June 22–26 2015
Raleigh NC

bnp 10

The 10th Conference on Bayesian Nonparametrics

Program & Abstracts

NC STATE UNIVERSITY

Duke UNIVERSITY

THE UNIVERSITY *of* NORTH CAROLINA *at* CHAPEL HILL

Time	Monday, Jun 22	Tuesday, Jun 23	Wednesday, Jun 24	Thursday, Jun 25	Friday, Jun 26
	8:45 - 9:00 AM: Opening				
9 AM	9:00 - 10:00 AM: Keynote I Harry van Zanten	9:00 - 10:00 AM: Contributed 2 A: F Gao, Miller, Ngyuen B: Roy, Griffin, Broderick	9:00 - 10:00 AM: Keynote II Emily Fox	9:00 - 10:00 AM: Contributed 5 A: Cai, Pati, Heller B: M Li, Seiler, Zhang	9:00 - 10:00 AM: Keynote III Igor Prünster
10 AM	Coffee Break	Coffee Break	Coffee Break	Coffee Break	Coffee Break
11 AM	10:30 AM - Noon: Invited 1 Scalable BNP I Blei Taddy Sudderth	10:30 AM - Noon: Invited 3 Multiscale BNP Ma Ray Quintana	10:30 AM - Noon: Invited 5 Sparsity Scott Carvalho Bhattacharya	10:30 AM - Noon: Invited 7 Asymptotics Szabo Rousseau Schmidt-Hieber	10:30 - 11:50 AM: Contributed 7 A: De Blasi, Tansey, Banerjee, Johndrow B: Favaro, Forbes, Martinez
Noon	Lunch Break	Lunch Break	Lunch Break	Lunch Break	Closing
1 PM					
2 PM	2:00 - 3:30 PM: Invited 2 Graphs Caron Orbanz Petrone	2:00 - 3:30 PM: Invited 4 Random Measures Zhou Nipoti Barrientos	2:00 - 3:30 PM: Invited 6 Scalable BNP II Duvenaud Dunson Teh	2:00 - 3:30 PM: Invited 8 Biostatistics Müller Holmes Nieto-Barajas	
3 PM	Coffee Break	Coffee Break	Coffee Break	Coffee Break	
4 PM	4:00 - 5:20 PM: Contributed 1 A: Arbel, Raykov, Lomeli, Srivastava B: Gal, Lee, Heaukulani, Ranganath	4:00 - 5:20 PM: Contributed 3 A: Bose, Kottas, Ferreira, Richardson B: C Gao, Norets, Salomond, van Waaij	4:00 - 5:20 PM: Contributed 4 A: Ge, Calderon, Boyles, Ruggiero B: H Li, Shen, Guha, Bhaumik	4:00 - 5:20 PM: Contributed 6 A: Serra, Naulet, Murray, McCulloch B: Liu, Bean, Canale, Soriano	
5 PM					
Evening	8:00 - 10:00 PM: Poster I	8:00 - 10:00 PM: Poster II		7:00 - 10:00 PM: Banquet Dinner	

Notes: 1) Keynote / Invited / Contributed A @ SAS Hall 2203; 2) Contributed B @ SAS Hall 1102; 3) Posters @ Embassy Suites Ballroom; 4) Dinner @ NC Museum of Art

Program

Monday, June 22, 2015

8:45 AM – 9:00 AM: Opening, SAS Hall

9:00 AM –10:00 AM: Keynote I, SAS Hall 2203 Chair: Subhashis Ghosal
Harry van Zanten. *Bayesian estimation of a smooth function on a large graph*

10:00 AM – 10:30 AM: Coffee Break

10:30 AM – 12:00 PM: Invited 1, SAS Hall 2203 Chair: Yize Zhao
David Blei. *Stochastic Variational Inference for Bayesian Nonparametrics*
Matt Taddy. *Bayesian and Empirical Bayesian Forests*
Erik Sudderth. *Scalable and Flexible Nonparametric Models of Sequential Data*

12:00 PM – 2:00 PM: Lunch Break

2:00 PM – 3:30 PM: Invited 2, SAS Hall 2203 Chair: Ye Liu
Francois Caron. *Sparse random graphs with exchangeable point processes*
Peter Orbanz. *Subsampling large graphs and invariance in networks*
Sonia Petrone. *Bayesian nonparametrics in bivariate evolutionary phenomena*

3:30 PM – 4:00 PM: Coffee Break

4:00 PM – 5:20 PM: Contributed 1A, SAS Hall 2203 Chair: Long Nguyen
Julyan Arbel. *Moment-based strategies for Bayesian nonparametrics*
Yordan Raykov. *Fast Approximate MAP Inference for Bayesian Non-Parametrics*
Maria Lomeli-Garcia. *A marginal sampler for σ-Stable Poisson-Kingman mixture models*
Sanvesh Srivastava. *WASP: Scalable Bayes via barycenters of subset posteriors*

4:00 PM – 5:20 PM: Contributed 1B, SAS Hall 1102 Chair: Haiming Zhou
Yarin Gal. *An Infinite Product of Sparse Chinese Restaurant Processes*
Ju Hee Lee. *Bayesian Inference for Tumor Subclones Accounting for Sequencing and Structural Variants*
Creighton Heaukulani. *Random partition-based black-box inference for feature allocations and their generalizations*
Rajesh Ranganath. *Correlated Random Measures*

8:00 PM – 10:00 PM: Poster I, Embassy Suites Ballroom

Tuesday, June 23, 2015

9:00 AM – 10:00 AM: CONTRIBUTED 2A, SAS HALL 2203 Chair: Mingyuan Zhou
Fengnan Gao. *Posterior contraction rates for deconvolution of Dirichlet-Laplace mixtures*
Jeff Miller. *An approach to inference under misspecification*
Long Nguyen. *Identifiability and posterior concentration of matrix-variate parameters in mixture models*

9:00 AM – 10:00 AM: CONTRIBUTED 2B, SAS HALL 1102 Chair: Emily Fox
Daniel Roy. *The continuum-of-urns scheme, generalized beta and Indian buffet processes, and hierarchies thereof*
Jim Griffin. *Compound random measures and their use in Bayesian nonparametrics*
Tamara Broderick. *Posteriors, conjugacy, and exponential families for completely random measures*

10:00 AM – 10:30 AM: COFFEE BREAK

10:30 AM – 12:00 PM: INVITED 3, SAS HALL 2203 Chair: Michele Guindani
Li Ma. *Probabilistic multi-resolution scanning for cross-sample differences*
Kolyan Ray. *Bernstein-von Mises theorems for adaptive Bayesian nonparametric procedures*
Fernando Quintana. *Predictions Based on the Clustering of Heterogeneous Functions via Shape and Subject-Specific Covariates*

12:00 PM – 2:00 PM: LUNCH BREAK

2:00 PM – 3:30 PM: INVITED 4, SAS HALL 2203 Chair: Matt Taddy
Mingyuan Zhou. *Priors for random count matrices with random or fixed row sums*
Bernardo Nipoti. *Modeling the association structure of clustered time-to-event data*
Andrés Felipe Barrientos. *Bayesian goodness-of-fit testing for density regression*

3:30 PM – 4:00 PM: COFFEE BREAK

4:00 PM – 5:20 PM: CONTRIBUTED 3A, SAS HALL 2203 Chair: Jarno Hartog
Sudip Bose. *Generating priors and Polya posteriors*
Athanasios Kottas. *Nonparametric Bayesian modeling for dynamic ordinal regression relationships*
Marco Ferreira. *Dynamic Multiscale Spatiotemporal Models for Poisson Data*
Robert Richardson. *Bayesian Non-parametric Modeling for Integro-Differential Equations*

4:00 PM – 5:20 PM: CONTRIBUTED 3B, SAS HALL 1102 Chair: Subharup Guha
Chao Gao. *Some Results on Posterior Contraction of Matrix Estimation*
Andriy Norets. *Adaptive Bayesian Estimation of Conditional Densities*
Jean-Bernard Salomond. *Gereral posterior contraction results for nonparametric ill posed incerse problems*
Jan van Waaij. *An adaptive prior for a one-dimensional diffusion model*

8:00 PM – 10:00 PM: POSTER II, EMBASSY SUITES BALLROOM

Wednesday, June 24, 2015

9:00 AM – 10:00 AM: KEYNOTE II, SAS HALL 2203 Chair: Peter Müller
Emily Fox. *Scalable Modeling and Inference for Complex Data Streams*

10:00 AM – 10:30 AM: COFFEE BREAK

10:30 AM – 12:00 PM: INVITED 5, SAS HALL 2203 Chair: Eunjee Lee
James Scott. *Multiscale spatial density smoothing*
Carlos Carvalho. *Utility Based Model Section for Bayesian Nonparametric Modeling using Trees*
Anirban Bhattacharya. *Bernstein von Mises theorems in Wasserstein distance*

12:00 PM – 2:00 PM: LUNCH BREAK

2:00 PM – 3:30 PM: INVITED 6, SAS HALL 2203 Chair: Surya Tokdar
David Duvenaud. *Scaling Up Bayesian Optimization*
David Dunson. *Scaling up Bayes computation to massive data*
Yee Whye Teh. *Mondrian Forests: Efficient Random Forests for Streaming Data via Bayesian Nonparametrics*

3:30 PM – 4:00 PM: COFFEE BREAK

4:00 PM – 5:20 PM: CONTRIBUTED 4A, SAS HALL 2203 Chair: David Mensah
Hong Ge. *Hierarchical infinite hidden Markov models*
Christopher Calderon. *Inferring Latent States and Refining Force Estimates via Hierarchical Dirichlet Process Modeling in Single Particle Tracking Experiments*
Levi Boyles. *Hierarchical Clustering with Aldous' Beta-Splitting Trees*
Matteo Ruggiero. *Filtering hidden Markov measures*

4:00 PM – 5:20 PM: CONTRIBUTED 4B, SAS HALL 1102 Chair: Pierpaolo de Blasi
Hanning Li. *Bayesian Variable Selection Using Continuous Shrinkage Priors*
Weining Shen. *Bayesian model selection for semi-nonparametric models*
Subha Guha. *Nonparametric Variable Selection, Clustering and Prediction for High-Dimensional Regression*
Prithwish Bhaumik. *Efficient Bayesian estimation and uncertainty quantification in differential equation models*

Thursday, June 25, 2015

9:00 AM – 10:00 AM: CONTRIBUTED 5A, SAS HALL 2203 Chair: Weining Shen
Diana Cai. *Priors on Exchangeable Directed Graphs*
Debdeep Pati. *Bayesian nonparametric graphon estimation*
Katherine Heller. *The Bayesian Echo Chamber: Modeling Influence in Conversations*

9:00 AM – 10:00 AM: CONTRIBUTED 5B, SAS HALL 1102 Chair: Maria Lomeli
Meng Li. *Bayesian Methods for Detecting Boundaries of Images*
Christof Seiler. *Parcellations of Vector Fields in Computational Anatomy*
Zhengwu Zhang. *Bayesian Clustering of Shapes of Curves*

10:00 AM – 10:30 AM: COFFEE BREAK

10:30 AM – 12:00 PM: INVITED 7, SAS HALL 2203 Chair: Debdeep Pati
Botond Szabo. *Asymptotic behaviour of the empirical Bayes posteriors associated to maximum marginal likelihood estimator*
Judith Rousseau. *Empirical Bayes procedures - understanding the behaviour of the MMLE in nonparametric models*
Johannes Schmidt-Hieber. *Adaptive posterior contraction*

12:00 PM – 2:00 PM: LUNCH BREAK

2:00 PM – 3:30 PM: INVITED 8, SAS HALL 2203 Chair: Antonio Canale
Peter Müller. *BNP Inference for Dynamic Treatment Regimes*
Chris Holmes. *Scalable Bayesian nonparametric regression for large data applications*
Luis Nieto-Barajas. *Markov constructions in Biostatistics applications*

3:30 PM – 4:00 PM: COFFEE BREAK

4:00 PM – 5:20 PM: CONTRIBUTED 6A, SAS HALL 2203 Chair: James Jonhdrow
Paulo Serra. *Adaptive empirical Bayesian smoothing splines*
Zacharie Naulet. *Adaptive Bayesian nonparametric regression using mixtures of kernels*
Jared Murray. *Loglinear Bayesian Additive Regression Trees for Classification, Counts, and Heteroscedastic Regression*
Robert McCulloch. *Nonparametric Heteroscedastic Regression Modeling, Bayesian Regression Trees and MCMC Sampling*

4:00 PM – 5:20 PM: CONTRIBUTED 6B, SAS HALL 1102 Chair: Juhee Lee
Linxi Liu. *Posterior concentration rate of a class of multivariate density estimators based on adaptive partitioning*
Andrew Bean. *Transformation and Bayesian Density Estimation*
Antonio Canale. *Multiscale Bernstein Polynomials for Densities*
Jacopo Soriano. *Variability decomposition across related mixture distributions with weight coupling and local perturbation kernels*

7:00 PM – 10:00 PM: BANQUET DINNER, NC MUSEUM OF ART

Friday, June 26, 2015

9:00 AM – 10:00 AM: KEYNOTE III, SAS HALL 2203 Chair: David Dunson
Igor Prünster. *Dependent Random Measures and Prediction*

10:00 AM – 10:30 AM: COFFEE BREAK

10:30 AM – 11:50 AM: CONTRIBUTED 7A, SAS HALL 2203 Chair: Anirban Bhattacharya
Pierpaolo De Blasi. *Posterior asymptotics under the sup-L_1 distance in nonparametric conditional estimation*
Wesley Tansey. *False Discovery Rate Smoothing*
Sayantan Banerjee. *Bayesian methods for learning relations in high dimensional graphical models*
James Johndrow. *Bayesian nonparametric methods for contingency tables*

10:30 AM – 11:50 AM: CONTRIBUTED 7B, SAS HALL 1102 Chair: Jeff Miller
Stefano Favaro. *Rediscovery of Good-Turing estimators via Bayesian nonparametrics*
Catherine Forbes. *Bayesian restricted likelihood-based instrumental variables regression*
Fabian Martinez. *Change-Point Detection on Dependent Data, a Random-Partition Approach*

11:50 AM – 12:30 PM: CLOSING, SAS HALL

Poster Session I

Monday, June 22, 2015, 8:00 PM - 10:00 PM, Embassy Suites Ballroom

Isadora Antoniano Villalobos. *Bayesian Inference for Multivariate Dependence Structures in Extreme Value Theory*
William Barcella. *Variable selection in covariate dependent random partition models: an application to urinary tract infection*
Marco Battiston. *Multi-armed bandit for species discovery: a Bayesian nonparametric approach.*
Annalisa Cadonna. *Bayesian nonparametric modeling for multiple spectral densities*
Federico Camerlenghi. *Nonparametric hierarchical models and their predictive structure*
Lu Cheng. *Integrating prior biological knowledge into machine learning models for predicting drug responses*
William Cipolli. *Multiple Testing via an Approximate Finite Polya Tree*
Glen Colopy. *Prior Regularization and Likelihood-Based Deterioration Detection Using Gaussian Processes in the Stepdown-ward*
David Dahl. *Attraction Indian Buffet Process*
Patrick Dallaire. *A Bayesian nonparametric prior on infinite directed acylic graphs*
Priyam Das. *Bayesian Quantile Regression Using Random Series*
Nicola Donelli. *Semiparametric Bayesian vector multiplicative error model*
Matt Edwards. *Bayesian semiparametric spectral density estimation in gravitational wave data analysis*
Gilbert Fellingham. *Predicting Home Run Production in Major League Baseball Using a Bayesian Semiparametric Model*
Yessica Fermin. *Inferring the Structure of Protein-Networks by Nonparametric Dynamic Bayesian Networks*
Tamara Fernandez. *Gaussian processes for survival analysis*
Nicholas Foti. *Streaming Variational Inference for Bayesian Nonparametric Mixture Models*
Chris Glynn. *Fully Bayesian Inference for a Dynamic Linear Topic Model*
Alessandra Guglielmi. *ϵ-NRMI mixtures for density and clustering estimation*
Rajarshi Guhaniyogi. *Bayesian Conditional Density Filtering for Streaming Data*
Luis Gutiérrez. *Bayesian nonparametric hypothesis testing for the two-sample problem*
Christoph Hellmayr. *Flexible Functional Clustering Using Dirichlet Processes*
Nhat Ho. *Posterior concentration of mixing parameters in some weakly identifiable finite mixture models*
Masaaki Imaizumi. *Bayesian estimation for nonparametric regression with low-rank tensor data*
Gwangsu Kim. *Bayesian analysis of Cox regression with a time-varying coefficient and its application*
Alisa Kirichenko. *Estimating a smoothly varying function on a large graph*
Christopher Krut. *A multivariate functional linear model with spatially varying coefficients.*

Poster Session II

Tuesday, June 23, 2015, 8:00 PM - 10:00 PM, Embassy Suites Ballroom

Dipankar Bandyopadhyay. *Nonparametric spatial models for clustered ordered periodontal data*
Eunjee Lee. *A Bayesian Functional Linear Cox Regression Model (BFLCRM) for Predicting Time to Conversion to Alzheimer's Disease*
Fabrizio Leisen. *A pseudo-marginal algorithm for MGF-like distributions*
Ramses Mena. *Ruin Probabilities for Bayesian Exchangeable Claims Processes*
David Kwamena Mensah. *Functional Longitudinal modelling with covariate dependent smoothness*
Thomas Murray. *Combining Functional or Survival Data Sources*
Ana Maria Ortega Villa. *Semiparametric Spatio-Temporal Varying Coefficient Model for Matched Case-Crossover Studies*
Garritt Page. *Spatial Product Partition Models*
Gavino Puggioni. *A Bayesian Nonparametric Approach for Spatiotemporal Point Processes*
Pedro Regueiro. *Stochastic variational Bayes in the stochastic blockmodel*
Boyu Ren. *Factor analysis of multi-population species sampling sequences*
Nilotpal Sanyal. *Bayesian wavelet analysis with nonlocal priors*
Terrance Savitsky. *Convergence Rates of Posterior Distributions under Informative Sampling*
Moritz Schauer. *Poster: Nonparametric Bayesian inference for diffusion processes - an integrated approach*
Aaron Schein. *Dynamic Bayesian Poisson Tensor Factorization*
Briana Joy Stephenson. *Bayesian nonparametric methods to describe high-dimensional exposures*
Adam Suarez. *Bayesian Estimation of Principal Components for Functional Data*
Qianwen Tan. *Bayesian inference in mixed models involving ordinary differential equations*
Victor Veitch. *On the general class of sparse exchangeable graphs on R*
Sara Wade. *A Bayesian nonparametric longitudinal model for predicting the temporal progression of Alzheimer's disease*
Claudia Wehrhahn. *Small sample properties of Dirichlet process mixture models for data supported on compact intervals*
Ran Wei. *The nonparametric Bayesian model with continuous shrinkage priors and its application in multiple pesticide exposures data*
Sinead Williamson. *Bayesian nonparametric models for prediction in networks*
Keisuke Yano. *Minimax Predictive Distributions in* l_2
Yize Zhao. *Bayesian Hierarchical Variable Selection for Genome-wide Association Studies*
Haiming Zhou. *Bayesian Causal Mediation Analysis for Survival Data*
Weixuan Zhu. *A multivariate extension of a vector of two-parameter Poisson-Dirichlet processes*

Abstracts

Isadora Antoniano Villalobos Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Bocconi University, Italy
▶ *Bayesian Inference for Multivariate Dependence Structures in Extreme Value Theory*

▷ In recent years, interest in high-dimensional and multivariate problems concerning extreme events has increased. In fields such as environmental and economical sciences, analysis of dependence structures is often required. Current dependence models for multivariate maxima are based upon max-stable distributions, characterized by the exponent measure function governing the dependence structure among the data. A change of variable allows an alternative characterization in terms of the Pickands dependence function, defined on a unit simplex of adequate dimension. A Pickands dependence function must satisfy certain conditions in order to properly define a max-stable distribution. In particular, it must be convex over its domain. A recent proposal exploits the shape-preserving properties of multivariate Bernstein polynomials in order to represent the projection of an initial, possibly non-convex estimate, onto the space of convex functions, thus constructing an estimator with improved theoretical properties. Two potential limitations of this method regard the choices of the initial estimation and the order of the Bernstein polynomials involved in the representation. In the present work, we propose a Bayesian approach in order to overcome the first issue, and briefly discuss a non-parametric extension for dealing with the second.

