information into the analysis thereby potentially decreasing uncertainty in conclusions. In regression analysis of a binary outcome, we discuss using a meta-analysis to set priors for the intercept and for those covariates for which a meta-analysis is possible. For coefficients of categorical covariates without meta-analytic information, we discuss setting priors that are self-consistent and compatible with the meta-analytic priors.

West, Mike

Multivariate Bayesian Predictive Synthesis for Macroeconomic Forecasting

I discuss Bayesian predictive synthesis (BPS) for multivariate time series with examples in multi-step ahead macroeconomic forecasting. BPS theory defines an over-arching framework for forecast synthesis whether density forecasts represent predictions from sets of models, individual forecasters, agencies, or other sources. The general result is a class of multivariate, dynamic latent factor models in which latent factor processes represent individual models or forecasters. The framework naturally allows modelling and estimation --- sequentially and adaptively over time --- of varying forecast biases and facets of miscalibration of individual forecast densities, and critically of time-varying inter-dependencies among models or forecasters over multiple series. The talk explores this in a specific subclass of BPS model, using a dynamic, seemingly-unrelated, dynamic factor and regression (DFSURE) approach in which each forecast model is linked to one set of multivariate dynamic latent factor processes. Bayesian simulation-based computation enables implementation. Applied studies with multiple macro-economic time series highlight questions of forecast synthesis methodology with respect to forecasting goals: interest in 12 or 24 month-ahead forecasting demands --- from a formal Bayesian perspective --- analysis customized to the horizon. The analyses bear out the potential to define fully Bayesian, interpretable models that can (i) adapt to time-varying biases and miscalibration of multiple models or forecasters, (ii) adaptively and practically account for while generating useful insights into ---patterns of time-varying relationships and dependencies among sets of models or forecasters, while (iii) improving forecast accuracy --- in some cases, most substantially for each of several multiple macroeconomic series together, at multiple horizons.

Whiteley, Nick

Sampling normalizing constants in high dimensions using inhomogeneous diffusions

Motivated by the task of computing normalizing constants and importance sampling in high dimensions, we study the dimension dependence of fluctuations for additive functionals of time-inhomogeneous Langevin-type diffusions on Rd. The main results are nonasymptotic variance and bias bounds, and a central limit theorem in the regime where d tends to infinity. We demonstrate that a temporal discretization inherits the fluctuation properties of the underlying diffusion, which are controlled at a computational cost growing at most polynomially with d. The key technical steps include establishing Poincare inequalities for time-marginal distributions of the diffusion and nonasymptotic bounds on deviation from Gaussianity in a martingale central limit theorem.Joint work with Christophe Andrieu (Bristol) and James Ridgway (AgroParisTech)

Williamson, Daniel

Climate model discrepancy: thinking outside of the UQ toolbox

The Uncertainty Quantification community has widely accepted the need to quantify and include model discrepancy when estimating the parameters of complex physical models. Due to well-known identifiability problems with traditional Bayesian approaches, structured prior knowledge is required for model discrepancy in order to reach appropriate parameter estimates. Climate models represent a special challenge for existing UQ methods. Climate model calibration (what the community calls “tuning”, if done the traditional Kennedy-O’Hagan way, would require structured prior specification for discrepancy over many of the massive spatio-temporal fields that make up the terabytes of output produced by climate models and satellite/station observations.We will argue that the way these models are constructed makes the specification of well informed structured discrepancy judgements over these fields infeasible. We present a “delayed-Bayesian”approach to parameter inference that involves first verifying the model’s capability to simulate emergent climate processes, then using the results of this analysis for structured prior discrepancy modeling and a formal calibration to complete tuning. These ideas will be presented in the form of a case study working with the Canadian climate model CanAGCM4.