⊕ Joint work with *Giulia Marcon, Simone Padoan*.

Julyan Arbel Contributed 1A, Monday, 6/22/15, 4:00 pm - 4:20 pm
Collegio Carlo Alberto, Moncalieri, Italy
▶ *Moment-based strategies for Bayesian nonparametrics*

▷ Marginal methods and the Ferguson and Klass algorithm are two well-established posterior sampling techniques for Bayesian nonparametric models. Though they proved to be successful techniques both in theory and in practice, it is also worth pointing out that they suffer from some drawbacks which we wish to address here. Under marginal method, the infinite dimensional parameter G is marginalized out, based on Blackwell–MacQueen Pólya urn schemes. This approach typically yields point estimates in the form of posterior expectations but prevents estimation of other non-linear functionals of posterior distributions such as credible intervals. In the Ferguson and Klass algorithm, one samples trajectories of jumps from G, leading to a truncation error which needs be controlled. In this talk, we show that exploiting readily available posterior moments of transforms of G within these techniques can lead to substantial estimation improvements, although at limited extra computational cost. This is first illustrated for marginal methods on hazard mixture models where we can draw approximate inference on survival functions that is not limited to the posterior mean but includes for instance credible bands and the median survival time. Second, in normalized random measures with independent increments models and the Indian buffet process model, we show that the truncation error of the Ferguson and Klass algorithm can be efficiently controlled by moment conditions.

⊕ Joint work with *Antonio Lijoi, Bernardo Nipoti, and Igor Prunster*.

Dipankar Bandyopadhyay Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Division of Biostatistics, School of Public Health, University of Minnesota, USA
▶ *Nonparametric spatial models for clustered ordered periodontal data*

▷ Clinical attachment level (CAL) is regarded as the most popular measure to assess periodontal disease (PD) status. Although CAL is continuous (in mm), these probed tooth-site level measures are usually rounded and recorded as whole numbers producing clustered (site measures within a mouth) error-prone responses. In addition, it is hypothesized that PD progression can be spatially-referenced. Although these rounded CAL values represent some ordering of PD progression, traditional analysis mostly consider linear mixed models with appropriate spatial random effect terms. In this talk, we consider a different route to model the ordinal CAL categories assuming a Bayesian multivariate probit framework where the cut-point parameters linking the observed ordinal CAL levels to the latent underlying disease process can be fixed in advance. The latent spatial association characterizing conditional independence under Gaussian graphs is introduced via a nonparametric Bayesian approach motivated by the probit stick-breaking process, where the components of the stick-breaking weights follows a multivariate Gaussian density with the precision matrix distributed as G-Wishart. This yields a computationally simple, yet robust and flexible framework to capture the latent disease progression leading to a natural clustering of tooth-sites and subjects with similar disease status (beyond spatial clustering) and improved parameter estimation through sharing of information. Both simulation studies and application to a motivating PD dataset reveal the advantages of considering this flexible nonparametric ordinal framework over other alternatives.
⊕ Joint work with *Antonio Canale*.

Sayantan Banerjee Contributed 7A, Friday, 6/26/15, 11:10 am - 11:30 am
The University of Texas MD Anderson Cancer Center, USA
► *Bayesian methods for learning relations in high dimensional graphical models*
▷ Objects of very high dimension are frequently encountered in many modern statistical applications, sometimes exceeding the available sample size. In such instances, it is extremely important to learn conditional independence relationship between different components of the observations. In spite of the very high complexity of the data, a key feature that allows valid statistical analysis is sparsity. In a Gaussian model, intrinsic relations between variables are neatly summarized by the precision matrix, given by the inverse of the covariance matrix. We discuss two Bayesian approaches for estimating the precision matrix in a high dimensional Gaussian model, one based on the so called graphical Wishart prior, and the other based on a Bayesian analog of the graphical lasso. In the former case, exploiting the conjugacy of the graphical Wishart prior, we can compute Bayesian estimates based on suitable loss functions and study their performance by simulation. For nearly banded true precision matrices, we obtain the convergence rate of the Bayesian procedure under an operator norm. In the latter case, lack of a conjugacy structure poses high challenge for Bayesian computation, since the standard approach using a (reversible jump) Markov chain Monte-Carlo sampling plan is not practically feasible in high dimensional situations. We derive the posterior convergence rate at a sparse true precision matrix under the Frobenius norm, and show that it agrees with the oracle convergence rate. We also device an approximate computing technique based on Laplace approximation avoiding Markov chain Monte-Carlo methods completely.
⊕ Joint work with *Subhashis Ghosal*.

William Barcella Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University College London, United Kingdom
► *Variable selection in covariate dependent random partition models: an application to urinary tract infection*

▷ Lower urinary tract symptoms (LUTS) can indicate the presence of urinary tract infection (UTI), a chronic condition that requires expensive and time consuming care as well as leading to reduced quality of life. Detecting the presence and gravity of an infection from the earliest symptoms is then highly valuable. Typically, white blood cell count (WBC) measured in a sample of urine is used to assess UTI. We consider clinical data from 1341 patients at their first visit in which UTI (i.e. WBC$\geq$ 1) is diagnosed. In addition, for each patient, a clinical profile of 34 symptoms was recorded. In this paper we propose a Bayesian nonparametric regression model based on the Dirichlet Process (DP) prior aimed at providing the clinicians with a meaningful clustering of the patients based on both the WBC (response variable) and possible patterns within the symptoms profiles (covariates). This is achieved by assuming a probability model for the symptoms as well as for the response variable. To identify the symptoms most associated to UTI, we specify a spike and slab base measure for the regression coefficients: this induces dependence of symptoms selection on cluster assignment. Posterior inference is performed through Markov Chain Monte Carlo methods.
⊕ Joint work with *Maria De Iorio, Gianluca Baio*.

Andrés Felipe Barrientos Invited 4, Tuesday, 6/23/15, 3:00 pm - 3:30 pm
Pontificia Universidad Católica de Chile, Chile
► *Bayesian goodness-of-fit testing for density regression*
▷ We present a novel Bayesian approach to perform hypotheses testing of goodness-of-fit for sets of predictor-dependent densities. We consider the problem of testing point null hypotheses versus non-parametric alternatives. The proposed method is based on an appropriated transformation of the response variable and a Bayesian non-parametric prior induced by a predictor-dependent Bernstein polynomial processes. We construct a Bayes factor for this problem and discuss its consistency. Finally, simulated and real-life data are used to assess the performance of the proposed approach.
⊕ Joint work with *Antonio Canale and Pierpaolo De Blasi - University of Torino and Collegio Carlo Alberto, Italy*.

Marco Battiston Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Bocconi University, Italy
► *Multi-armed bandit for species discovery: a Bayesian nonparametric approach.*
▷ Let $\{\Pi_1, \ldots, \Pi_J\}$ be J discrete populations of animals supported on the same unknown set and let $\{G_1, \ldots, G_J\}$ the corresponding unknown distribution functions. The problem we explore in this work is how to sequentially sample from these populations, in order to maximize the total number of distinct animals observed. This problem is relevant, for instance, in applications in genomics, ecology and biology.
We adopt a Bayesian nonparametric approach and we use a Hierarchical Pitmar Yor prior as nonparametric prior for $\{G_1, \ldots, G_J\}$. As a consequence of the hierchical structure, the J unknown distributions share the same random support (that of their common random base measure). Given this prior choice, we propose a sequential rule that, at every time step, given the information available up to that point, selects the population from which to collect the next observation. Rather than simply picking the population with the highest posterior estimate of producing a new value, the proposed rule includes a Thompson sampling step to better balance the exploration-exploitation trade-off, inherent in any bandit problem. We propose two versions of the algorithm, one for incidence data and another for abudance one.
Performances of the algorithm are assessed through simulations and compared to other possible

strategies. Finally, an application with a real dataset of phylogenetic data is presented.
⊕ Joint work with *Stefano Favaro, Yee Whye Teh*.

Andrew Bean Contributed 6B, Thursday, 6/25/15, 4:20 pm - 4:40 pm

Ohio State University, United States

► *Transformation and Bayesian Density Estimation*

▷ Dirichlet-process mixture models, favored for their large support and for the relative ease of their implementation, are popular choices for Bayesian density estimation. However, despite the models' flexibility, the performance of density estimates suffers in certain situations, in particular when the true distribution is skewed or heavy tailed. We detail a method that improves performance in a variety of settings by initially transforming the sample, choosing the transformation to facilitate the subsequent density estimation procedure. The effectiveness of the method is demonstrated under a variety of simulated scenarios, and in an application to body mass index (BMI) observations from Ohio counties.
⊕ Joint work with *Xinyi Xu, Steven MacEachern*.

Anirban Bhattacharya Invited 5, Wednesday, 6/24/15, 11:30 am - 12 Noon

Assistant Professor, USA

► *Bernstein von Mises theorems in Wasserstein distance*

▷ We study Bernstein von-Mises (BvM) phenomenon in the function estimation setting using the Wasserstein distance as a measure of discrepancy between the posterior and its Gaussian approximation. A general scheme of obtaining rates of convergence results from Wasserstein BvMs is described and applied to random design Gaussian process regression on the entire real line.
⊕ Joint work with *Debdeep Pati*.

Prithwish Bhaumik Contributed 4B, Wednesday, 6/24/15, 5:00 pm - 5:20 pm

North Carolina State University, United States

► *Efficient Bayesian estimation and uncertainty quantification in differential equation models*

▷ In engineering, physics, biomedical sciences and many other fields the regression function is known to satisfy a system of ordinary differential equations (ODEs). Our interest lies in the unknown parameters involved in the ODEs. When the analytical solution of the ODEs is not available, one approach is to use numerical methods to solve the system. A four stage Runge-Kutta (RK4) method is one such method. The approximate solution can be used to construct an approximate likelihood. We assign a prior on the parameters and then draw posterior samples, but this method may be computationally expensive. Bhaumik and Ghosal (2014) considered a two-step approach of parameter estimation based on integrated squared error, where a posterior is induced on the parameters using a random series based on the B-spline basis functions. The parameter is estimated by minimizing the distance between the nonparametrically estimated derivative and the derivative suggested by the ODE. Although this approach is computationally fast, the Bayes estimator is not asymptotically efficient. In this paper we also suggest a modification of the two-step method by directly considering the distance between the function in the nonparametric model and that obtained from RK4 method. We study the asymptotic behavior of the posterior distribution of $\boldsymbol{\theta}$ in both RK-4 approximate likelihood based and modified two-step approaches and establish a Bernstein-von Mises theorem which assures that Bayesian uncertainty quantification matches with the frequentist one. We allow the model to be misspecified in that the true regression function may lie outside the ODE model. Unlike in the original two-step procedure, the precision matrix matches with the Fisher information matrix.

⊕ Joint work with *Subhashis Ghosal*.

David Blei Invited 1, Monday, 6/22/15, 10:30 am - 11:00 am

Columbia University, USA

► *Stochastic Variational Inference for Bayesian Nonparametrics*

▷ The central computational problem of Bayesian statistics is posterior inference, the problem of approximating the conditional distribution of latent variables given observations. Approximate posterior inference algorithms have revolutionized Bayesian statistics, revealing its potential as a usable and general-purpose language for data analysis. Bayesian statistics, however, has not yet reached this potential. Statisticians and scientists regularly encounter massive data sets but existing approximate inference algorithms do not scale to solve such problems.

In this talk, I will describe stochastic variational inference, a general algorithm for approximating posterior distributions that are conditioned on massive data sets. Stochastic inference is easily applied to a large class of hierarchical models, including time-series models, factor models, and Bayesian nonparametric models. I will demonstrate its application to Bayesian nonparametric topic models fit with millions of articles. Stochastic variational inference opens the door to scalable Bayesian computation for modern data analysis.

⊕ Joint work with *Matt Hoffman, John Paisley, and Chong Wang*.

Sudip Bose Contributed 3A, Tuesday, 6/23/15, 4:00 pm - 4:20 pm

Department of Statistics, The George Washington University, USA

► *Generating priors and Polya posteriors*

▷ The Polya posterior approach (Ghosh and Meeden, 1997) is a Bayesian approach to finite population sampling.

Levi Boyles Contributed 4A, Wednesday, 6/24/15, 4:40 pm - 5:00 pm

University of Oxford, United Kingdom

► *Hierarchical Clustering with Aldous' Beta-Splitting Trees*

▷ Beta-splitting trees, as described by Aldous, is a distribution on binary tree structures defined by a discrete time continuous state process. The process is described by a hierarchy of "splitting points" which partitions the unit interval. A set of individuals live on the interval, and each splitting event thus partitions the individuals. For a finite set of individuals, and for particular choices of splitting distributions, integrating out the splitting points recovers many familiar distributions on trees, for example the uniform distribution, and also the Yule tree distribution (the distribution on tree structures induced by Kingman's Coalescent). This process stops just short of defining a de Finetti representation for discrete tree distributions, as a splitting point can fail to split a finite set of points into two nonempty sets. We show that we can account for these "trivial splits" in the Yule tree distribution case to recover an analytic description of the de Finetti representation. We define consistent time variables using this representation. This representation is more amenable to MCMC inference (as compared to Kingman's Coalescent), and we perform inference using this model on a tumor subclonal reconstruction task.

⊕ Joint work with *Max Welling, Yee Whye Teh*.

Tamara Broderick Contributed 2B, Tuesday, 6/23/15, 9:40 am - 10:00 am

UC Berkeley/MIT, USA

► *Posteriors, conjugacy, and exponential families for completely random measures*

▷ We demonstrate how to calculate posteriors for general Bayesian nonparametric priors and likelihoods based on completely random measures (CRMs). We further show how to represent Bayesian nonparametric priors as a sequence of finite draws using a size-biasing approach—and how to represent full Bayesian nonparametric models via finite marginals. Motivated by conjugate priors based on exponential family representations of likelihoods, we introduce a notion of exponential families for CRMs, which we call exponential CRMs. This construction allows us to specify automatic Bayesian nonparametric conjugate priors for exponential CRM likelihoods. We demonstrate that our exponential CRMs allow particularly straightforward recipes for size-biased and marginal representations of Bayesian nonparametric models. Along the way, we prove that the gamma process is a conjugate prior for the Poisson likelihood process and the beta prime process is a conjugate prior for a process we call the odds Bernoulli process. We deliver a size-biased representation of the gamma process and a marginal representation of the gamma process coupled with a Poisson likelihood process.
⊕ Joint work with *Ashia C. Wilson, Michael I. Jordan*.

Annalisa Cadonna Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University of California, Santa Cruz, USA
► *Bayesian nonparametric modeling for multiple spectral densities*
▷ We develop a new Bayesian nonparametric approach to spectral density estimation. The main idea is to approximate the Whittle likelihood with a local mixture of Normal distributions, with frequency dependent weights and means. To achieve this, we use a Dependent Dirichlet process mixture or a geometric stick breaking process mixture. The main application is to neuroscience. Electroencephalogram (EEG) data are typically recorded at multiple location. Hence, there is need for a model that can be extended to multiple spectral densities. The fact that the spectral density estimation model has been cast in a nonparametric mixture modeling framework is key for effective inference methods as we seek to extend to modeling for multiple spectral densities, because it allows convenient implementation and inference.
⊕ Joint work with *Athanasios Kottas, Raquel Prado*.

Diana Cai Contributed 5A, Thursday, 6/25/15, 9:00 am - 9:20 am
Gamalon Labs, USA
► *Priors on Exchangeable Directed Graphs*
▷ Directed graphs occur throughout statistical modeling of networks, and exchangeability is a natural assumption when the ordering of vertices does not matter. Exchangeable undirected graphs admit certain key structural results, many of which extend to the directed case – although with additional complexities arising from the 4 possibilities for directed edges among a pair of vertices. Exchangeable directed graphs are characterized by a sampling procedure given by the Aldous-Hoover theorem. Such a procedure is determined by specifying a distribution on measurable objects known as digraphons. Most of the existing work on exchangeable graphs models has focused on undirected graphs, and little attention has been placed on priors for exchangeable directed graphs. We present several Bayesian nonparametric models for exchangeable directed random graphs, including infinite discrete and continuous models, such as a directed variant of the infinite relational model (Kemp et al., 2006) and Gaussian process priors on digraphons, extending work of Lloyd et al. (2012). We demonstrate applications of inference for these models on directed network data. We also examine analogous priors on other exchangeable structures such as posets and hypergraphs.
⊕ Joint work with *Nathanael Ackerman, Cameron Freer*.

Christopher Calderon Contributed 4A, Wednesday, 6/24/15, 4:20 pm - 4:40 pm

Ursa Analytics, United States

► *Inferring Latent States and Refining Force Estimates via Hierarchical Dirichlet Process Modeling in Single Particle Tracking Experiments*

▷ Recent advances in microscopy allow researchers to accurately monitor the motion of individual biomolecules (proteins, DNA, etc.) in living cells. These experiments produce multiple trajectories characterizing complex heterogeneous *in vivo* environments. High-throughput trajectory analysis has potential to answer numerous open questions in molecular biology. However, cataloging the rich spatio-temporal dynamics experienced in the cell poses several time series modeling challenges. Latent factors (e.g., due to interactions of the molecule with organelles in the cell) modulate the dynamics, "process" (or thermal) noise is often commensurate with measurement noise, and nonlinear spatially dependent force fields govern the dynamics of the trajectory; reliably quantifying all of these factors is of interest to researchers [1]. Although global dynamical models are not known *a priori*, the temporal and spatial resolution allow data-driven analysis techniques. Bayesian nonparametric techniques accounting for temporal dependence and an unknown number of states [2] show great promise as a tool for processing "single-molecule" data [3-4]. Elicitation of priors poses substantial challenges [3]. Computation of posteriors and quantifying uncertainty of parameters is not as important as testing model assumptions against data via goodness-of-fit hypothesis testing in this domain [1,3,4]. In this talk, we provide several examples illustrating where a pragmatic nonparametric Bayesian / frequentist approach can help in analyzing molecular motion from position vs. time data measured *in vivo*. Open issues facing this domain where new nonparametric Bayesian techniques will likely make a large impact will also be sketched.

[1] Calderon, C.P. , Thompson, M.A., Casolari, J.M., Paffenroth, R.C., and Moerner, W.E., J Phys Chem B 117 (49), 2013.

[2] Fox E., Sudderth E.B., Jordan, M.I., and Willsky A.S., IEEE Transactions on Signal Processing 59 (4), 2011.

[3] Calderon, C.P., Molecules, 19, 2014.

[4] Calderon, C.P. and Bloom, K., PLOS Comp. Bio. (under review).

⊕ Joint work with *Kerry Bloom*.

Federico Camerlenghi Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

Universitá degli Studi di Pavia, Italy

► *Nonparametric hierarchical models and their predictive structure*

▷ Hierarchical models are becoming increasingly popular in Bayesian nonparametric inference, being especially appropriate to analyze data from different and related studies. The hierarchy has the effect to induce a borrowing strength phenomenon across diverse groups of observations. A large variety of applications can be handled through these models, the most popular being classification of documents in a corpus, on the basis of different topics. The hierarchical models introduced so far are based on the Dirichlet and the Pitman-Yor processes. We will show how to construct them on the basis of a broader class of discrete random probability measures that can be represented as transformations of random measures, extending the results already known in literature. From the theoretical point of view, we aim to investigate the properties of the induced partition structure and the prediction rules. On the other hand we propose applications to different problems in genomics and survival analysis. In the first case, conditionally on a basic population, we will devise a novel Gibbs sampler algorithm in order to determine: the number of new species detected in an additional sample of arbitrary length, the probability of discovering a new species and the sample coverage. With respect to inference on survival data, we will develop a Bayesian nonparametric model to manage heterogeneous groups of observations, such as in clinical trials where groups of patients undergo different treatments in different hospitals.

Antonio Canale Contributed 6B, Thursday, 6/25/15, 4:40 pm - 5:00 pm
University of Turin and Collegio Carlo Alberto, Italy
► *Multiscale Bernstein Polynomials for Densities*
▷ Our focus is on constructing a multiscale nonparametric prior for densities. The Bayes density estimation literature is dominated by single scale methods, with the exception of Polya trees, which favor overly-spiky densities even when the truth is smooth. We propose a multiscale Bernstein polynomial family of priors, which produce smooth realizations that do not rely on hard partitioning of the support. At each level in an infinitely-deep binary tree, we place a beta dictionary density; within a scale the densities are equivalent to Bernstein polynomials. Using a stick-breaking characterization, stochastically decreasing weights are allocated to the finer scale dictionary elements. A slice sampler is used for posterior computation, and properties are described. The method characterizes densities with locally-varying smoothness, and can produce a sequence of coarse to fine density estimates. An extension for Bayesian testing of group differences is introduced and applied to DNA methylation array data.
⊕ Joint work with *David B. Dunson*.

Francois Caron Invited 2, Monday, 6/22/15, 2:00 pm - 2:30 pm
University of Oxford, United Kingdom
► *Sparse random graphs with exchangeable point processes*
▷ Statistical network modeling has focused on representing the graph as a discrete structure, namely the adjacency matrix, and considering the exchangeability of this array. In such cases, the Aldous-Hoover representation theorem (Aldous, 1981; Hoover, 1979) applies and informs us that the graph is necessarily either dense (the number of edges increases quadratically with the number of nodes) or empty. Here, we instead consider representing the graph as a point process on the positive quadrant. For the associated definition of exchangeability in this continuous space, we rely on the Kallenberg representation theorem (Kallenberg, 2005). We show that for certain choices of the specified graph construction, our network process is both exchangeable and sparse with power-law degree distribution. In particular, we build on the framework of completely random measures (CRMs) and use the theory associated with such processes to derive important network properties, such as an urn representation for network simulation. The CRM framework also provides for interpretability of the network model in terms of node-specific sociability parameters, with properties such as sparsity and power-law behavior simply tuned by three hyperparameters.
⊕ Joint work with *Emily B. Fox*.

Carlos Carvalho Invited 5, Wednesday, 6/24/15, 11:00 am - 11:30 am
The University of Texas, USA
► *Utility Based Model Section for Bayesian Nonparametric Modeling using Trees*
▷ Over the last several years, dramatic advances in Bayesian modeling and computation have given us powerful tools for flexible fitting of high dimensional relationships. However, the flexibility and complexity of the modeling procedures comes at a price: we may have difficulty under-

standing what our models have found. In particular, we are often interested in finding a simple model that works well, with variable selection being an important special case. Traditionally Bayesian approaches to search for a simple model have emphasized the specification of priors on models and computation of the posterior on models. In this paper we emphasize the role of utility in choosing a model. We use fits of the posterior predictive using binary tree models to search for simple structure. Tree models are computationally fast and capable of capturing complex structure so that we can feasibly search for model simplifications that are not too simple in that important variables and complexity (e.g. nonlinearity) are not lost.
⊕ Joint work with *P. Richard Hahn and Robert McCulloch*.

Lu Cheng Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Department of Computer Science, Aalto University, Finland
► *Integrating prior biological knowledge into machine learning models for predicting drug responses*
▷ For complex genetic diseases such as rheumatoid arthritis, treatment effects can vary significantly among different patients. We integrate prior biological knowledge into machine learning models to explain the differences. We use GWAS, CCA based selection, PharmGKB, differential gene expression analyses to select the SNPs. We use GEMMA (linear mixture model) and BEMKL (multiple kernel methods) for our predictions. Our results show that the clinical information contributes the most for explaining the differences. Genetic information contributes a relatively small amount for the predictions, which is validated by comparing with clinical only predictions and random SNP set predictions. Correctly utilizing the methods is also very important. Our results show that changing the settings within these methods can create huge differences in predictions.
⊕ Joint work with *Lu Cheng, Gopal Peddinti, Muhammad Ammad-ud-din, Alok Jaiswal, Himanshu Chheda, Suleiman Ali Khan, Kerstin Bunte, Jing Tang, Matti Pirinen, Pekka Marttinen, Janna Saarela, Jukka Corander, Krister Wennerberg, Samuel Kaski, Tero Aittokallio*.

William Cipolli Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University of South Carolina, USA
► *Multiple Testing via an Approximate Finite Polya Tree*
▷ Multiple testing, or multiplicity, problems often require testing multiple means with the assumption that we will reject infrequently, as motivated by the need to analyze DNA microarray data. The goal is to keep the combined rate of false discoveries and non-discoveries as small as possible. We propose a discrete approximation to a Polya tree prior – that enjoys fast, conjugate updating, centered at the usual normal distribution, thus generalizing Scott and Berger (2006) and Ausin et al. (2011) to a nonparametric setting. This new technique and the advantages of our approach are demonstrated using simulation accompanied by a Java web application. The numerical studies demonstrate that our new procedure shows promising FDR and better estimation of key values in the mixture model with very reasonable computational speed.
⊕ Joint work with *Timothy Hanson, Alexander McLain*.

Glen Colopy Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University of Oxford, United Kingdom
► *Prior Regularization and Likelihood-Based Deterioration Detection Using Gaussian Processes in the Stepdown-ward*
▷ 1-in-7 electronically monitored patients in the stepdown-ward (a high-acuity hospital environment) suffer cardiac arrest or readmission to the ICU under emergency conditions. Advanced detection of patient deterioration could prevent many of these adverse events. The current state-

of-the-art relies on heuristic threshold-based ”Early Warning Scores” (which are often manual), or kernel density estimates of a representative population’s vital-sign data. Both methods assume an i.i.d. relationship between vital-sign data and can only assess a patients present risk. They are unable to forecast future risk. Gaussian Process (GP) regression addresses both the i.i.d. assumption and forecasting: dependencies between vital-sign data within a timeseries are modelled by the GP, allowing explicit modelling of vital-sign dynamics; the plausible range of future values is represented as a distribution over a function of time. Additionally, models are constructed on-line, in a patient-specific manner, offering the potential for increased sensitivity to ”abnormal” physiological events in comparison to existing population-based approaches. A key factor in the applicability of the GPs to vital-sign data is ensuring a suitably regularised model. Robust fitting of the data is obtained by defining prior distributions over the hyperparameters of the GP covariance function, while the latter may be defined in terms of linear combinations of basic kernels according to the particular dynamics of the vital-sign being modelled. Once fitted to a patient’s vital-sign data, GP models are amenable to both traditional regression (for detecting abnormalities in newly-observed data) or timeseries-clustering (to describe patient dynamics in terms of previously observed clusters within representative populations).
⊕ Joint work with *Stephen J. Roberts, David A. Clifton*.

David Dahl Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Brigham Young University, United States
▶ *Attraction Indian Buffet Process*
▷ We propose a Bayesian latent feature model for non-exchangeable data indexed by pairwise similarity between items. We formulate the distribution such that similar items are more likely to share features, while still preserving many of the familiar properties of the original Indian buffet process (IBP) for exchangeable data. For example, the expected number of non-zero columns is equal to that of the IBP. Posterior inference is straightforward since our proposal has an explicit probability mass function. We compare our method to the distance dependent Indian buffet process (dd-IBP) and demonstrate an application to non-exchangeable data.
⊕ Joint work with *Arthur Lui*.

Patrick Dallaire Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Laval University, Canada
▶ *A Bayesian nonparametric prior on infinite directed acylic graphs*
▷ We define a probability distribution on infinite directed acyclic graphs. The model is based on a Beta process prior controlling the outgoing connection probabilities in the continuous space of orders. The Beta process is marginalized out, yielding a probability distribution on the joint space of adjacency matrices and orders. Another marginalization is done by defining a set of observables nodes, finding their ancestors and marginalizing all connections leading to a non-ancestor node. The remaining adjacency submatrix is sparse as for finite set of observable nodes, the process generates graphs having a finite number of ancestor with probability 1. The new stochastic process is intended for Bayesian nonparametric structure learning in Bayesian networks, but is general enough to be applied in other structure learning problems involving directed acyclic graphs. A reversible jump Markov Chain Monte Carlo algorithm is proposed to perform structure inference.
⊕ Joint work with *Ludovic Trottier, Philippe Giguére, Brahim Chaib-draa*.

Priyam Das Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

North Carolina State University, United States

▶ *Bayesian Quantile Regression Using Random Series*

▷ We consider Bayesian methods for simultaneous quantile regression. A representation of quantile function is given by a convex combination of two monotone increasing function not depending on the prediction variables. In a Bayesian approach, a prior is put on quantile functions by putting prior distribution on ξ_1 and ξ_2 independently. To comply with the monotonicity constraint of the monotone curves which are used to estimate the slope and the intercept, we consider two possible approaches to conduct a finite random series which naturally obeys the shape restrictions when the coefficients are ordered accordingly within [0,1].In the first method, we used Bernstein polynomial. In the second approach, we consider a finite random series prior based on B-splines with coefficients in increasing order between 0 and 1. We compare both approaches with a Bayesian method using Gaussian process prior and a non-Bayesian approach through an extensive simulation study. An application to a data on hurricane activities in the Atlantic is given.

⊕ Joint work with *Dr. Subhashis Ghoshal*.

Pierpaolo De Blasi Contributed 7A, Friday, 6/26/15, 10:30 am - 10:50 am

University of Torino and Collegio Carlo Alberto, Italy

▶ *Posterior asymptotics under the sup-L_1 distance in nonparametric conditional estimation*

▷ In this paper we study posterior consistency in conditional density estimation with respect to the supremum L_1 norm. Compared to the integrated L_1, which is predominant in the literature on posterior asymptotics, it allows consistency for prediction at any designated conditional density. We model the conditional density as in a regression tree model, that is via a sample size dependent sequence of increasingly finer partitions of the predictor space and by letting the conditional density be the same across all predictor values in a partition set. Each conditional density is modeled independently by a Dirichlet process mixture, hence, the prior specifies a type of dependence between conditional densities which disappears after a certain number of observations has been observed. The rate at which the number of partition sets increases with the sample size determines when the dependence between pairs of conditional densities is set to zero and, ultimately, drives posterior consistency. We also provide some results on what type of convergence rates the posterior can achieve.

⊕ Joint work with *Stephen G. Walker*.

Nicola Donelli Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

Universitá degli Studi Milano-Bicocca, Italy

▶ *Semiparametric Bayesian vector multiplicative error model*

▷ The purpose of our research is to generalize the semiparametric Bayesian MEM of Mira and Solgi (2013) to a multivariate setting, namely vector MEM (vMEM). In the proposed vMEM, the positive vector x_t, given the information set available at time $t-1$, is modeled as

$$x_t = \mu_t \odot \varepsilon_t.$$

In the vMEM literature the components of the vector of innovations, ε_t, are typically modeled with Gamma distributions (with a single free parameter to ensure unit mean), and the specification of the conditional mean μ_t depends on the specific application. In our approach, the vector of innovations ε_t is an i.i.d. process from a unit mean multivariate distribution supported on the positive orthant and we model this distribution nonparametrically, using a Dirichlet process mixture of multivariate log-Normal distributions. We also propose a general parametric model for

the conditional mean μ_t that nests most of the specifications used in the vMEM literature. This model could be used to study the dynamic interactions among different volatility measures or among volatilities in different market indices (just to mention two possible applications). To perform Bayesian inference we first expand the parameter space, considering an unidentifiable model with a non-unit mean parameter for the innovations. We then apply the slice sampler algorithm to the parameter-expanded model and finally post-process this sample to obtain a sample from the posterior of the original model. In order to efficiently sample from the full-conditional distribution of the parameters of the conditional mean, an adaptive version of the Metropolis Adjusted Langevin Algorithm is used.
⊕ Joint work with *Reza Solgi, and Antonietta Mira*.

David Dunson Invited 6, Wednesday, 6/24/15, 2:30 pm - 3:00 pm
Duke University, USA
► *Scaling up Bayes computation to massive data*
▷ This talk focuses on recent developments in scaling up Bayesian computation to massive data settings, motivated in particular by nonparametric models involving large numbers of parameters and multimodal posteriors. We focus on a new idea based on the Wasserstein barycenter of subset posteriors, which can be shown to accurately approximate the full data posterior. This motivates an extremely fast algorithm in which we run MCMC in parallel for different subset posteriors and then use a sparse linear program to combine the results. We illustrate this in an example involving recommender systems, involving 100 million users and tens of thousands of items. This approach dominates scalable variational methods, avoiding problems with underestimating posterior convergence and only characterizing a single mode. We also discuss other promising directions in scaling up Bayes to truly massive settings.

David Duvenaud Invited 6, Wednesday, 6/24/15, 2:00 pm - 2:30 pm
Harvard University, USA
► *Scaling Up Bayesian Optimization*
▷ Bayesian optimization is a method for performing optimization of expensive and noisy objective functions. The core idea is to use a Bayesian nonlinear regression model, such as a Gaussian process, for the expensive objective and use this model to make information-theoretically grounded decisions about where to evaluate next. Although these ideas are quite old, recent advances in inference techniques have resulted in renewed interest. Scalability to large parallel systems has remained challenging, however. In this talk, I will discuss recent work to deal with the case where hundreds or thousands or thousands of evaluations can be performed in parallel. I will also describe a new nonparametric regression model (and positive definite kernel) that is well-suited for projecting the outcome of long-running iterative objectives, allowing them to be stopped early when appropriate.
⊕ Joint work with *Jasper Snoek, Kevin Swersky*.

Matt Edwards Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University of Auckland, New Zealand
► *Bayesian semiparametric spectral density estimation in gravitational wave data analysis*
▷ Astronomy is entering a new and exciting era, with the second generation of ground-based gravitational wave (GW) interferometers (Advanced LIGO, Advanced Virgo, and KAGRA) expected to come on-line in the next few years. The importance of extracting astrophysical informa-

tion encoded in GW signals is paramount, and since observations are extremely noisy, accurate predictions rely on an honest characterisation of detector noise. The default model in the GW data analysis literature assumes detector noise is stationary and Gaussian distributed, with a known power spectral density (PSD) that is usually estimated using off-source data. Real GW data often experience departures from such assumptions, and so we propose a Bayesian semiparametric approach to improve this. Using a nonparametric Bernstein polynomial prior on the PSD with weights attained via a Dirichlet process distribution, we update the approximate Whittle likelihood, and sample from the posterior of the PSD using a Metropolis-within-Gibbs sampler. We simultaneously estimate the reconstruction parameters of a rotating core collapse GW burst that has been embedded in simulated Advanced LIGO noise. We also discuss an approach to deal with non-stationary data by breaking longer data streams into smaller and locally stationary components.
⊕ Joint work with *Renate Meyer, Nelson Christensen*.

Stefano Favaro Contributed 7B, Friday, 6/26/15, 10:30 am - 10:50 am
Universitá di Torino and Collegio Carlo Alberto, Italy
► *Rediscovery of Good-Turing estimators via Bayesian nonparametrics*
▷ A full range of statistical approaches, parametric and nonparametric as well as frequentist and Bayesian, have been proposed for estimating discovery probabilities. These approaches have found many applications in ecology, and their importance has grown considerably in recent years driven by challenging applications arising from bioinformatics, genetics, linguistics, designs of experiments, machine learning, combinatorics, etc. In this talk we present the interplay between the celebrated Good-Turing approach, which is a frequentist nonparametric approach developed in the 1940s, and a Bayesian nonparametric approach recently introduced in the literature. Under the assumption of Gibbs-type priors we show that Bayesian nonparametric estimators of discovery probabilities are approximately equal, for a large sample size, to suitably smoothed Good-Turing estimators. The smoothing rule is induced by the Gibbs-type prior and it depends on the order of approximations. We present an explicit characterization of the smoothing rules induced by considering first and second order approximations with respect to two noteworthy examples of Gibbs-type priors, namely the two parameter Poisson-Dirichlet prior and the normalized generalized Gamma prior. Our results are illustrated through a simulation study and the analysis of a dataset arising from linguistic.
⊕ Joint work with *Julyan Arbel, Bernardo Nipoti, and Yee Whye Teh*.

Gilbert Fellingham Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Dept. of Statistics, Brigham Young University, USA
► *Predicting Home Run Production in Major League Baseball Using a Bayesian Semiparametric Model*
▷ We predict home run hitting performance for Major League Baseball players using a Bayesian semiparametric model. Following previous research in optimal performance in baseball, we include in the model effects for era of birth, season of play, and home ball park. We estimate performance curves for each player using quartic polynomials. We use a Dirichlet process prior on the unknown distributions for the coefficients of the polynomials, and parametric priors for the other effects. Dirichlet process priors are useful in prediction for two reasons: (1) an increased probability of obtaining more precise prediction accompanies the increased flexibility of the prior specification, and (2) the clustering inherent in the Dirichlet process provide the means to share information across players. Data from 3735 players who competed from 1871 to 2008 were used to fit the model. We used data for 22 players who played from 2009 to 2013 to evaluate the predictive

performance of the model. We used what we called 'pure performance' curves to predict future performance.
⊕ Joint work with *Jared Fisher*.

Yessica Fermin Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

TU Dortmund University, Faculty of Statistics, Germany

► *Inferring the Structure of Protein-Networks by Nonparametric Dynamic Bayesian Networks*

▷ In systems biology, probabilistic graphical models such as Bayesian networks and dynamic Bayesian networks have been widely applied as a useful tool to model complex biochemical systems. However, these two particular network approaches have the drawback of making parametric assumptions, which in real-world applications are often violated. To address this, Ickstadt et al. (2011) developed a nonparametric version of Gaussian Bayesian networks and in the present study, we propose a nonparametric version of dynamic Bayesian networks by a combination of dynamic Bayesian networks (Ghahramani, 1997) and of nonparametric Bayesian networks. The suitability of the new network approach in inferring the structure of protein-networks in a non-Gaussian dynamic situation is compared with the suitability of Gaussian dynamic Bayesian networks and of nonparametric Bayesian networks in a simulation study. Furthermore, the new approach is illustrated using multicolor live cell imaging data sets, in which the levels of four different proteins were measured over multiple time points in individual cell-matrix adhesion sites. References: [1] Ghahramani, Z. (1997) Learning Dynamic Bayesian Networks. Lecture Notes in Computer Science 1387, 168-197. [2] Ickstadt, K., Bornkamp, B., Grzegorczyk, M., Wieczorek, J., Sheriff, M.R., Grecco, H.E. and Zamir, E(2011) Nonparametric Bayesian Networks. In: Bernardo, Bayarri, Berger, Dawid, Heckerman, Smith and West (eds.): Bayesian Statistics 9, Oxford University Press, 283-316.
⊕ Joint work with *Katja Ickstadt, Malik R. Sheriff, Sarah Imtiaz, Eli Zamir*.

Tamara Fernandez Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

Department of Statistics, University of Oxford, United Kingdom

► *Gaussian processes for survival analysis*

▷ We introduce a non-parametric Bayesian model for survival analysis data which places a function of a Gaussian process as prior over the hazard function. One of the main advantages of this model is that it flexibly incorporates covariates without imposing any additional constraints or parametric forms over the hazard function. Exact inference in this model is made feasible by using a data augmentation approach based on Poisson thinning. The approach allows any kind of censorship to be present on the data.
⊕ Joint work with *Yee Whye Teh*.

Marco Ferreira Contributed 3A, Tuesday, 6/23/15, 4:40 pm - 5:00 pm

Virginia Tech, United States

► *Dynamic Multiscale Spatiotemporal Models for Poisson Data*

▷ We propose a new class of dynamic multiscale models for Poisson spatiotemporal processes. Specifically, we use a multiscale spatial Poisson factorization to decompose the Poisson process at each time point into spatiotemporal multiscale coefficients. We then connect these spatiotemporal multiscale coefficients through time with a novel Dirichlet evolution. Further, we propose a simulation-based full Bayesian posterior analysis. In particular, we develop filtering equations for updating of information forward in time and smoothing equations for integration of information backward in time, and use these equations to develop a forward filter backward sampler for

the spatiotemporal multiscale coefficients. Because the multiscale coefficients are conditionally independent a posteriori, our full Bayesian posterior analysis is scalable, computationally efficient, and highly parallelizable. Moreover, the Dirichlet evolution of each spatiotemporal multiscale coefficient is parametrized by a discount factor that encodes the relevance of the temporal evolution of the spatiotemporal multiscale coefficient. Therefore, the analysis of discount factors provides a powerful way to identify regions with distinctive spatiotemporal dynamics. Finally, we illustrate the usefulness of our multiscale spatiotemporal Poisson methodology with two applications. The first application examines mortality ratios in the state of Missouri, and the second application considers tornado reports in the American Midwest.
⊕ Joint work with *Thais Fonseca*.