Wilson, Amy

Varying coefficient models for Bayes linear emulation

Computer simulators of complex systems are widely used to help make policy decisions in industry and government. These computer simulators are often computationally intensive and have high-dimensional input spaces, making it difficult to use traditional Monte Carlo simulation to assess the impact of input uncertainty. Instead, a statistical model of the computer simulator, known as an emulator, can be used in combination with Monte Carlo simulation to model uncertainty in simulator output. This presentation describes methodology for emulation of computer simulators, motivated by a real-world example in energy policy. The computer simulator studied was an economic model of investment in electricity generation, used to study a government policy designed to incentivise investment in renewable technologies. Limited computing time meant that few runs of the simulator were available to fit an emulator. The statistical methodology developed was therefore focussed on accurately capturing the uncertainty in computer simulator output arising from the small number of available runs. A varying coefficient model was used to emulate the computer simulator to allow for increased uncertainty when extrapolating away from simulator runs. To maximise use of the small number of runs available, this varying coefficient emulator was paired with a criterion-based procedure for design selection.

Woody, Spencer

Bayesian post-selection inference in spatial modeling with application to hotspot detection

In problems of spatial modeling, it is often of interest to find spatially contiguous regions of anomalous signal, or hotspots. When estimating the magnitude of these signals after discovery, it is necessary to account for selection bias. That is, unadjusted estimates of hotspot signal intensity will be biased owing to the fact that the target of inference is chosen adaptively. While there is a wealth of literature on Bayesian approaches to spatial modeling, there is currently little work done in this realm addressing this "look-elsewhere" problem. Existing Bayesian approaches to post-selection inference utilize a likelihood truncated by the selection event for performing selection-adjusted posterior inference. However, computing the exact truncated likelihood is usually only possible in simple settings; in the case of spatial models, doing so requires calculating an intractable integral over a very high-dimensional joint distribution. Therefore we propose a thresholding rule for identification of hotspots, and present a method for valid post-selection Bayesian inference of the magnitude of those hotspots, based on a tractable approximation to the truncated likelihood function. We then apply this method to a problem in evolutionary biology, with the goal to jointly identify alleles under natural selection and show geographically where the selection occurs.

Xu, Xinyi

A Bayesian nonparametric approach for estimating Heterogeneous Causal Effects

Inferring a causal relationship is an important task in both social science and health research.In a large population, different subgroups of individuals might respond differently to certain treatments.Identifying and estimating the heterogeneous effects can help researchers improve treatments or better allocate resources to meet the needs.In an observational study, propensity score is often used as a dimension reduction tool to aid the inference under the ignorable treatment assignment assumption.We propose a nonparametric Bayesian approach that utilizes propensity scores and observable factors to capture heterogeneous treatment effects.We show that our model produces estimators that take the same form as traditional matching estimators under certain prior specifications, and outperform the matching estimators with improved efficiency and better identification of heterogeneous effects.A sensitivity analysis strategy is proposed to assess the impact due to potential unmeasured confounders.Furthermore, we apply our method to investigate the impact of college attendance on women fertility, which is known to suffer from the potential heterogeneous effects.

Xu, Yanxun

Bayesian Repulsive Gaussian Mixture Model

We develop a general class of Bayesian repulsive Gaussian mixture models that encourage well-separated clusters, aiming at reducing potentially redundant components produced by independent priors for locations (such as the Dirichlet process). The asymptotic results for theposterior distribution of the proposed models are derived, including posterior consistency and posterior contraction rate in the context of nonparametric density estimation. More importantly, we show that, as a measurement of the model complexity, the posterior number of necessary components to fit the data well grows sub-linearly with respect to the sample size asymptotically. In addition, an efficient and easy-to-implement blocked-collapsed Gibbs sampler is developed based on the exchangeable partition distribution and the corresponding urn model. We evaluate the performance and demonstrate the advantages of the proposed model through extensive simulationstudies and real data analysis. The R code is freely available online.

Yang, Chiao-Yu

Dirichlet Process Mixture Model: Limiting Distribution for the Number of Components

Dirichlet process mixture (DPMM) is a widely used model for density estimation as well as statistical inference on the number of components. Despite its success in density estimation, DPMM was shown to exhibit inconsistency with respect to estimation of the number of true components, especially in the regime of finite components. In this work, we show the limiting distribution for DPMM’s posterior number of components in the regime of finite components, and then give several extensions and simulations.