Catherine Forbes Contributed 7B, Friday, 6/26/15, 10:50 am - 11:10 am
Monash University, Australia
► *Bayesian restricted likelihood-based instrumental variables regression*
▷ In recent work, Lewis, Lee and MacEachern (2012, 2014) suggest a Bayesian Restricted Likelihood framework for obtaining posterior inference for structural parameters based on summary statistics rather than a complete sample of observations. In this talk, an application of the Bayesian Restricted Likelihood approach will be explored as a robust method for handling the presence of outliers in an instrumental variables regression setting.
⊕ Joint work with *Zhichao Liu*.

Nicholas Foti Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University of Washington, United States
► *Streaming Variational Inference for Bayesian Nonparametric Mixture Models*
▷ In theory, Bayesian nonparametric (BNP) models are well suited to streaming data scenarios due to their ability to adapt model complexity with the observed data. Unfortunately, such benefits have not been fully realized in practice; existing inference algorithms are either not applicable to streaming applications or not extensible to BNP models. For the special case of Dirichlet processes, streaming inference has been considered. However, there is growing interest in more flexible BNP models building on the class of normalized random measures (NRMs). We work within this general framework and present a streaming variational inference algorithm for NRM mixture models. Our algorithm is based on assumed density filtering (ADF), leading straightforwardly to expectation propagation (EP) for large-scale batch inference as well. We demonstrate the efficacy of the algorithm on clustering documents in large, streaming text corpora.
⊕ Joint work with *Alex Tank, Emily Fox*.

Emily Fox Keynote II, Wednesday, 6/24/15, 9:00 am - 10:00 am
University of Washington, United States
► *Scalable Modeling and Inference for Complex Data Streams*
▷ Data streams of increasing complexity are being collected in a variety of fields ranging from neuroscience, genomics, and environmental monitoring to e-commerce based on technologies and infrastructures previously unavailable. With the advent of MCMC combined with the computational power to implement such algorithms, deploying increasingly expressive models has been a focus in recent decades. However, these modern datasets typically pose a challenge to such, now classical, approaches. For example, (i) individual data streams may provide scarce observations relative to the complexity of models used to describe them (p ¿¿ n), (ii) large collections of data streams have intricate relationships to capture (again, p¿¿n), and (iii) the data streams are very

large or stream in continuously (large n). In theory, Bayesian nonparametrics play a critical role in addressing such problems by adapting model complexity. We examine two intertwined and complementary aspects of these recent challenges: scalable modeling and inference, specifically for data with complex dependencies. Our models focus on capturing sparse dependencies and lower-dimensional embeddings whereas the computational tools focus on algorithms that subsample data or rely solely on sufficient statistics of past observations. We also discuss some open challenges in seamlessly integrating modeling and computational choices.

Yarin Gal Contributed 1B, Monday, 6/22/15, 4:00 pm - 4:20 pm
University of Cambridge, United Kingdom
► *An Infinite Product of Sparse Chinese Restaurant Processes*
▷ We define a new process that gives a natural generalisation of the Indian buffet process (used for binary feature allocation) into categorical latent features. We take advantage of different limit parametrisations of the Dirichlet process and its generalisation the Pitman-Yor process. A specific sparse parametrisation is of interest. This parametrisation results in a draw of a Dirac measure in the limit. In the corresponding sparse Chinese restaurant process, a.s. all customers sit at a single table. We derive a new non-parametric process by taking the limit of an infinite product of these sparse Chinese restaurant processes. The derived process, as it turns out, exhibits very different behaviour from its individual components behaviour. Among the infinitely many restaurants, almost all restaurants are degenerate (have all customers sitting at a single table) and finitely many follow the law of an adapted Pitman-Yor process. This adapted process can be described by removing the first stick from the stick-breaking construction of the Pitman-Yor process. We describe the theoretical properties and introduce an urn scheme for the process. This generative procedure results in a partition product in which a set of points is partitioned in multiple ways into a finite number of partitions. We use the process to model multi-view clustering, in which we attempt to discover multiple systems of clusters over a single data set, and can be seen as a categorical generalisation of binary feature allocation models.
⊕ Joint work with *Tomoharu Iwata, Zoubin Ghahramani*.

Fengnan Gao Contributed 2A, Tuesday, 6/23/15, 9:00 am - 9:20 am
Leiden University, Netherlands
► *Posterior contraction rates for deconvolution of Dirichlet-Laplace mixtures*
▷ We study nonparametric Bayesian inference with location mixtures of the Laplace density and a Dirichlet process prior on the mixing distribution. We derive a contraction rate of the corresponding posterior distribution, both for the mixing distribution relative to the Wasserstein metric and for the mixed density relative to the Hellinger and Lq metrics.
⊕ Joint work with *Aad van der Vaart*.

Chao Gao Contributed 3B, Tuesday, 6/23/15, 4:00 pm - 4:20 pm
Yale University, USA
► *Some Results on Posterior Contraction of Matrix Estimation*
▷ In this talk, we present results on Bayes matrix estimation in covariance model, location model and network model. We discuss the limits of the classical prior-mass-and-testing technique to prove posterior contraction rates. When this technique is applied to loss functions such as matrix operator norm, it is impossible to obtain optimal rates. Novel techniques on proving sharp minimax posterior contraction rates using matrix functionals are developed to overcome this intrinsic

difficulty.
⊕ Joint work with *Harrison H. Zhou*.

Hong Ge Contributed 4A, Wednesday, 6/24/15, 4:00 pm - 4:20 pm

University of Cambridge, U.K.

▶ *Hierarchical infinite hidden Markov models*

▷ Hidden Markov models (HMMs) are an extremely popular class of models in sequential data analysis, with applications in fields as diverse as language modelling, computational biology, and finance. Despite this popularity standard HMMs are awkward in two regards: firstly, the number of latent states must be selected manually, and secondly, it is challenging to interpret the resulting latent states. We introduce the hierarchical infinite HMM (H-iHMM) to address these problems. Through use of a hierarchical generalisation of the Dirichlet process, the H-iHMM is able to learn an appropriate number of states from data analogously to the infinite HMM, but in addition organises the states in a tree structure. While a large number of states may optimally explain the data, the practitioner can choose to look at a lower granularity by considering subtrees rather than individual states. Organising hidden states in a tree structure also helps constrain the space of transition and emission probability matrices, thus potentially leading to a model that generalises better with limited data Neal (2003).
⊕ Joint work with *David Knowles, Zoubin Ghahramani*.

Chris Glynn Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

Duke Statistical Science, USA

▶ *Fully Bayesian Inference for a Dynamic Linear Topic Model*

▷ The proportion of a document characterized by a semantic topic may depend on properties of the document itself including the publisher and time of publication. We extend the Dynamic Topic Model of Blei and Lafferty (2006) by explicitly modeling document-level topic proportions with covariates and dynamic structure that includes time-trend and periodicity. A Markov Chain Monte Carlo algorithm that utilizes Polya-Gamma data augmentation is developed for posterior inference. Conditional independencies in the model and sampling are made explicit, and our MCMC algorithm is parallelized where possible to allow for inference in large corpora. To address computational bottlenecks associated with Polya-Gamma sampling, we appeal to the Central Limit Theorem to develop a Gaussian approximation to the Polya-Gamma random variable. This approximation is fast and reliable for parameter values relevant in the text- mining domain. Our model and inference algorithm are demonstrated on a corpus of Epigenetics abstracts from PubMed.
⊕ Joint work with *Surya Tokdar, David Banks, Brian Howard*.

Jim Griffin Contributed 2B, Tuesday, 6/23/15, 9:20 am - 9:40 am

University of Kent, United Kingdom

▶ *Compound random measures and their use in Bayesian nonparametrics*

▷ Many nonparametric priors have been proposed for related distribution and have found a wide-range of applications in statistics and machine learning. I will describe a new class of dependent random measures which we call compound random measures. These priors can be constructed with gamma, stable and generalized gamma process marginals and their dependence can be characterized using both the Levy copula and correlation function. Normalized version of these random measures can be used as dependent priors for related distributions and inference can be made using a slice sampling algorithm. The approach will be illustrated with data examples.

⊕ Joint work with *Fabrizio Leisen*.

Alessandra Guglielmi Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Politecnico di Milano, Italy
► *ϵ-NRMI mixtures for density and clustering estimation*
▷ The first aim of this work is the definition of a new class of nonparametric priors, which can be considered as an approximation of the distribution of a homogeneous normalized random measure with independent increments (HNRMI). We extend the construction of the ϵ-NGG processes in Argiento, Bianchini, Guglielmi (2015) to the more general class of HNRMIs. The new process is defined from the representation of HNRMIs as discrete measures where the weights are obtained by normalization of the jumps of a Poisson process, and the support consists of iid points, however considering only jumps larger than a positive threshold ϵ. Therefore, the number of jumps of the new process, called ϵ-NRMI, is a.s. finite. A prior distribution for ϵ can be elicited. We prove convergence in distribution, as ϵ goes to zero, of the ϵ-NRMI to the infinite-dimensional process, and derive its first moments. We assume this new discrete random probability as the mixing measure in a mixture model, with the ultimate purpose of density and cluster estimation. The main achievement of this work is the construction of a Gibbs sampler scheme to simulate from the posterior; in particular, we have built a conditional algorithm that uses a finite random number of jumps, but, on the other hand, it is easy to implement. Some hints on the choice of the prior distribution for ϵ will be given. Argiento, Bianchini, Guglielmi (2015). A blocked Gibbs sampler for NGG-mixture models via a priori truncation. Statistics and Computing.
⊕ Joint work with *Raffaele Argiento, Ilaria Bianchini*.

Subha Guha Contributed 4B, Wednesday, 6/24/15, 4:40 pm - 5:00 pm
University of Missouri, USA
► *Nonparametric Variable Selection, Clustering and Prediction for High-Dimensional Regression*
▷ The development of parsimonious models for reliable inference and prediction of responses in high-dimensional regression settings is often challenging due to relatively small sample sizes and the presence of complex interaction patterns between a large number of covariates. We propose an efficient, nonparametric framework for simultaneous variable selection, clustering and prediction in high-throughput regression settings with continuous or discrete outcomes, called VariScan. The VariScan model utilizes the sparsity induced by Poisson-Dirichlet processes (PDPs) to group the covariates into lower-dimensional latent clusters consisting of covariates with similar patterns among the samples. The data are permitted to direct the choice of a suitable cluster allocation scheme, choosing between PDPs and their special case, a Dirichlet process. Subsequently, the latent clusters are used to build a nonlinear prediction model for the responses using an adaptive mixture of linear and nonlinear elements, thus achieving a balance between model parsimony and flexibility. We investigate theoretical properties of the VariScan procedure that differentiate the allocations patterns of PDPs and Dirichlet processes both in terms of the number and relative sizes of their clusters. Additional theoretical results guarantee the high accuracy of the model-based clustering procedure, and establish model selection and prediction consistency. Through simulation studies and analyses of benchmark data sets, we demonstrate the reliability of VariScan's clustering mechanism and show that the technique compares favorably to, and often outperforms, existing methodologies in terms of the prediction accuracies of the subject-specific responses.
⊕ Joint work with *Veera Baladandayuthapani*.

Rajarshi Guhaniyogi Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

UNIVERSITY OF CALIFORNIA, SANTA CRUZ, USA

► *Bayesian Conditional Density Filtering for Streaming Data*

▷ We propose a Conditional Density Filtering (C-DF) algorithm for efficient online Bayesian inference. C-DF adapts MCMC sampling to the online setting, sampling from approximations to conditional posterior distributions obtained by propagating surrogate conditional sufficient statistics (a function of data and parameter estimates) as new data arrive. These quantities eliminate the need to store or process the entire dataset simultaneously and offer a number of desirable features. Often, these include a reduction in memory requirements and runtime and improved mixing, along with state-of-the-art parameter inference and prediction. These improvements are demonstrated through several illustrative examples including an application to high dimensional compressed regression. Finally, we show that C-DF samples converge to the target posterior distribution asymptotically as sampling proceeds and more data arrives.

⊕ Joint work with *Shaan Qamar, David Dunson*.

Luis Gutiérrez Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm

University of Chile, Chile

► *Bayesian nonparametric hypothesis testing for the two-sample problem*

▷ The two-sample problem is defined as the comparison between two populations on the basis of two independent samples, one from each population. In the field of biostatistics, it is of interest to identify whether a certain feature acts in the same way in both populations. Formally, let $Y_{1,1},\ldots,Y_{1,n_1}$ and $Y_{2,1},\ldots,Y_{2,n_2}$ be two samples measuring such feature, with G_1 and G_2 being the corresponding distribution functions. This problem has been traditionally addressed by evaluating if $G_1(Y_1) = G_2(Y_1+\theta)$ so that any comparison to be made between the two populations then depends on the parameter θ. However, the above approach can be embedded in a more general setting by replacing θ by $\Delta(Y_1)$, where $\Delta(\cdot)$ is referred to as the shift function. This extended approach can be used not only to test the hypothesis that $G_1 = G_2$ but also to make inferences about the set $\{y : \Delta(y) \neq 0\}$, which allows, for instance, to identify the members of the population for which a given treatment is beneficial. We propose to develop a novel and fully Bayesian nonparametric (BNP) model to make inferences on shift functions. Specifically, we develop a procedure to test the hypothesis that $G_1 = G_2$ and to make inferences about $\{y : G_1(y) \neq G_2(y)\}$, we also extend the procedures by considering confounders and predictors. We develop efficient Markov chain Monte Carlo schemes for exploring the corresponding posterior distributions. Comparisons based on simulations between the proposal and the parametric and nonparametric alternatives commonly used in Biostatistics are provided. Finally, the proposal is illustrated with simulated and real-life datasets.

⊕ Joint work with *Jorge Gonzalez, Andrés Felipe Barrientos*.

Creighton Heaukulani Contributed 1B, Monday, 6/22/15, 4:40 pm - 5:00 pm

University of Cambridge, United Kingdom

► *Random partition-based black-box inference for feature allocations and their generalizations*

▷ In recent work, Roy (2014) presented a black-box construction for random feature allocations (overlapping subsets) that takes the law of any random partition as an input. For example, if the random partition is a Dirichlet process, then the resulting feature allocation is an Indian buffet process (IBP; Griffiths and Ghahramani, 2006). Other special cases corresponding to various choices of random partitions can be seen to match different generalizations of the IBP. Taking advantage of these constructions, we develop corresponding black-box inference algorithms that only require

an inference procedure for the random partition. Furthermore, these constructions straightforwardly extend to feature allocations induced by so-called hierarchical base measures, and so too do the inference procedures. Similar constructions for other generalizations, such as random multisets (feature allocations accompanied by counts) and time-evolving feature allocations, produce corresponding black-box inference schemes. Inference techniques for random partitions have been well-developed in the literature, which now enable a wide variety of inference options for random feature allocations and their generalizations.
⊕ Joint work with *Daniel M. Roy*.

Katherine Heller Contributed 5A, Thursday, 6/25/15, 9:40 am - 10:00 am
Duke University, USA
► *The Bayesian Echo Chamber: Modeling Influence in Conversations*
▷ We present the Bayesian Echo Chamber, a new Bayesian generative model for social interaction data. By modeling the evolution of people's language usage over time, this model discovers latent influence relationships between them. Unlike previous work on inferring influence, which has primarily focused on simple temporal dynamics evidenced via turn-taking behavior, our model captures more nuanced influence relationships, evidenced via linguistic accommodation patterns in interaction content. The model, which is based on a multivariate Hawkes process, presents a new kind of language model, and allows us to jointly predict the temporal flow and content of conversation. We validate our model's ability to discover latent influence patterns using transcripts of arguments heard by the US Supreme Court and the movie "12 Angry Men". We showcase our model's capabilities by using it to infer latent influence patterns in Federal Open Market Committee meeting transcripts, demonstrating state-of-the-art performance at uncovering social dynamics in group discussions.
⊕ Joint work with *Richard Guo, Charles Blundell, Hanna Wallach*.

Christoph Hellmayr Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Duke University, USA
► *Flexible Functional Clustering Using Dirichlet Processes*
▷ Clustering as a technique is regularly employed in a variety of contexts, most often in exploratory data analysis. It's usefulness can be limited by the lack of an underlying statistical model. Clustering using a statistical model can be achieved by employing Dirichlet process mixtures over the quantities to be clustered. In a Bayesian setting using conjugate priors a Gibbs sampler can be built that results in a very natural idea of clustering – a latent class structure is built into the sampler that assigns each observation to a realization of the Dirichlet process mixture at each iteration of the sampler. We consider clustering multivariate function data, i.e., model based clustering where each observation is a collection of surfaces. We model multivariate functional data by using an appropriate multivariate Gaussian process as a basis function for the Dirichlet process. This framework enables prediction of new observations and interpolation at unobserved points in the underlying input space. We explore the usefulness and problems that arise in this method through a number of simulation studies. We also use it on two real data sets, comparing it to other methods for multivariate functional clustering.
⊕ Joint work with *Alan E Gelfand*.

Nhat Ho Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Department of Statistics, University of Michigan, Ann Arbor, United States
► *Posterior concentration of mixing parameters in some weakly identifiable finite mixture models*

▷ We establish posterior concentration rates and MLE rates of mixing parameters for mean-covariance Gaussian mixtures, shape rate Gamma mixtures, and some variants of finite mixture models, including the overfitted setting where the number of mixing components is bounded but unknown. These mixture models belong to what we call weakly identifiable classes, which exhibit specific interactions among mixing parameters driven by the algebraic structures of the class of kernel densities and their partial derivatives. Accordingly the optimal convergence rates of parameters in these models, as stated in terms of optimal transport distance metrics on the space of mixing measures, are shown to be typically much slower than the usual $n^{-1/2}$ or $n^{-1/4}$ rates of convergence.
⊕ Joint work with *XuanLong Nguyen*.

Chris Holmes Invited 8, Thursday, 6/25/15, 2:30 pm - 3:00 pm
University of Oxford,
► *Scalable Bayesian nonparametric regression for large data applications*
▷ Applications of regression analysis involving 1,000,000's of data points are increasing across the physical and life sciences. On the one hand the availability of large numbers of observations should assist the statistical analyst in relaxing traditional assumptions, such as non-Gaussianity of the response distribution. On the other hand such data dimensions can raise problems of scalability of computation and inference. In this talk we present a Bayesian nonparametric regression modelling approach designed for large data applications. The model is specified via a factorisation involving a nonparametric prior to characterise the marginal distributions of the response Y and covariates X, and then a regression function that captures the stochastic ordering of the predictive distribution of Y conditional on X using a Plackett-Luce ordering model. The approach supports a composite likelihood approximation which decouples the regression model from the estimation of the marginal densities allowing for standard software to be applied. The method is demonstrated on a data set of over 1,300,000 data observations and 100 predictor variables.

Masaaki Imaizumi Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
University of Tokyo, Japan
► *Bayesian estimation for nonparametric regression with low-rank tensor data*
▷ We consider a nonparametric regression model, which treats covariates as multidimensional arrays (tensors). A tensor-variate regression problem appears in many application fields such as object recognition and brain data analysis. Because the data with tensor form requires numerous cost to analyze it, low-rank approximation of the tensor structure is often used when analyzing the tensor data. In addition, a linear regression model is common to handle the tensor data and its low-rank structure. In this paper, we propose a regression model constituted by nonparametric local functions. The form is based on low-rank decomposition of tensor data, thus it can capture an effect from each components without a parametric function form. We also provide Bayesian nonparametric estimation method for the model with the Gaussian process prior. We provide theoretical validity of the estimation method, and how the nonparametric model can provide better inference. We investigate the performance of our method by both synthetic and real data.
⊕ Joint work with *Kohei Hayashi*.

James Johndrow Contributed 7A, Friday, 6/26/15, 11:30 am - 11:50 am
Duke University, USA
► *Bayesian nonparametric methods for contingency tables*

▷ Latent structure analysis is a common alternative to classical log-linear models for contingency table analysis. Latent structure models lead to a reduced rank tensor factorization of the probability mass function for multivariate categorical data. We present a number of results relating the rank of a probability tensor to sparsity in the equivalent log-linear model. A novel class of tensor decompositions is proposed and a Bayesian approach to inference in this class of models is described and applied to testing for marginal and conditional independence relationships between groups of categorical variables. We also present asymptotic rate and distribution theory results for the recovery of log-linear model parameters from the posterior under a Bayesian tensor decomposition prior. Identification, uniqueness, and estimation of the nonnegative rank of probability tensors is considered, and conditions for rank identifiability are provided in the case of PARAFAC and Tucker decompositions. We provide practical implications of these results for estimation of the number of ancestral populations of species using pure and mixed membership models of genetic variation commonly employed in population genetics.
⊕ Joint work with *David Dunson, Anirban Bhattacharya, Jeff Miller*.

Gwangsu Kim Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
Korea university, Korea
▶ *Bayesian analysis of Cox regression with a time-varying coefficient and its application*
▷ Bayesian analysis of Cox regression with a time-varying coefficient was studied theoretically. Posterior consistency for a time-varying coefficient and Bayes factor consistency were shown by the B-spline expansion and appropriate priors. In Bayes factor consistency, we consider the test of parametric against non-parametric for a coefficient, and theoretical results imply that using only partial likelihood is sufficient to obtain these consistency. Based on theoretical results, test for the equivalence of two groups hazards functions was proposed. It worked well in simulations studies for various hazards ratios including crossing hazards.

Alisa Kirichenko Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
KdVI, University of Amsterdam, Netherlands
▶ *Estimating a smoothly varying function on a large graph*
▷ We propose a nonparametric Bayesian procedure for estimating a smooth function on an expanding graph. In particular, we investigate how the convergence rates of such procedures depend on the smoothness of the function and the geometry of the graph. Here both notions of ”geometry” and ”smoothness” are quantified using the Laplacian of the graph. We prove that using a rescaled Gaussian prior we can obtain an estimator that adapts to the degree of smoothness of the unknown function. Finally, we discuss the families of the graphs that satisfy our condition on the spectrum of the Laplacian.
⊕ Joint work with *Harry van Zanten*.

Athanasios Kottas Contributed 3A, Tuesday, 6/23/15, 4:20 pm - 4:40 pm
University of California, Santa Cruz, USA
▶ *Nonparametric Bayesian modeling for dynamic ordinal regression relationships*
▷ We will present a Bayesian nonparametric framework for modeling ordinal regression relationships which evolve in discrete time. The methodology builds from nonparametric mixture modeling for the joint stochastic mechanism of covariates and latent continuous responses. This approach yields flexible inference for ordinal regression functions while at the same time avoiding the computational challenges of parametric models. In particular, the nonparametric mixture

model has full Kullback-Leibler support under fixed cut-off points that relate the latent responses with the ordinal responses. A novel dependent Dirichlet process prior for time-dependent mixing distributions extends the model to the dynamic setting. The motivating application involves study of dynamically evolving natural selection surfaces in evolutionary biology. More specifically, the methodology is applied to estimate maturity, recorded on an ordinal scale, of chilipepper rockfish as a function of age and length, using data collected over 20 years along the coast of California.
⊕ Joint work with *Maria DeYoreo*.

Christopher Krut Poster I, Monday, 6/22/15, 8:00 pm - 10:00 pm
North Carolina State University, USA
► *A multivariate functional linear model with spatially varying coefficients.*
▷ Functional data is an interesting and active area of modern statistical research with a variety of important applications. The following work proposes a multivariate functional linear model for functional predictors and scalar responses. Spatially varying coefficients are incorporated and add flexibility to the model. In addition the natural hierarchical structure of the model easily accommodates noisy functional predictors and the inclusion of derivatives of functional predictors. Coefficient functions are represented using a tensor product of B-splines and Gaussian basis functions. The careful selection of prior distributions results in inference that can be performed using a Gibbs sampler. Thus, no tuning is required and we obtain a computationally efficient algorithm. The resulting model is applied to the analysis of hurricane trajectories and intensities.
⊕ Joint work with *Montserrat Fuentes, Brian Reich*.

Ju Hee Lee Contributed 1B, Monday, 6/22/15, 4:20 pm - 4:40 pm
University of California, Santa Cruz, USA
► *Bayesian Inference for Tumor Subclones Accounting for Sequencing and Structural Variants*
▷ Tumor samples are heterogeneous. They consist of different subclones that are characterized by differences in DNA nucleotide sequences and copy numbers on multiple loci. Heterogeneity can be measured through the identification of the subclonal copy number and sequence at a selected set of loci. Understanding that the accurate identification of variant allele fractions greatly depends on a precise determination of copy numbers, we develop a Bayesian feature allocation model for jointly calling subclonal copy numbers and the corresponding allele sequences for the same loci. The proposed method utilizes three random matrices, L, Z and w to represent subclonal copy numbers (L), num- bers of subclonal variant alleles (Z) and cellular fractions of subclones in samples (w), respectively. The unknown number of subclones implies a random number of columns for these matrices. We use next-generation sequencing data to estimate the subclonal structures through inference on these three matrices. Using simulation studies and a real data analysis, we demonstrate how posterior inference on the subclonal struc- ture is enhanced with the joint modeling of both structure and sequencing variants on subclonal genomes.
⊕ Joint work with *Peter Mueller, Subhajit Sengupta, Kamalakar Gulukota, Yuan Ji*.

Eunjee Lee Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
UNC at Chapel Hill, USA
► *A Bayesian Functional Linear Cox Regression Model (BFLCRM) for Predicting Time to Conversion to Alzheimer's Disease*
▷ The aim of this paper is to develop a Bayesian functional linear Cox regression model (BFLCRM) with both functional and scalar covariates. This new development is motivated by establishing the likelihood of conversion to Alzheimers disease (AD) in 346 patients with mild cognitive im-

pairment (MCI) enrolled in the Alzheimers Disease Neuroimaging Initiative 1 (ADNI1) and the optimal early markers of conversion. These 346 MCI patients were followed over 48 months, with 161 MCI participants progressing to AD at 48 months. The functional linear Cox regression model was used to establish that the conversion time to AD can be accurately predicted by functional covariates including hippocampus surface morphology and scalar covariates including brain MRI volumes, cognitive performance (ADAS-Cog), and APOE status. Posterior computation proceeds via an efficient Markov chain Monte Carlo algorithm. A simulation study is performed to evaluate the finite sample performance of BFLCRM.
⊕ Joint work with *Hongtu Zhu, Dehan Kong, Yalin Wang, Kelly Sullivan Giovanello, Joseph Ibrahim*.

Fabrizio Leisen Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
University of Kent, United Kingdom
► *A pseudo-marginal algorithm for MGF-like distributions*
▷ Pseudo-marginal Metropolis (Andrieu and Roberts, 2009) allows to sample from a target distribution which can't be pointwise evaluated. The minimum requirement is the availability of a non-negative unbiased estimator of the target distribution. In this work, we propose a pseudo-marginal algorithm to sample from target distributions which have a form similar to a Moment Generating Function. For instance, the slice sampling for normalized random measures (Griffin and Walker (2011), involves full conditional distributions of this type. Illustrations of the method are provided with a particular focus on Bayesian non-parametrics. This is a joint work with Jim Griffin.
⊕ Joint work with *Jim Griffin*.

Hanning Li Contributed 4B, Wednesday, 6/24/15, 4:00 pm - 4:20 pm
Departent of Statistics, Florida State University, USA
► *Bayesian Variable Selection Using Continuous Shrinkage Priors*
▷ In the Bayesian paradigm, sparsity is routinely induced through two-component mixture priors having a probability mass at zero, but such priors encounter daunting computational problems in high dimensions. This has motivated continuous shrinkage priors, which can be expressed as global-local scale mixtures of Gaussians, facilitating computation. While such priors are widely used for estimating high-dimensional sparse vectors, selecting a subset of variables for linear prediction using continuous shrinkage remains an area of active research. In this article, we propose a default method for variable selection using continuous shrinkage priors. The absence of any tuning parameters make our method attractive in comparison to adhoc thresholding approaches. Theoretical properties of the proposed approach are investigated and the method is shown to have good performance in a wide range of synthetic data and real data examples.
⊕ Joint work with *Debdeep Pati*.

Meng Li Contributed 5B, Thursday, 6/25/15, 9:00 am - 9:20 am
North Carolina State University, USA
► *Bayesian Methods for Detecting Boundaries of Images*
▷ Detecting boundary of an image based on noisy observations is a fundamental problem of image processing and image segmentation. For a d-dimensional image ($d = 2, 3, \ldots$), the boundary can often be described by a closed smooth $(d-1)$-dimensional manifold. Typically, it is difficult to represent the boundary parametrically. In this talk, we propose a nonparametric Bayesian approach by a Gaussian process indexed by $\mathbb{S}^{d-1}$, the $(d-1)$-dimensional sphere. We describe efficient Markov Chain Monte Carlo methods for computing the posterior distribution of the boundary. As

the image resolution n increases, we show that the posterior of the boundary concentrates around the true boundary at the minimax optimal rate $n^{-\alpha/(\alpha+d-1)}$ up to a logarithm factor, adaptively in the smoothness level α of the boundary. We discuss particular examples of priors constructed by a rescaled squared exponential process (for any d), or finite random series generated by the Fourier bases (for $d = 2$) and spherical harmonics (for $d = 3$).
⊕ Joint work with *Subhashis Ghosal and Aad van der Vaart*.

Linxi Liu Contributed 6B, Thursday, 6/25/15, 4:00 pm - 4:20 pm
Stanford University, USA
► *Posterior concentration rate of a class of multivariate density estimators based on adaptive partitioning*
▷ Density estimation is a fundamental problem in Statistics. In this paper, we examine a class of non-parametric density estimators. These estimators are piecewise constant functions on binary partitions. We construct a prior on such functions which results in feasible algorithms for the inference of the posterior distribution, and calculate the posterior concen- tration rate under this prior. We demonstrate that parametric rate can be achieved by this class of density estimators, and the rate does not directly depend on the dimension of the problem.
⊕ Joint work with *Wing Hung Wong*.

Maria Lomeli-Garcia Contributed 1A, Monday, 6/22/15, 4:40 pm - 5:00 pm
Gatsby Unit, University College London, United Kingdom
► *A marginal sampler for σ-Stable Poisson-Kingman mixture models*
▷ Infinite mixture models reposed on random probability measures like the Dirichlet process allow for flexible modelling of densities and for clustering applications where the number of clusters is not fixed a priori. This is due to the fact that we can formulate the problem as a hierarchical model where the top level is a discrete random probability measure. In recent years, there has been a growing interest in using different random probability measures, beyond the classical Dirichlet process, for extending modelling flexibility. Some examples include Pitman-Yor processes, normalised generalised Gamma processes, and normalized random measures. Our understanding of these models has grown significantly over the last decade: there is an increasing realisation that while these models are nonparametric in nature and allow an arbitrary number of components to be used, they do impose significant prior assumptions regarding the clustering structure. In this talk we will present a wide class of random probability measures, called σ-Stable Poisson-Kingman processes, and discuss its use for Bayesian nonparametric mixture modelling. This class of processes encompasses most known random probability measures proposed in the literature so far and we argue that it forms a natural class to study. We will review certain characterisations which lead us to propose a tractable and exact posterior inference algorithm for the whole class. Specifically, we are able to derive a marginal sampler in an augmented space that has a fixed number of auxiliary variables per iteration. We illustrate the algorithm's performance with a multidimensional experiment and compare it against a conditional sampler.
⊕ Joint work with *Stefano Favaro and Yee Whye Teh*.

Li Ma Invited 3, Tuesday, 6/23/15, 10:30 am - 11:00 am
Duke University, U.S.A
► *Probabilistic multi-resolution scanning for cross-sample differences*
▷ We propose a multi-resolution scanning approach to identifying two-sample differences. Windows of multiple scales are constructed through nested dyadic partitioning on the sample space and a local hypothesis regarding the two-sample difference is defined on each window. Instead of

testing the local hypotheses independently, we adopt a joint graphical model on the null/alternative states of these hypotheses to incorporate spatial correlation across windows. This dependency leverages on the spatial clustering of two-sample differences to borrow information across nearby and nested windows, which we show is critical for detecting high-resolution, local differences. We evaluate the performance of the method through simulation and show that it substantially outperforms other state-of-the-art two-sample tests when the two-sample differences is local, involving only a small subset of the data. We then apply it to a flow cytometry data set from immunology, in which it successfully identifies highly local differences. In addition, we show how to properly control for multiple testing in a decision theoretic approach as well as how to summarize and report the inferred two-sample difference by inferring a representative partition of the sample space.

Fabian Martinez Contributed 7B, Friday, 6/26/15, 11:10 am - 11:30 am
Universidad Nacional Autónoma de México, Mexico
► *Change-Point Detection on Dependent Data, a Random-Partition Approach*
▷ Change-point detection models aim to find abrupt changes in a given sample indexed on an ordered set. We tackle this problem under a clustering approach restricted by such indexing set. Our proposed methodology is based on exchangeable partition probability functions, specifically on Pitman's sampling formula, also known as two-parameter Poisson-Dirichlet process. Moreover, and due to the model-based nature of the problem, emphasis will be given to the case where data inside each cluster is modelled by a Markovian process, in particular we will concentrate on discretely observed Ornstein-Uhlenbeck diffusion processes. Some properties of the resulting model are explained and posterior results are obtained via a novel Markov chain Monte Carlo algorithm.
⊕ Joint work with *Ramses Mena*.

Robert McCulloch Contributed 6A, Thursday, 6/25/15, 5:00 pm - 5:20 pm
U of Chicago., USA
► *Nonparametric Heteroscedastic Regression Modeling, Bayesian Regression Trees and MCMC Sampling*
▷ Bayesian additive regression trees (BART) have become increasingly popular as flexible and scalable non-parametric models useful in many modern applied statistics regression problems. They bring many advantages to the practitioner dealing with large datasets and complex nonlinear response surfaces, such as the matrix-free formulation and the lack of a requirement to specify a regression basis a priori. However, there are some known challenges to this modeling approach, such as poor mixing of the MCMC sampler and inappropriate uncertainty intervals when the assumed homoscedastic variance model is violated. In this talk, we introduce a new Bayesian regression tree model that allows for possible heteroscedasticity in the variance model and devise novel MCMC samplers that appear to adequately explore the posterior tree space of this model.
⊕ Joint work with *Matt Pratola Hugh Chipman Ed George*.

Ramses Mena Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
IIMAS UNAM, Mexico
► *Ruin Probabilities for Bayesian Exchangeable Claims Processes*
▷ Among the driving assumptions in classical collective risk models, the independence among claims is frequently violated by real applications. Therefore, there is an evident need of models that relax such a restriction. We undertake the exchangeable claims platform and obtain some results for the infinite time ruin probability. The main result is that the ruin probability under the

exchangeable claims model can be represented as the expected value of the ruin probabilities corresponding to certain independent claims cases. This allows us to extend some classical results to this dependent claims scenario. The main tool is based on the de Finettis representation theorem for exchangeable random variables, and as a consequence a natural Bayesian modeling feature for risk processes becomes available. In particular, an interesting redefinition of the net profit condition is necessary.
⊕ Joint work with *Arrigo Coen*.

David Kwamena Mensah Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Department of Statistics and Applied Probability, National University of Singapore, Singapore
► *Functional Longitudinal modelling with covariate dependent smoothness*
▷ Recent technological advances have made the collection of functional data common in a range of scientific applications and this has engineered a growing interest in functional data analysis within the statistical community. We consider functional longitudinal models with subject and group specific trends modelled using Gaussian processes. Fitting Gaussian process regression models is well known to be a computationally challenging task, and various sparse approximations to Gaussian processes have been considered in the literature to ease the computational burden. We focus on Bayesian methods, and in particular, we build on a fast non-standard variational approximation based on a sparse spectral representation and which is able to treat uncertainty in the covariance function hyperparameters. This allows fast variational computational methods to be extended to complex models where there are many functions to be estimated and where there exist a hierarchical model involving the covariance function parameters. This idea is implemented in the context of functional longitudinal models by allowing individual specific smoothness related to covariates for different subjects. Understanding the relationship of smoothness to individual specific covariates is of great interest in some applications. The methods are illustrated with simulated data and a dataset of streamflow curves generated by a rainfall runoff model.
⊕ Joint work with *David Nott*, *Linda Tan*, *Lucy Mashall*.

Jeff Miller Contributed 2A, Tuesday, 6/23/15, 9:20 am - 9:40 am
Duke University, USA
► *An approach to inference under misspecification*
▷ Due to small departures from the assumed model, the posterior of a flexible Bayesian model such as a nonparametric mixture sometimes concentrates on overly complex submodels, particularly on large datasets. This can significantly reduce the interpretability of the results, and further, can greatly increase the computational burden. We explore an approach to coherent Bayesian inference allowing for minor misspecification, based on conditioning on a neighborhood of the data, rather than conditioning on the data exactly. We provide theoretical results and empirical demonstrations with real and simulated data.
⊕ Joint work with *David Dunson*.

Peter Müller Invited 8, Thursday, 6/25/15, 2:00 pm - 2:30 pm
UT Austin, United States
► *BNP Inference for Dynamic Treatment Regimes*
▷ We develop a BNP inference approach for dynamic treatment regimes that include a sequence of treatments and transition times between disease states. A common example arises in cancer therapies where the transition times between various sequences of treatments, disease remission, disease progression, and death characterize overall survival time. For the general setting, we

propose estimating mean overall outcome time by assuming a Bayesian nonparametric survival regression for each of the transition times. We use a simple dependent Dirichlet process prior. The assumed model allows us to infer treatment effects as differences between expected outcomes under alternative treatments. This is important to address the lack of randomization in some of the treatment assignments. We compare the proposed approach with two commonly used frequentist methods, parametric likelihood-based G-estimation and inverse probability of treatment weighting. These comparisons are done by simulation studies of both single-stage and multi-stage regimes, with treatment assignment depending on baseline covariates. The nonparametric method compares favorably, especially when both models, response model and treatment assignment, are mis-specified
⊕ Joint work with *Yanxun Xu, Peter Thall, and Abdus Wahed*.

Jared Murray Contributed 6A, Thursday, 6/25/15, 4:40 pm - 5:00 pm
Carnegie Mellon University, USA
► *Loglinear Bayesian Additive Regression Trees for Classification, Counts, and Heteroscedastic Regression*
▷ Bayesian additive regression trees (BART) have been applied to nonparametric mean regression and classification problems in a range of applications. To date BART has been limited to models for Gaussian "data", observed or latent, and with good reason - the Bayesian backfitting MCMC algorithm for BART is remarkably efficient in conditionally Gaussian models. But while many useful models are naturally cast in terms of latent Gaussian variables, many others are not. I extend BART to loglinear models for multinomial logistic regression, Poisson and negative binomial regression, and Gaussian mean-variance regression. To accomplish this I introduce a novel prior distribution over BART's leaf parameters, as the original Gaussian leaf priors are no longer tractable in these models. Like the original BART leaf parameter prior, this new prior is carefully constructed and calibrated to induce a sensible prior over the regression function. With this new prior distribution and some judicious data augmentation I am able to implement an efficient generalized Bayesian backfitting algorithm for MCMC in these new models. I illustrate with synthetic and real datasets.

Thomas Murray Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
The University of Texas MD Anderson Cancer Center, United States
► *Combining Functional or Survival Data Sources*
▷ Conventional approaches to statistical inference preclude structures that facilitate incorporation of supplemental information acquired from similar circumstances. This work develops a hierarchical model structure that leverages all available information about a curve, using penalized splines, while accommodating important between-source features. Our proposed methods flexibly borrow strength from the supplemental data to a degree that reflects the commensurability of the supplemental curve with the primary curve. We apply our method in two contexts, functional data analysis and survival analysis. We analyze functional data obtained using perfusion computed tomography to characterize imaging biomarkers in cancerous regions of the liver that may benefit from partially informative data collected concurrently in non-cancerous regions. We also analyze survival data from a clinical trial that compares time to disease progression for three colorectal cancer treatments and may benefit from information collected in a previous trial on the current standard of care.
⊕ Joint work with *Brian P. Hobbs, and Bradley P. Carlin*.