Yang, Ziheng

Statistical and computational challenges for Bayesian inference in comparative genomics

The coalescent is a stochastic process model in population genetics that describes the genealogical history of a sample taken from a population, and forms the basis for inference in modern population genetics. The multispecies coalescent (MSC) is its simple extension to multiple species, and provides a probabilistic description of the genealogical relationships of DNA sequences sampled from multiple closely-related species. In the past 5-10 years, the MSC has emerged as a powerful framework for statistical inference using genomic sequence data to address a number of problemsin evolutionary biology and comparative genomics, such as species delimitation, species tree estimation, and inference of population demographic changes or migration histories. Full-likelihood implementations of the MSC have relied on Bayesian Markov chain Monte Carlo (MCMC) algorithms, which integrate over the latent variables (the unknown genealogical histories or gene trees at different genomic regions). While those methods have better statistical properties than heuristic methods or approximate Bayesian computation, they involve heavy computation. In particular, the cross-model MCMC algorithms often suffer from poor mixing in large datasets. In this talk, I will review the MSC model and discuss our efforts to implement it through smart transmodel MCMC proposals.

Yeager, David

Nonparametric regression models of multilevel treatment effect moderation: The National Study of Learning Mindsets

Bayesian nonparametric approaches to causal inference have recently become popular. However, current approaches fail to address three important features in applications: Accounting for multilevel structure, allowing for targeted regularization, and providing interpretable summaries of scientifically meaningful quantities. We extend recently proposed BART-based methods to include all of these features. A key point of this model is a parameterization that allows treatment heterogeneity to be regularized separately from prognostic effects, and also parsimoniously incorporates multilevel structure. In an application to the National Study of Learning Mindsets (Yeager et. al., 2017) we use these new tools to provide meaningful insights about effect modification at the school and individual level. Our novel posterior summarization strategy avoids pitfalls common to existing approaches relying on post-hoc data snooping or large collections of hypothesis tests.

Zanella, Giacomo

Bayesian complementary clustering, MCMC and Anglo-Saxon place-names

This work is motivated by the study of Anglo-Saxon settlements, where administrative clusters involving complementary names tend to appear. We investigate such phenomena by developing a Bayesian Random Partition Model for multi-type point processes where clusters are formed by points of different types (complementary clustering). The resulting posterior distribution, defined over the space of matchings contained in a k-partite hypergraph, presents computational challenges even for moderate data sizes. Although this is a common situation in the context of discrete-valued high-dimensional parameter spaces, the computational methodology applicable in these contexts is still limited. We propose novel, generic samplers for posterior distributions with discrete state spaces, based on the idea of informed locally-balanced proposal distributions. We provide theoretical analysis of the proposed schemes, including a characterization of the class of asymptotically-optimal proposal distributions, an explicit convergence analysis in some simplified scenarios, and an elegant connection with classical gradient-based schemes for continuous spaces. The resulting algorithms are straightforward to implement and provide significantly improved scalability for Bayesian models with discrete parameters, which we illustrate on record linkage and variable selection problems. Finally, the methodology is used to study a dataset including locations and place-names of 1316 Anglo-Saxon settlements dated around 750-850 AD.

Zhang, Junni

Bayesian Estimation of Demographic Systems

Many problems in applied demography consist of estimating demographic systems. We present a general approach to estimating demographic systems based on the idea of a demographic account. A demographic account is a set of linked tabulations of demographic series, such as fertility, mortality, migration, and population. The estimation methods are fully Bayesian and allow for multiple noisy data sources. We simultaneously estimate (i) the true finite-population counts of events and populations, (ii) the demographic rates underlying these counts, and (iii) indicators of data quality. The methods also generate uncertainty measures for all estimated quantities. We illustrate the methods using examples from New Zealand and China.