Zacharie Naulet Contributed 6A, Thursday, 6/25/15, 4:20 pm - 4:40 pm

CEA, LIST, Laboratory of Modeling, Simulation and Systems, France

► *Adaptive Bayesian nonparametric regression using mixtures of kernels*

▷ Recently, interest in a Bayesian nonparametric approach to the sparse regression problem emerged after initial works from Abramovich et al. (2000) and Wolpert et al. (2011). The underlying probabilistic model is analogous to infinite dimensional density mixture models and parsimony is intrinsically induced by the choice of the mixing measure. In contrast with density estimation, mixture components must be replaced by suitable kernels function, chosen so that they span the function space of interest. We believe that kernels arising from representations of topological groups are good candidates (see Ali et al., 2000), because they are overcomplete dictionaries for large families of Banach spaces, which is often advantageous to get sparse representations of functions. Famous examples treated include continuous wavelets, and Gabor atoms. We should insist on the fact that the set of such kernels is not countable and must not be confused with discrete wavelets or Gabor basis. Modeling the regression function as a mixture of kernels also requires the use of generalized mixing measures, taking reals or complex values. We propose a simple construction based on Completely Random Measures (Kingman, 1967), allowing us to make simple analysis of prior distribution and to get posterior consistency results in gaussian regression setting. Due to similarities with density mixture models, efficient sampling schemes can be proposed, requiring some adaptions because of the lack of connection between observations and atoms of the mixing measure. For a particular class of generalized mixing measures we propose a Gibbs sampler based on algorithm from Neal (2000). It is worth noting that the algorithm allows to sample the full posterior distribution, permitting the estimation of credible bands. As an application, the algorithm is compared to classical wavelets thresholding methods in gaussian regression on the collection of test functions from Marron et al. (1998) and methodology from Antoniadis et al. (2001). We also consider the problem of Quantum Homodyne Tomography (Artiles et al., 2005), a non linear inverse problem for which no Bayesian approach has been proposed so far.

⊕ Joint work with *Eric Barat*, *Judith Rousseau*, *Trong Tuong Truong*.

Long Nguyen Contributed 2A, Tuesday, 6/23/15, 9:40 am - 10:00 am

University of Michigan, U.S.A

► *Identifiability and posterior concentration of matrix-variate parameters in mixture models*

▷ This paper studies identifiability and posterior concentration behavior of matrix-variate parameters and multiple parameter types that arise in finite mixtures, including the setting of overfitted mixture models. We revisit and generalize several notions of strong identifiability, and use them to establish sharp inequalities relating the distance of mixture densities and the Wasserstein distances in various orders of the corresponding mixing measures. Characterization of identifiability is given for a broad range of matrix-variate mixture models commonly employed in practice, including location-covariance mixtures and location-scale-covariance mixtures, mixtures of symmetric densities, as well as some asymmetric ones. We also show that in the case of overfitted mixture models, the convergence behavior depends in a crucial way on the nature of the kernel densities. In particular, for finite location-covariance Gaussian mixtures, the posterior concentration rate of the mixing measure on the locations and covariances can get arbitrarily worse (in the polynomial scale of $(1/n)$, where n is the sample size) as the number of extra mixing components is allowed to increase. Identifiability and concentration behaviors of the mixing measure are illustrated by a simulation study on a broad range of multivariate and matrix-variate kernel densities.

⊕ Joint work with *Nhat Ho*.

Luis Nieto-Barajas Invited 8, Thursday, 6/25/15, 3:00 pm - 3:30 pm

Department of Statistics, ITAM, Mexico

▶ *Markov constructions in Biostatistics applications*

▷ In this talk we review some discrete and continuous time Markov constuctions that are useful to build Bayesian nonparametric models. We illustrate the use of the models in different contexts: Survival data analysis, time series of proteins expression, DNA copy number comparison and disease mapping.

Bernardo Nipoti Invited 4, Tuesday, 6/23/15, 2:30 pm - 3:00 pm

University of Torino and Collegio Carlo Alberto, Italy

▶ *Modeling the association structure of clustered time-to-event data*

▷ A standard approach for dealing with correlated time-to-event data within the proportional hazards (PH) context has been the introduction of a random effect (frailty) that is common to subjects within the same cluster. PH models with shared random effects have been widely discussed because they provide useful summary information in the absence of estimates of a baseline survival distribution and may be formulated in a semiparametric fashion. However, when it is assumed that the frailty variables are independent and identically distributed, they induce a particular marginal association structure for the clustered variables, implying equal intra-cluster dependence as well as between-cluster heterogeneity. We propose an alternative Bayesian semiparametric model that naturally accommodates for different degrees of association, allowing for covariate-dependent association structures within each cluster. The proposal is based on the introduction of cluster-dependent random hazard functions and on the use of mixture models induced by conditionally independent σ-stable completely random measures. We show that the proposed model class has the appealing property of preserving marginally the PH structure. Further borrowing of strength between clusters can be favored by introducing dependence across cluster-specific random hazards. The model is illustrated by means of an application to a rich last survivor life insurance policy dataset.

⊕ Joint work with *Alejandro Jara, Michele Guindani*.

Andriy Norets Contributed 3B, Tuesday, 6/23/15, 4:20 pm - 4:40 pm

Brown University, US

▶ *Adaptive Bayesian Estimation of Conditional Densities*

▷ We consider a non-parametric Bayesian model for conditional densities. The model is a finite mixture of normal distributions with covariate dependent multinomial logit mixing probabilities. A prior for the number of mixture components is specified on positive integers. The marginal distribution of covariates is not modeled. We study asymptotic frequentist behavior of the posterior in this model. Specifically, we show that when the true conditional density has a certain smoothness level, then the posterior contraction rate around the truth is equal up to a log factor to the frequentist minimax rate of estimation. As our result holds without a priori knowledge of the smoothness level of the true density, the established posterior contraction rates are adaptive. Moreover, we show that the rate is not affected by inclusion of irrelevant covariates in the model.

⊕ Joint work with *Debdeep Pati*.

Peter Orbanz Invited 2, Monday, 6/22/15, 2:30 pm - 3:00 pm

Columbia University,

▶ *Subsampling large graphs and invariance in networks*

▷ Consider a very large graph—say, the link graph of a large social network. Now specify a randomized algorithm that extracts a smaller subgraph as a sample. What can we learn from a statistical analysis of this sample? Clearly, that should depend on the subsampling algorithm. I show how the choice of algorithm defines a notion of distributional invariance. This invariance is precisely exchangeability for a certain choice of algorithm, but it changes drastically when seemingly minor modifications are made to the subsampler; I will give a concrete example. Thus, the invariance properties a statistical model should satisfy can be regarded as a consequence of a of the data acquisition method. Under suitable conditions, the invariant distributions defined by an algorithm automatically admit a de Finetti-style representation theorem. From this algorithmic perspective, various known pathologies that make exchangeable graph models problematic for network analysis can be explained as a selection bias.
⊕ Joint work with *NA*.

Ana Maria Ortega Villa Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Student, United States of America
► *Semiparametric Spatio-Temporal Varying Coefficient Model for Matched Case-Crossover Studies*
▷ In matched case-crossover studies, it is generally accepted that the covariates on which a case and associated controls are matched cannot exert a confounding effect on independent predictors included in the conditional logistic regression model. However, some matching covariates such as time often play an important role as an effect modification which makes incorrect statistical estimation and prediction. Furthermore there might also be an effect due to the different spatial locations among the subjects. Hence, we propose two semiparametric, spatio-temporal varying coefficient models to evaluate effect modification by time and spatial location in order to make correct statistical inference. Our approaches allow us to evaluate parametric relationships between the predictor and binary response, semiparametric relationships between the predictor and time as well as effect modifications due to spatial location.
⊕ Joint work with *Inyoung Kim*.

Garritt Page Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Ponticia Universidad Catolica de Chile, Chile
► *Spatial Product Partition Models*
▷ When modeling geostatistical or areal data, spatial structure is commonly accommodated via a covariance function for the former and a neighborhood structure for the latter. In both cases the resulting spatial structure is a consequence of implicit spatial grouping in that observations near in space tend to have similar values. It would be desirable to develop spatial methods that explicitly model the partitioning of spatial locations providing more control over resulting spatial structures and being able to better balance global vs. local spatial dependence. To this end, we extend product partition models to a spatial setting so that the partitioning of locations into spatially dependent clusters is explicitly modeled. ⊕ Joint work with *Fernando A. Quintana*.

Debdeep Pati Contributed 5A, Thursday, 6/25/15, 9:20 am - 9:40 am
Florida State University, United States
► *Bayesian nonparametric graphon estimation*
▷ Data available in the form of networks are increasingly becoming common in modern applications ranging from brain remote activity, protein interactions, web applications, social networks to name a few. Estimating large networks and the underlying function (graphon) generating the network calls for structured dimension reduction and estimation in stylized domains, necessitating

new tools for model based inference and theory. In this talk, we focus on Bayesian nonparametric graphon estimation via a transformed tensor product of Bernstein polynomials. We also establish minimax optimal rate of posterior convergence in estimating the network and the graphon. An efficient Gibbs sampler is developed and the proposed graphon estimation approach is illustrated via simulated examples and comparing real social networks.
⊕ Joint work with *Anirban Bhattacharya*.

Sonia Petrone Invited 2, Monday, 6/22/15, 3:00 pm - 3:30 pm
Bocconi University, Milano, Italy
► *Bayesian nonparametrics in bivariate evolutionary phenomena*
▷ Many basic tools for Bayesian nonparametrics are developed in the framework of exchangeability. Exchangeability refers to an evolutionary process. Think of the Chinese Restaurant process, where customers enter sequentially. It is the representation theorem that guarantees that an exchangeable evolutionary phenomena is probabilistically equivalent to a static sampling where first, the population distribution is drawn at random from a prior law, then observations are sampled independently. However, in some situations the assumption of exchangeability is restrictive, and the evolutionary nature of the phenomena cannot be ignored. The phenomena we have in mind for this talk are such that a weight should be associated to each observation: in the CRP, the capacity of attracting new customers at a table depends not only on the number of customers already sitting there, but also on their importance, say. Problems of this nature include, for example, competition in evolving networks and evolutionary linguistic. The need of allowing for weighted versions of BNP procedures underlines recent developments in species sampling and extensions of the Indian Buffet Process. Recently, Berti, Pratelli and Rigo (Annals of Probability, 2004) introduced a notion of conditionally identically distributed sequences that extends exchangeability and can describe many situations like the above. In this talk, we propose a notion of partially identically distributed (p-cid) sequences, that extends partial exchangeability. We show main theoretical properties of p-cid sequences and illustrate how they can describe dependent evolutionary phenomena. Examples include weighted generalizations of the bivariate Dirichlet process and of the hierarchical Dirichlet process.
⊕ Joint work with *Sandra Fortini*, *Polina Sporysheva*.

Igor Prünster Keynote III, Friday, 6/26/15, 9:00 am - 10:00 am
University of Torino & Collegio Carlo Alberto, Italy
► *Dependent Random Measures and Prediction*
▷ Within the framework of partial exchangeability, we consider models based on dependent completely random measures (or suitable transformations thereof). Some of their distributional properties are presented with focus on additive, hierarchical and nested constructions. This is done in de Finetti's spirit by adopting a predictive viewpoint, which allows to highlight some key features of the models and, more generally, of partial exchangeability itself.
⊕ Joint work with *Federico Camerlenghi, Antonio Lijoi, and Bernardo Nipoti*.

Gavino Puggioni Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
University of Rhode Island, United States
► *A Bayesian Nonparametric Approach for Spatiotemporal Point Processes*
▷ We propose a nonparametric method to estimate the intensity of a point process observed in space and time. The modeling procedure, treated as a dynamic density estimation problem, involves the specification of a prior based on a Dirichlet process mixture of bivariate normal distri-

butions at each point in time. Temporal dependence is introduced through the atoms that evolve according to a dynamic linear model structure. Comparison with existing methods and an application to wildlife infectious disease spread are provided to illustrate the methodology.
⊕ Joint work with *Lance Waller, Luca Gerardo-Giorda, Leslie Real*.

Fernando Quintana Invited 3, Tuesday, 6/23/15, 11:30 am - 12 Noon

Pontificia Universidad Católica de Chile, Chile

► *Predictions Based on the Clustering of Heterogeneous Functions via Shape and Subject-Specific Covariates*

▷ Motivated by a study of players employed by teams in the National Basketball Association, we consider data structures where units of observation are functional curves. We are especially interested in a case where the observed functional output displays large amounts of between-subject heterogeneity in the sense that some individuals produce curves that are fairly smooth while others are (much) more erratic. We argue that this variability in curve shape is a feature that can be exploited to guide decision-making, learn about processes under study and improve prediction. We develop a methodology that takes advantage of this feature when clustering functional curves. Individual curves are flexibly modeled, while a hierarchical structure allows the clustering to be guided by the smoothness of individual curves. In a sense, the hierarchical structure balances the desire to fit individual curves well while still producing meaningful clusters that are used to guide prediction. We seamlessly incorporate available covariate information to guide the clustering of curves non-parametrically through the use of a product partition model prior for a random partition of individuals. Clustering based on curve smoothness and subject-specific covariate information is particularly important in carrying out the two types of predictions that are of interest for the motivating data, namely, those that complete a partially observed curve from an active player, and those that predict the entire career curve for a player yet to play in the NBA.
⊕ Joint work with *Garritt Page*.

Rajesh Ranganath Contributed 1B, Monday, 6/22/15, 5:00 pm - 5:20 pm

Princeton University, USA

► *Correlated Random Measures*

▷ Many hierarchical Bayesian nonparametric models are built from completely random measures, in which atom weights are independent. This leads to implicit independence assumptions in the corresponding hierarchical model, assumptions that are often misplaced in real-world settings. In this work, we address this limitation. We develop correlated random measures, a class of random measures where the measures on two disjoint sets can exhibit both positive and negative dependence. Correlated random measures model correlation within the measure by using a Gaussian process in concert with a Poisson process. With this construction, for example, we can develop a latent feature model for which we can infer both the properties of the latent features and their correlation. We develop several examples of correlated random measures. We study a hierarchical model of pairwise count data constructed from correlated random measures and derive a variational inference algorithm to approximate the corresponding posterior. We show improved predictive performance on large collections of text and large collections of medical diagnostic codes.
⊕ Joint work with *David Blei*.

Kolyan Ray Invited 3, Tuesday, 6/23/15, 11:00 am - 11:30 am

Leiden University, Netherlands

► *Bernstein-von Mises theorems for adaptive Bayesian nonparametric procedures*

▷ A key aspect of statistical inference is uncertainty quantification and the Bayesian approach to this problem is to use the posterior distribution to generate a credible set, that is a region of prescribed posterior probability (often 95

Yordan Raykov Contributed 1A, Monday, 6/22/15, 4:20 pm - 4:40 pm
Aston University, United Kingdom

► *Fast Approximate MAP Inference for Bayesian Non-Parametrics*

▷ Full probabilistic inference for Bayesian non-parametric (BNP) models is in most cases analytically intractable and computationally intensive MCMC techniques are required. As a result, BNP-based methods, which have considerable potential, are restricted to applications in which computational resources and time for inference is plentiful. For example, they would not be practical for digital signal processing on embedded hardware, where computational resources are at a serious premium. We present a simplified yet statistically rigorous approximate maximum a-posteriori inference algorithm and demonstrate its applicability on the the Dirichlet process mixture model and its hierarchical extension, the HDP, and the Pitman-Yor process. This class of algorithms is as simple as K-means clustering, performs in experiments as well as Gibb's sampling, while requiring only a fraction of the computational effort. Unlike related small variance asymptotics, this class of algorithms is non-degenerate and so inherits the 'rich get richer' property of the underlying stochastic process. It also retains a non-degenerate closed-form likelihood which enables standard tools such as cross-validation to be used. The generality and scalability of the approach is demonstrated on a large longitudinal dataset using a semi-parametric random effects model.

⊕ Joint work with *Alexis Boukouvalas*, *Max Little*.

Pedro Regueiro Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
University of California, Santa Cruz, USA

► *Stochastic variational Bayes in the stochastic blockmodel*

▷ The class of Bayesian stochastic blockmodels has become a popular approach to relational data. This is due, in part, to the fact that inference on structural properties of networks follows naturally in this framework. We use a latent variable augmentation scheme to develop a Markov chain Monte Carlo algorithm that allows to fit this model. However, Markov Chain Monte Carlo algorithms for stochastic blockmodels do not scale well as the network grows. Here, we develop a stochastic variational algorithm for the stochastic blockmodel. Illustrations are provided both through simulated and real datasets.

⊕ Joint work with *Abel Rodriguez*.

Boyu Ren Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Harvard T. H. Chan School of Public Health, United States

► *Factor analysis of multi-population species sampling sequences*

▷ For species sampling sequences, the latent factors that are associated with the underlying probability measure over the species space are of great interest. In practice, we usually have only limited prior knowledge about such factors and an unsupervised learning method is then desired to discover the candidates from the observed data. We present here a Bayesian nonparametric factor analysis framework to address such problem with high flexibility. Through a new construction of Dirichlet Process prior, we can make a one-to-one correspondence between the correlation structure of the underlying multinomial probabilities among populations to the covariance matrix of a multivariate normal component in the model. By doing factor analysis for the covariance matrix, we can uncover the latent factors that are associated with variability of the population specific probability measures defined in the species space. The assumed construction possesses a

log-concave posterior density, which makes the sampling easy to carry out. We perform simulation study to verify the model can recover the predefined latent factors. A few real data examples in Microbiome data are also presented to show the usefulness of this model in discovering underlying factors that contribute to the variability of human microbiome profile over time and body sites.
⊕ Joint work with *Sergio Bacallado, Stefano Favaro, Lorenzo Trippa*.

Robert Richardson Contributed 3A, Tuesday, 6/23/15, 5:00 pm - 5:20 pm
University of California - Santa Cruz, USA
► *Bayesian Non-parametric Modeling for Integro-Differential Equations*
▷ Integro-Differential Equations (IDEs) are a novel way of dynamically modeling spatio-temporal data. The choice of kernel in an IDE spatio-temporal model relates directly to the underlying physical process the model is assuming, and more flexible kernels will allow more flexibility in the physical process explained by the IDE model. With the wide variety of spatio-temporal data arising from environmental science, ecology, and other fields, it is important to use a model which can accommodate a wide array of physical processes. We introduce Bayesian non-parametrics to the IDE literature as a way to allow complete flexibility. We use a Dirichlet process mixture of normals to model the kernel and then expand to a spatial Dirichlet process mixture of normals. The latter allows the model to capture non-stationarity in space and reflect a changing physical process across the spatial field. We address computational concerns involved with learning this model, including using Hermite polynomials for a basis representation of the kernel to reduce the dimensionality and sampling from the posterior distributions of the spatial Dirichlet process parameters using Hamiltonian Markov Chain Monte Carlo. Simulations will show that the model can successfully recapture complicated dynamics which vary across space and then we compare the model fit against various other models in a real data example. We will show that the spatial Dirichlet process mixture of normals kernel is an excellent choice for prediction and model accuracy.
⊕ Joint work with *Athanasios Kottas, Bruno Sanso*.

Judith Rousseau Invited 7, Thursday, 6/25/15, 11:00 am - 11:30 am
Université Paris Dauphine and CREST-ENSAE, France
► *Empirical Bayes procedures - understanding the behaviour of the MMLE in nonparametric models*
▷ In this work we investigate frequentist properties of Empirical Bayes procedures. Empirical Bayes procedures are very much used in practice in more or less formalized ways as it is common practice to replace some hyperparameter in the prior by some data dependent quantity. There are typically two ways of constructing these data dependent quantities : using some king of moment estimator or some quantity whose behaviour is well understood or using a maximum marginal likelihood estimator. In this work we caracterize the behaviour of the maximum marginal likelihood estimator (MMLE) in nonparametric models, generalizing the findings fo Szabo et al. (2014) which were obtained in the white noise models with Gaussian process priors, to more general models and families of priors.
⊕ Joint work with *Botond Szabo*.

Daniel Roy Contributed 2B, Tuesday, 6/23/15, 9:00 am - 9:20 am
University of Toronto, Canada
► *The continuum-of-urns scheme, generalized beta and Indian buffet processes, and hierarchies thereof*
▷ This talk pertains to a recent preprint http://arxiv.org/abs/1501.00208 I will discuss how a

new stochastic process—the continuum-of-urns scheme—sheds new light on the key stochastic processes underlying nonparametric latent feature models that have attracted a great deal of attention since the introduction of the Indian buffet process by Griffiths and Ghahramani and the subsequent characterization of the IBP as the combinatorial structure of an exchangeable sequence of Bernoulli processes directed by a beta process. One of the key results shows how every exchangeable partition gives rise to a "generalized" beta process. Indeed, Dirichlet processes yield beta processes; Pitman–Yor processes yield stable beta process; and Gibbs-type processes yield novel Gibbs-type beta processes. Specifically, I show that the ordinary component of a generalized beta process is determined by the structural distribution of the exchangeable partition, i.e., the distribution of first size-biased pick. It is then the ordinary component that gives rise to a generalized Indian buffet process. Beyond the structural distribution, the law of the exchangeable partition gives rise to a generalized hierarchical beta process. In particular, a linear functional of the ranked masses gives rise to a probability kernel that determines the law of fixed atoms. It is this probability kernel that defines the law of a conditional generalized beta process whose mean is another generalized beta process. The continuum-of-urns scheme yields other insights, including stick-breaking constructions and finite truncations with a posteriori approximation bounds.

Matteo Ruggiero Contributed 4A, Wednesday, 6/24/15, 5:00 pm - 5:20 pm
University of Torino and Collegio Carlo Alberto, Italy

► *Filtering hidden Markov measures*

▷ We consider the problem of learning two families of time-evolving random measures from indirect observations. In the first model, the signal is a Fleming–Viot diffusion, which is reversible with respect to the law of a Dirichlet process, and the data is a sequence of random samples from the state at discrete times. In the second model, the signal is a Dawson–Watanabe diffusion, which is reversible with respect to the law of a gamma random measure, and the data is a sequence of Poisson point configurations whose intensity is given by the state at discrete times. A common methodology is developed to obtain the filtering distributions in a computable form, which is based on the projective properties of the signals and duality properties of their projections. The filtering distributions take the form of mixtures of Dirichlet processes and gamma random measures for each of the two families respectively, and an explicit algorithm is provided to compute the parameters of the mixtures. Hence, our results extend classic characterisations of the posterior distribution under Dirichlet process and gamma random measures priors to a dynamic framework.

⊕ Joint work with *Omiros Papaspiliopoulos, Dario Spanó*.

Jean-Bernard Salomond Contributed 3B, Tuesday, 6/23/15, 4:40 pm - 5:00 pm
CWI, Amsterdam, The netherlands

► *Gereral posterior contraction results for nonparametric ill posed incerse problems*

▷ We consider the problem of estimating a function f from indirect observations. More precisely, we consider the case where the available data comes from a transformation of f by a non-invertible linear operator K. In the recent years, nonparametric approach to these models, has become more and more popular in the statistical literature. Standard examples include regression models with error in variables, numerical differentiation or the computerized tomography problem. Bayesian methods for these models have received a growing interest in the recent years, but only very specific priors such as conjugate Gaussian priors have been considered so far from a theoretical point of view. In this work, we present a general approach to derive posterior contraction results for

ill posed inverse problems based on regularity property of the model. An interesting aspect of this approach is that it allows us to derive contraction rates for priors that are not related to the singular value decomposition of the operator. We recovered (a subset of) the existing results from the Bayesian literature. Our approach should be viewed as a generalization of the ideas presented in the latter paper. Furthermore, we were able to derive posterior contraction rates for prior distributions that were not covered by the existing theory. We also treated an operator that does not admit singular value decomposition. In this sense, the approach proposed in this paper is more general, and we elieve more natural, that the existing ones.
⊕ Joint work with *Bartek Knapik*.

Nilotpal Sanyal Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Indian Statistical Institute-Calcutta, India
► *Bayesian wavelet analysis with nonlocal priors*
▷ We propose a Bayesian hierarchical wavelet methodology for nonparametric regression based on nonlocal priors. Our methodology assumes for the wavelet coefficients a prior that is a mixture of a point mass at zero and a Johnson-Rossell nonlocal prior. We have considered two choices of Johnson-Rossell nonlocal priors, the moment prior and the inverse moment prior. To borrow strength across wavelet coefficients, in addition to more traditional specifications from the wavelet literature, we consider a logit specification for the mixture probability and a polynomial decay specification for the scale parameter. These specifications depend on few hyperparameters. To estimate these hyperparameters, we have developed an empirical Bayes methodology based on a Laplace approximation. In addition, we have developed Laplace approximation-based analytic posterior inference for the wavelet coefficients. We then use the inverse wavelet transform to obtain a nonparametric estimate of the function of interest. To assess performance, we have conducted a simulation study using the four Donoho-Johnstone test functions to compare our methodology to several other wavelet-based methods available in the literature. Compared with the competing methods, in terms of mean squared error our methodology with the inverse moment prior and the novel specifications for the prior hyperparameters yields comparable results for samples of size 512 and superior results for larger sample sizes.
⊕ Joint work with *Marco A. R. Ferreira*.

Terrance Savitsky Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
U.S. Bureau of Labor Statistics, USA
► *Convergence Rates of Posterior Distributions under Informative Sampling*
▷ An informative sampling design assigns probabilities of inclusion that are correlated with the response of interest and often induces a dependence among sampled observations. It is well-known that model inference performed on data acquired under an informative sampling design will be biased for estimation of the joint distribution of model parameters supposed to generate the population from which the sample was drawn. Known marginal inclusion probabilities assigned through the sampling design may be used to weight the likelihood contribution of each observed unit in the sample with the practical intent to "undo" the design for inference about the population. This paper extends a theoretical result on consistency of the posterior distribution at the true generating distribution to the weighted "pseudo" posterior distribution used to account for an informative sampling design. We construct conditions on known marginal and pairwise inclusion probabilities that define a class of sampling designs where consistency is achieved, in probability. We perform a simulation study and provide an application for data acquired under the allowable class of sampling designs that demonstrate the result.

Moritz Schauer Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

University of Amsterdam, The Netherlands

► *Poster: Nonparametric Bayesian inference for diffusion processes - an integrated approach*

▷ The poster presents an integrated approach to nonparametric Bayesian estimation of the drift function of a diffusion process. We propose a Gaussian prior derived via the midpoint-displacement procedure from a Gaussian Markov process. The prior obtained by midpoint-displacement can alternatively be seen as an expansion in the Faber-Schauder basis. The infinite series is truncated and scaled. We derive truncation and scaling hyper priors that give adaptive posterior contraction rates for diffusion with periodic drift. Numerically, the compact support of the basis elements give a sparse Gramian matrix, which can be inverted in quasi linear-time using a divide and conquer algorithm. Moreover, by the Markov-property of the prior a corresponding sparsity property for the prior precision matrix holds. Truncation and scaling are handled by a reversible jump MCMC procedure. The implementation is available as software package and can be adapted to allow for discretely observed diffusion processes with unknown diffusion coefficents.

⊕ Joint work with *Frank van der Meulen and Jan van Waaij*.

Aaron Schein Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

UMass Amherst, USA

► *Dynamic Bayesian Poisson Tensor Factorization*

▷ We propose a nonparametric Bayesian Poisson tensor factorization model for dynamic, multiway count data observed sequentially through time. Each of the latent components inferred by our model consists of a set of latent factor vectors—one for each mode of the observed tensor and with one mode being time. To model the sequential nature of time, and thus dependencies between time steps, our model imposes a Markov chain as a prior over each component's temporal factor vector. This prior specifies that the component's factor at time t is drawn from a gamma distribution whose shape parameter is equal to the component's factor at time t-1. Finally, to encourage a parsimonious representation, we place a gamma process prior over the component weights. We apply our model to a large-scale international relations data set, consisting of over a quarter of a billion pairwise interactions between countries, and show that the inferred latent components correspond to multilateral relations and events. We also apply our model to a data set of almost two million US State Department cables from 1973 to 1977. This data set provides a unique view into the diplomatic history of the US—a topic of considerable interest to both political scientists and historians.

⊕ Joint work with *Mingyuan Zhou, John Paisley, David Blei, Hanna Wallach*.

Johannes Schmidt-Hieber Invited 7, Thursday, 6/25/15, 11:30 am - 12 Noon

Leiden University, Netherlands

► *Adaptive posterior contraction*

▷ We investigate the problem of deriving posterior contraction rates in nonparametric models under different loss functions (in particular sup-norm loss). We derive a lower bound on posterior coverages of shrinking neighbourhoods. In a second part, feasible priors are constructed that lead to adaptive rates of contraction under L^2 or sup-norm loss and that moreover achieve our lower bound.

⊕ Joint work with *Judith Rousseau, Marc Hoffmann*.

James Scott Invited 5, Wednesday, 6/24/15, 10:30 am - 11:00 am

University of Texas at Austin, USA

► *Multiscale spatial density smoothing*

▷ We consider the estimation of a spatially varying density function, motivated by problems that arise in large-scale radiological survey and anomaly detection. Four challenges make this a difficult problem. First, the density at any given spatial location may have both smooth and non-smooth features. Second, the spatial correlation is neither stationary nor isotropic. Third, the spatial correlation decays at different length scales for different parts of the density. Finally, at some spatial locations, there is very little data. We present a method called multiscale spatial density smoothing that successfully addresses these challenges. The method is motivated by the same construction that underlies a Polya-tree prior, in that it is based on a recursive dyadic partition of the underlying density function. We also describe an efficient algorithm for finding a maximum a posteriori (MAP) estimate that leverages recent advances in convex optimization for non-smooth functions.

⊕ Joint work with *Wesley Tansey, Alex Athey, Alex Reinhart*.

Christof Seiler Contributed 5B, Thursday, 6/25/15, 9:20 am - 9:40 am

Stanford University, Department of Statistics, USA

► *Parcellations of Vector Fields in Computational Anatomy*

▷ A common preprocessing task to many medical imaging studies is the registration of subject images to a common template image. Registration describes the task of finding a deformation that warps one image into another. To avoid biologically implausible deformations (like folding and tearing), we work with diffeomorphic deformations modeled with differential equations. In this talk, we take a first step towards data-driven identification of anatomical regions of interest. We propose a probabilistic model of geometric variability and describe individual patients as noisy deformations of a random spatial structure (modeled as regions) from a common template. The random regions are generated using the distance dependent Chinese Restaurant Process. We employ the Gibbs sampler to infer regions from a set of noisy deformation fields. We show preliminary results on a dataset of 400 CT images of the human spine.

⊕ Joint work with *Susan Holmes*.

Paulo Serra Contributed 6A, Thursday, 6/25/15, 4:00 pm - 4:20 pm

Institute for Mathematical Stochastics, University of Goettingen, Germany

► *Adaptive empirical Bayesian smoothing splines*

▷ Adaptive empirical Bayesian smoothing splines (arXiv:1411.6860 [math.ST]) In this talk I present empirical Bayesian smoothing splines in the context of non-parametric regression. These are smoothing splines with both smoothing parameter and penalty order determined via the empirical Bayes method. They can also be interpreted as the posterior mean corresponding to a certain empirical prior. We endow the regression function and the variance of the noise with a conjugate prior. Estimates for the smoothing parameter and the penalty order are obtained from the resulting marginal likelihood via appropriate estimating equations. We describe the asymptotic distri- bution of the smoothing parameter corresponding to each penalty order, and the asymptotic behaviour of the penalty order for different models (including parametric models). The corresponding empirical Bayesian smoothing spline is an adaptive estimator for the underlying regression function. The selected order and respective smoothing parameter are used to con- struct adaptive credible sets with good frequentist coverage over large families of models. Since the

prior is conjugate, obtaining the estimates and credible sets is straightforward. Smoothing parameters selected using unbiased risk minimisers (e.g., via generalised cross validation) also adapt to the underlying smoothness of the regression function (up to the order of the smoothing spline). However, this comes at the cost of large asymptotic variance which has a negative impact on the performance of the resulting frequentist smoothing splines. We use our credible sets as a proxy to show the superior performance of the adaptive empirical Bayesian smoothing splines compared to frequentist smoothing splines.
⊕ Joint work with *Tatyana Krivobokova*.

Weining Shen Contributed 4B, Wednesday, 6/24/15, 4:20 pm - 4:40 pm
The University of Texas MD Anderson Cancer Center, USA
► *Bayesian model selection for semi-nonparametric models*
▷ We conduct Bayesian analysis on a class of semi-nonparametric regression models with high-dimensional parametric covariates. In particular, we show (1) strong model selection consistency, where the posterior probability of the true model converges to one; and (2) joint BvM theorem, where the posterior of the selected parametric and nonparametric components jointly converges to a Gaussian vector. This is a joint work with Guang Cheng from Purdue.
⊕ Joint work with *Guang Cheng*.

Jacopo Soriano Contributed 6B, Thursday, 6/25/15, 5:00 pm - 5:20 pm
Duke University, USA
► *Variability decomposition across related mixture distributions with weight coupling and local perturbation kernels*
▷ Multi-sample comparison across mixture distributions involving local differences in some mixture components is considered. We propose a nonparametric Bayesian method that achieves high statistical power through effectively combining information across the mixture distributions. Our method incorporates local dependencies across the models through spiked priors on both the weight and the location parameters. This model allows one to test for variations in the mixture proportions and shifts in the mixture locations respectively. Simulation studies show that our method compares favorably in detecting local cross-sample differences to other state-of-the-art methods. We apply the method to a flow cytometry study, and it successfully identifies rare cell subpopulations that differentiate multiple cell samples.
⊕ Joint work with *Li Ma, Cliburn Chan*.

Sanvesh Srivastava Contributed 1A, Monday, 6/22/15, 5:00 pm - 5:20 pm
Duke University and SAMSI, USA
► *WASP: Scalable Bayes via barycenters of subset posteriors*
▷ The promise of Bayesian methods for big data sets has not fully been realized due to the lack of scalable computational algorithms. For massive data, it is necessary to store and process subsets on different machines in a distributed manner. We propose a simple, general, and highly efficient approach, which first runs a posterior sampling algorithm in parallel on different machines for subsets of a large data set. To combine these subset posteriors, we calculate the Wasserstein barycenter via a highly efficient linear program. The resulting estimate for the Wasserstein posterior (WASP) has an atomic form, facilitating straightforward estimation of posterior summaries of functionals of interest. The WASP approach allows posterior sampling algorithms for smaller data sets to be trivially scaled to huge data. We provide theoretical justification in terms of posterior consistency and algorithm efficiency. Examples are provided in complex settings including

Gaussian process regression and nonparametric Bayes mixture models.
⊕ Joint work with *Volkan Cevher, Quoc Tran-Dinh, David Dunson*.

Briana Joy Stephenson Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

The University of North Carolina at Chapel Hill, United States of America

► *Bayesian nonparametric methods to describe high-dimensional exposures*

▷ Modern data collection techniques have allowed for an increase in granularity of subject-level data. This results in a large number of potential factors for analysis. As an example, in public health research, dietary intake can be detailed down to a specific nutrient or food item consumed. Cluster analysis is used to aggregate subject data by patterns displayed in a large set of exposure variables. As the population grows, the number of clusters increases. Bayesian nonparametrics can be used to flexibly cluster exposures in an ever-increasing and dynamic population. Using data obtained from the 1997-2009 National Birth Defects Prevention Study, a population-based case-control study, this analysis examines various Bayesian nonparametric methods to describe high dimensional exposures. The different methods are evaluated to determine which approach best describes these exposures in a heterogeneous population, such as the United States.
⊕ Joint work with *Amy Herring, Andrew Olshan*.

Adam Suarez Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

North Carolina State University, USA

► *Bayesian Estimation of Principal Components for Functional Data*

▷ The area of principal components analysis (PCA) has seen relatively few contributions from the Bayesian school of inference. In this paper, we propose a Bayesian method for PCA in the case of functional data observed with error. We suggest modeling the covariance function by use of an approximate spectral decomposition, leading to easily interpretable parameters. We study in depth the choice of using the implied distributions arising from the inverse Wishart prior and prove a convergence theorem for the case of an exact finite dimensional representation. We also discuss computational issues as well as the care needed in choosing hyperparameters. A simulation study is used to demonstrate competitive performance against a recent frequentist procedure, particularly in terms of the principal component estimation. Finally, we apply the method to a real dataset, where we also incorporate model selection on the dimension of the finite basis used for modeling.
⊕ Joint work with *Subhashis Ghosal*.

Erik Sudderth Invited 1, Monday, 6/22/15, 11:30 am - 12 Noon

Brown University,

► *Scalable and Flexible Nonparametric Models of Sequential Data*

▷ We describe a novel variational bound for learning hierarchical Dirichlet process hidden Markov models. This objective favors the sticky state persistence found in many applications while penalizing models with redundant or irrelevant states, allowing Bayesian learning of an appropriate model complexity. We maintain scalability via incremental model updates based on batches of data, and increase flexibility via online structural changes to the learned state space. We apply our method to speaker diarization, motion capture, and epigenetic data, discovering models that are more compact and interpretable than competitors.
⊕ Joint work with *Michael Hughes and William Stephenson*.

Botond Szabo Invited 7, Thursday, 6/25/15, 10:30 am - 11:00 am

University of Amsterdam, The Netherlands

► *Asymptotic behaviour of the empirical Bayes posteriors associated to maximum marginal likelihood estimator*

▷ In Bayesian nonparametrics it is common to consider a family of prior distribution indexed by some hyper parameters. The best choice of the prior out of this collection crucially depends on certain characteristics (e.g. smoothness, sparseness,...) of the unknown functional parameter of interest, which are usually not available. Therefore in practice it is common to apply data dependent choices for the hyper-parameters. Arguably, the maximum marginal likelihood empirical Bayes method is one of the most well-known data-dependent Bayesian procedure. The frequentist behaviour of this method was investigated only in a few papers and mostly in specific models. Our aim is to investigate the performance of this method in a general nonparametric framework. We provide a general theorem describing the asymptotic behaviour of the empirical Bayes posterior distribution under "standard" and "natural" assumptions. Then we apply the main theorem for various models (Gaussian white noise model, nonparametric regression, density function estimation) and families of prior distributions, recovering some of the existing results in the literature, along side with new ones.

⊕ Joint work with *Judith Rousseau*.

Matt Taddy Invited 1, Monday, 6/22/15, 11:00 am - 11:30 am

Chicago Booth, USA

► *Bayesian and Empirical Bayesian Forests*

▷ We interpret random forests via the framework of distribution-free nonparametric Bayesian analysis, so that the ensemble average is an approximation to posterior mean inference for the population CART tree. This insight motivates a class of fully Bayesian Forest (BF) algorithms that provide small gains in predictive performance and large gains in interpretability (from a Bayesian perspective) over their classically bagged predecessors. The framework is then applied to derive Empirical Bayesian Forests (EBF), in which a single short tree "trunk" is estimated and Bayesian Forests are fit to the data at each leaf of the trunk. We are able to derive conditions under which this fixed trunk has high posterior probability, and demonstrate that in such settings the EBF performs nearly as well in out-of-sample prediction as the full BF. The advantage of pre-partitioning a fixed trunk is that the EBF ensembles can be fit independently for each initial partition. This implies a novel strategy for fitting tree ensemble predictors on data stored in a distributed file system (such as HDFS), and we show that it strongly outperforms the common strategy of fitting forests to without-replacement data subsets. The work is illustrated on a number of publicly available examples, and on massive data from eBay.com.

Qianwen Tan Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

Department of Statistics, NC State University, United States

► *Bayesian inference in mixed models involving ordinary differential equations*

▷ In the context of nonlinear modeling, the within-subject mechanism are often represented by a system of ordinary differential equations (ODEs), whose parameters depend on the specific subject in the underlying population. Often there is no closed form analytic solution of the equations and hence we cannot use the usual non-linear least squares technique to estimate the unknown parameters. For fixed effects a two step approach to solve this problem, consists of first fitting the data nonparametrically, and then estimating the parameter by minimizing the distance between

the nonparametrically estimated derivative and the derivative suggested by the system of ODEs. We consider a Bayesian analog of the two step method by allowing a part of the parameters to be subject-specific. These are modeled as i.i.d random variables with a fixed unknown distribution G on which we wish to make inference. In several real applications, such as pharmacokinetics and pharmacodynamics (PK/PD) models, the differential equations are generally linear in the parameters. Under this setting, we consider an efficient approach to inducing posterior distribution on the subject-specific parameters and hence on G from the posterior distribution on the nonparametric regression function obtained from a random series prior based on B-splines.
⊕ Joint work with *Subhashis Ghosal*.

Wesley Tansey Contributed 7A, Friday, 6/26/15, 10:50 am - 11:10 am
University of Texas at Austin, United States
► *False Discovery Rate Smoothing*
▷ In this talk, we present false discovery rate smoothing, a nonparametric Bayesian method for exploiting spatial structure in large multiple-testing problems. FDR smoothing extends Efron's classic two groups model to non-IID cases where there are spatially localized regions of greater signal density. It detects these regions by minimizing a non-standard loss function incorporating a graph-structured penalty, for which we detail an efficient augmented-Lagrangian algorithm. The approach then relaxes the threshold of statistical significance within these regions, and tightens it elsewhere, in a manner that controls the overall false-discovery rate at a given level. This results in increased power and cleaner spatial separation of signals from noise. Through the use of nonparametric density estimation and solution path hyperparameter tuning techniques, FDR smoothing is fully-automated and scales well to modern scientific datasets. We validate the approach quantitatively on a series of simulated examples and qualitatively by applying the method to a data set from a real-world fMRI experiment.
⊕ Joint work with *Oluwasanmi Koyejo, Russell A. Poldrack, James G. Scott*.

Yee Whye Teh Invited 6, Wednesday, 6/24/15, 3:00 pm - 3:30 pm
Statistics, Oxford, United Kingdom
► *Mondrian Forests: Efficient Random Forests for Streaming Data via Bayesian Nonparametrics*
▷ Ensembles of randomized decision trees are widely used for classification and regression tasks in machine learning and statistics. They achieve competitive predictive performance and are computationally efficient to train (batch setting) and test, making them excellent candidates for real world prediction tasks. However, the most popular variants (such as Breiman's random forest and extremely randomized trees) work only in the batch setting and cannot handle streaming data easily. In this talk, I will present Mondrian Forests, where random decision trees are generated from a Bayesian nonparametric model called a Mondrian process (Roy and Teh, 2009). Making use of the remarkable consistency properties of the Mondrian process, we develop a variant of extremely randomized trees that can be constructed in an incremental fashion efficiently, thus making their use on streaming data simple and efficient. Experiments on real world classification tasks demonstrate that Mondrian Forests achieve competitive predictive performance comparable with existing online random forests and periodically retrained batch random forests, while being more than an order of magnitude faster, thus representing a better computation vs accuracy tradeoff.
⊕ Joint work with *Balaji Lakshminarayanan and Daniel Roy*.

Jan van Waaij Contributed 3B, Tuesday, 6/23/15, 5:00 pm - 5:20 pm

University of Amsterdam, The Netherlands

► *An adaptive prior for a one-dimensional diffusion model*

▷ Pokern, Stuart and Van Zanten (2012) prove in their paper a posterior convergence result for continuous observations of an SDE $dX_t = b(X_t) + dW_t$, where b is a one-periodic diffusion and $(W_t)_{t\geq 0}$ is a Brownian motion. Their prior is not adaptive and they get suboptimal convergence rates. We extend their results with a computationally attractive prior which is adaptive for every Sobolev smoothness and has minimax convergence rates up to a log-factor.

⊕ Joint work with *Prof.dr. J.H. van Zanten*.

Harry van Zanten Keynote I, Monday, 6/22/15, 9:00 am - 10:00 am

University of Amsterdam, The Netherlands

► *Bayesian estimation of a smooth function on a large graph*

▷ We consider the problem of learning a ”smooth” function on a ”large” given graph. A proper statistical procedure should somehow incorporate the geometry of the graph (which we assume to be known) and the degree of smoothness of the function of interest (which will typically not be known). In this talk we investigate Bayesian procedures that achieve this by using a scaled Gaussian prior with a precision operator involving the Laplacian of the graph. We discuss asymptotic results that give convergence rates for the corresponding posterior under an asymptotic shape condition on the graph formulated in terms of the eigenvalues of the graph Laplacian and a smoothness condition on the function formulated in terms of the Laplacian as well.

⊕ Joint work with *Alisa Kirichenko*.

Victor Veitch Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

University of Toronto, Canada

► *On the general class of sparse exchangeable graphs on R*

▷ It is a consequence of the well-known Aldous–Hoover theorem that any random graph on N=1,2,... that is both projective and satisfies a simple probabilistic symmetry, exchangeability, must be either empty or dense. This means that the majority of random graph models currently in use are inappropriate for modelling real-world random network phenomena that result in sparse structures. A recent paper of Caron and Fox circumvents this problem by exploiting a connection between certain discrete random measures and random graphs on the reals, R, giving rise to a family of sparse, projective random graph models. In this work we extend this insight by establishing a relationship between exchangeable random measures on the plane and random graphs. We give a simple representation theorem for random graphs of this type and derive a number of their basic properties, including the expected number of nodes, and edges; the asymptotic degree structure; and the asymptotic connectivity structure. This results in a general statistical framework suitable for the analysis of real-world networks; with both power-law degree distributions and small-world behaviour arising naturally in particular examples.

Sara Wade Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

University of Cambridge, U.K.

► *A Bayesian nonparametric longitudinal model for predicting the temporal progression of Alzheimer's disease*

▷ Alzheimer's disease (AD) is a major public health concern due its damaging effects, increasing prevalence, and costs to society. In this work, we develop a Bayesian nonparametric (BNP) longi-

tudinal model to predict the temporal progression of AD. Classical longitudinal models, such as linear mixed effects models, are inappropriate as they cannot capture characteristics of the data such as non-linear trajectories or non-normal random effects with covariate-dependent behaviour. BNP methods offer an intriguing alternative, where complexity is data driven. We construct a joint BNP longitudinal model for the temporal evolution of many AD biomarkers, including the Alzheimer's disease assessment scale-cognition (ADAS-Cog), which will be used as a response to measure disease progression. We assume sigmoidal trajectories and include additional covariate information in the random effect distribution. Data is obtained from the Alzheimer's Disease Neuro-Initiative (ADNI) database, an extensive database containing neuroimaging, clinical, and biological data for cognitively normal, mild cognitive impairment, and AD patients, followed up to five years. Due to high number of predictors, we explore dimension reduction and feature selection approaches. To ensure the inference is computationally tractable, we investigate variational approaches, including non-trivial factorisations and non-conjugate priors, as well as practical approaches to speed up MCMC calculations.
⊕ Joint work with *Alexis Boukouvalas*.

Claudia Wehrhahn Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Pontificia Universidad Católica de Chile, Chile
▶ *Small sample properties of Dirichlet process mixture models for data supported on compact intervals*
▷ In this work we study the small sample properties of two Bayesian nonparametric models for single density estimation. Both models correspond to Dirichlet process mixture (DPM) of beta models and have different posterior convergence rates. Specifically, we consider the Bernstein-Dirichlet prior proposed by Petrone (1999. Scandinavian Journal of Statistics, 26: 373–393) and the DPM of mixtures of beta models proposed by Kruijer & Van der Vaart (2008. Journal of Statistical Planning and Inference 138: 1981–1992). Under the Hellinger metric, the posterior distribution of the Bernstein-Dirichlet prior and the DPM of mixtures of beta models is minimax suboptimal and optimal, respectively. Here, it is assumed that the true density to be estimated belongs to a Hölder class with α-regularity of at most $\alpha = 1$. The comparison of the model is performed using simulated data under different scenarios, sample sizes and considering different metrics. A. Jara's research is supported by Fondecyt grant 1141193.
⊕ Joint work with *Alejandro Jara, Andrés F. Barrientos*.

Ran Wei Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
North Carolina State University, USA
▶ *The nonparametric Bayesian model with continuous shrinkage priors and its application in multiple pesticide exposures data*
▷ Choosing an appropriate subset of variables related to the outcome measurements is one of the important steps in studying the overall effects of multiple factors on individual outcomes. We propose a Bayesian nonparametric regression model with multivariate continuous shrinkage priors for variable selection and model prediction. This regression model first reduces the nonparametric regression function to additive linear settings by basis expansion in a way that each chemical main effect and interaction effect is associated with a basis function and coefficient vector. The multivariate horseshoe priors, as well as other multivariate continuous shrinkage priors are imposed to shrink the coefficients of trivial variables. An auxiliary thresholding step can then discard these variables while keeping the important variables. This model is applied in the multiple pesticide exposures data to select a subset of pesticide chemicals associated with human neurobehavioral measurement. The simulation study has demonstrated the advantages of our method in variable

selection and model prediction in terms of false positives, false negatives as well as prediction mean squared errors.
⊕ Joint work with *Subhashis Ghoshal, Brian Reich*.

Sinead Williamson Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
University of Texas at Austin, USA
► *Bayesian nonparametric models for prediction in networks*
▷ We are often interested in studying the interactions between entities – for example, friendships or communications in social networks, interactions in biological networks, traffic patterns in road and communication networks. We are often interested in representing these networks in terms of some latent structure – for example, explaining the pattern of emails sent between a collection of individuals in terms of the social groups of these individuals. Most existing Bayesian approaches to network modeling focus on elucidating this underlying structure. While the resulting models are useful for characterization, they tend to perform poorly in terms of prediction of new interactions. Prediction is an important modeling goal: For example we are often interested in who a person in a social network will interact with next. This may involve out-of-sample prediction: my next email may be to someone who is not represented in the training set. In this work, we focus on predictive models for networks, where both the number of nodes and the dimensionality of the latent structure are unbounded. We demonstrate that we can capture latent structure while maintaining predictive power, and discuss possible extensions.

Keisuke Yano Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
The University of Tokyo, Japan
► *Minimax Predictive Distributions in* l_2
▷ We consider prediction in the infinite sequence model. We observe R^∞-random variable x and predict the distribution of target R^∞-random variable y. We measure the performance of the predictive distribution by the Kullback–Leibler risk and consider the minimax rate of the Kullback–Leibler risk under the parameter constraint among all the predictive distributions. We treat the problem under the ellipsoid constraint as it remains a nonparametric prediction.
⊕ Joint work with *Fumiyasu Komaki*.

Zhengwu Zhang Contributed 5B, Thursday, 6/25/15, 9:40 am - 10:00 am
Florida State University, USA
► *Bayesian Clustering of Shapes of Curves*
▷ Unsupervised clustering of curves according to their shapes is an important problem with broad scientific applications. The existing model-based clustering techniques either rely on simple probability models (e.g., Gaussian) that are not generally valid for shape analysis or assume the number of clusters. We develop an efficient Bayesian method to cluster curve data using an elastic shape metric that is based on joint registration and comparison of shapes of curves. The elastic-inner product matrix obtained from the data is modeled using a Wishart with parameters that follow certain prior models. This allows for automatic inference on the number of clusters. Posterior is sampled through an efficient Markov chain Monte Carlo procedure based on the Chinese restaurant process to infer (1) the posterior distribution on the number of clusters, and (2) clustering configuration of shapes. This method is demonstrated on a variety of synthetic data and real data examples on protein structure analysis, cell shape analysis in microscopy images, and clustering of shaped from MPEG7 database.

⊕ Joint work with *Debdeep Pati, Anuj Srivastava*.

Yize Zhao Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

SAMSI, USA

► *Bayesian Hierarchical Variable Selection for Genome-wide Association Studies*

▷ It becomes increasingly important in the genome-wide association studies (GWAS) to select important genetic information in relation to a dichotomous variable or a quantitative trait. For instance, in the ADNI study, determination of disease related genetic factors plays a vital role in the prevention of Alzheimers Disease at early stage. Currently, the discovery of biological association among SNPs motivates the strategy to construct the SNP-sets along the genome, which motivates the strategy to incorporate the grouping information into the selection procedure to increase the selection power and facilitate more biological meaningful result. To this end, we proposed a unified Bayesian framework which allows the hierarchical variable selection at both SNP-set (group) level and SNP (within group) level while simultaneously encouraging the grouping effect among SNPs. To accommodate the ultra high-dimensionality, we overcome the limitation of existing methods and propose a novel sampling scheme. By constructing an auxiliary variable selection model under SNP-set level, the new procedure utilizes the posterior samples of the auxiliary model to guide the posterior inference for the target model. We apply the proposed method to a variety of simulation studies and show that our method is computational efficient and achieve substantially better performance than completing approaches in both selection of SNP-sets and SNPs. Applying the method to the ADNI data, we identify meaningful genetic information that are highly associated with several different neuroimaging phenotypes. Our method is general and readily to be applied to a wide range of biomedical studies.

⊕ Joint work with *Fei Zou, Hongtu Zhu*.

Mingyuan Zhou Invited 4, Tuesday, 6/23/15, 2:00 pm - 2:30 pm

University of Texas at Austin, USA

► *Priors for random count matrices with random or fixed row sums*

▷ We define a family of probability distributions for random count matrices with a potentially unbounded number of rows and columns. Although random count matrices within the family are defined by a row-wise construction, their columns can be shown to be independent and identically distributed. Moreover, by analyzing these matrices' combinatorial structure, we derive the predictive distribution of a new row count vector that brings additional columns. As an example of our framework, we construct a naive-Bayes text classifier to categorize a count vector to one of several existing random count matrices of different categories. The classifier supports an unbounded number of features, and unlike most existing methods, it does not require a predefined finite vocabulary to be shared by all the categories. We further extend this framework to describe how to partition a column count vector into a latent random count matrix, whose number of nonzero columns is random and whose each row sums to a fixed integer, providing a novel way to construct exchangeable partition probability functions for mixed-membership modeling.

⊕ Joint work with *Oscar-Hernan Madrid-Padilla, James Scott*.

Haiming Zhou Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm

University of South Carolina, USA

► *Bayesian Causal Mediation Analysis for Survival Data*

▷ In this talk, Bayesian causal mediation analysis is presented for interval-censored time-to-event data. Under the sequential ignorability assumption of Imai, Keele and Yamamoto (2010), the

causal mediation analysis requires two models: one for the distribution of event times conditional on exposure, mediator and covariates (S-model), and one for the distribution of mediator conditional on the exposure and covariates (M-model). For the S-model, a general method for fitting proportional hazards, proportional odds, accelerated failure time, and accelerated hazards models via mixtures of Polya trees is presented. For the M-model, a Dependent Dirichlet process mixture prior is used to capture the possible complex distribution of mediator. An efficient Markov chain Monte Carlo algorithm is proposed for posterior computation of causal direct and indirect effects. Bayesian sensitivity analysis on the sequential ignorability assumption will also be discussed.
⊕ Joint work with *Timothy Hanson*.

Weixuan Zhu Poster II, Tuesday, 6/23/15, 8:00 pm - 10:00 pm
Universidad Carlos III de Madrid, Spain
► *A multivariate extension of a vector of two-parameter Poisson-Dirichlet processes*
▷ In the big data era there is a growing need to model the main features of large and non trivial datasets. This paper proposes a Bayesian non-parametric prior for modelling situations where data is divided into different units with different densities, allowing information pooling across the groups. Leisen and Lijoi (2011) introduced a bivariate vector of random probability measures with Poisson-Dirichlet marginals where the dependence is induced through a L évy Copula. In this paper the same approach is used for generalizing such a vector to the multivariate setting. A first important contribution is the derivation of the Laplace functional transform which is non-trivial in the multivariate setting. The Laplace transform is the basis to derive the Exchangeable Partition Probability function (EPPF) and, as a second contribution, we pro- vide an expression of the EPPF for the multivariate setting. Finally, a novel Markov Chain Monte Carlo algorithm for evaluating the EPPF is introduced and tested. In particular, numerical illustrations of the clustering behaviour of the new prior are provided.
⊕ Joint work with *Fabrizio Leisen*.

List of Participants

A

Agnelli, Robert
Ahmed, Anwar
Antoniano Villalobos, Isadora
Arbel, Julyan

B

Bandyopadhyay, Dipankar
Banerjee, Sayantan
Barcella, William
Barrientos, Andres Felipe
Battiston, Marco
Bean, Andrew
Berger, James
Bhattacharya, Anirban
Bhaumik, Prithwish
Blei, David
Bontemps, Dominique
Bose, Sudip
Boukouvalas, Alexis
Boyles, Levi
Broderick, Tamara

C

Cadonna, Annalisa
Cai, Diana
Calderon, Christopher
Camerlenghi, Federico
Canale, Antonio
Caron, Francois
Carvalho, Carlos
Castillo, Ismael
Chamberlain, Benjamin
Chen, Xilong
Cheng, Lu
Cipolli III, William
Clyde, Merlise
Colopy, Glen Wright

D

Dahl, David
Dallaire, Patrick
Das, Priyam
Davalos, Angel
De Blasi, Pierpaolo
Donelli, Nicola
Du, Xingqi
Dunson, David
Duvenaud, David

E

Edwards, Matt
Elahi, Ehtsham

F

Fall, Mame Diarra
Fan, Kai
Favaro, Stefano
Fellingham, Gilbert
Fermin, Yessica
Fernandez Aguilar, Tamara Alejandra
Ferreira, Marco
Forbes, Catherine
Foti, Nicholas
Fox, Emily
Freer, Cameron
Fuglstad, Geir-Arne

G

Gal, Yarin
Gao, Chao
Gao, Fengnan
Ge, Hong
Ghosal, Subhashis
Gillooly, Diana
Glynn, Chris
Griffin, Jim

Guglielmi, Alessandra
Guha, Aritra
Guha, Subharup
Guhaniyogi, Rajarshi
Guindani, Michele
Guo, Richard
Gutiérrez, Luis

H

Harms, Frauke
Hartog, Jarno
Heaukulani, Creighton
Heller, Katherine
Hellmayr, Christoph
Herring, Amy
Ho, Nhat
Holmes, Chris

I

Imaizumi, Masaaki

J

Johndrow, James

K

Kanazawa, Yuichiro
Kim, Gwangsu
Kirichenko, Alisa
Koehler, Elizabeth
Kottas, Athanasios
Krut, Christopher
Kundu, Suprateek

L

Lee, Eunjee
Lee, Jaeyong
Lee, Ju Hee
Lee, Taiyeong
Leisen, Fabrizio
Li, Hanning
Li, Meng
Lijoi, Antonio
Liu, Linxi
Liu, Ye
Lomeli-Garcia, Maria

M

Ma, Li
MacEachern, Steven
Martínez, Asael Fabian
McCulloch, Robert
Mena, Ramses
Mensah, David
Miller, Jeff
Mistler, Stephen
Müller, Peter
Murray, Jared
Murray, Thomas

N

Naulet, Zacharie
NeCamp, Timothy
Nguyen, Long
Nieto-Barajas, Luis
Nipoti, Bernardo
Norets, Andriy

O

Oladele, Abiodun Moruf
Orbanz, Peter
Ortega Villa, Ana Maria
Ossa, Felipe

P

Page, Garritt
Pati, Debdeep
Pearson, John
Petrone, Sonia
Pratola, Matthew

Prünster, Igor
Puggioni, Gavino

Q

Quintana, Fernando

R

Ranganath, Rajesh
Ray, Kolyan
Raykov, Yordan
Regueiro, Pedro
Reich, Brian
Ren, Boyu
Richardson, Robert
Rousseau, Judith
Roy, Daniel
Ruggiero, Matteo

S

Salomond, Jean-Bernard
Sanyal, Nilotpal
Savitsky, Terrance
Schauer, Moritz
Schein, Aaron
Schmidt-Hieber, Johannes
Scott, James
Seiler, Christof
Serra, Paulo
Shen, Weining
Soriano, Jacopo
Srivastava, Sanvesh
Stephenson, Briana
Suarez, Adam
Sudderth, Erik
Szabo, Botond

T

Taddy, Matt
Tan, Qianwen
Tansey, Wesley
Teh, Yee Whye
Thomson, Steve
Tokdar, Surya

V

van den Boom, Willem
van der Vaart, Aad
van Waaij, Jan
van Zanten, Harry
Veitch, Victor

W

Wade, Sara
Wehrhahn, Claudia
Wei, Ran
Williamson, Sinead

X

Xu, Richard

Y

Yano, Keisuke
Yazdani, Akram
Yurochkin, Mikhail

Z

Zhang, Zhengwu
Zhao, Yize
Zhou, Haiming
Zhou, Mingyuan
Zhu, Hongtu
Zhu, Weixuan

Acknowledgements

- ▶ International Society for Bayesian Analysis (Sponsoring society)
- ▶ Institute of Mathematical Statistics (Sponsoring society)
- ▶ National Science Foundation (Grant agency)
- ▶ Office of Naval Research (Grant agency)
- ▶ US Army Research Office (Grant agency)
- ▶ SAS Institute (Gold sponsor)
- ▶ Google, Inc. (Silver sponsor)
- ▶ Cambridge University Press (Supporter)