CRiSM WORKSHOP ON 'INTRACTABLE LIKELIHOOD', 15-17 MAY 2013

WEDNESDAY, 15 MAY 2013, ROOM MS.01, ZEEMAN BUILDING

14.00 – 14.50 **Richard Samworth, University of Cambridge**
High-dimensional variable selection
The fact that modern technology allows us to collect and store data on previously unimaginable scales presents significant new challenges for statisticians. One of the main topics that has received enormous attention over the last 15-20 years is variable selection, i.e. the problem of choosing which of many variables are needed in the model, even when there may be more variables than observations. I will outline some of the historical developments in this area, which have recently centred around the idea of sparsity. I will also describe some of my own work, in collaboration with Rajen Shah, where we propose Complementary Pairs Stability Selection. This very general technique, which comes with theoretical guarantees, is designed to improve the performance of any existing variable selection method under any data generating mechanism.

14.50 – 15.40 **Matthew Stephens**
Adaptive shrinkage, False Discovery Rates, and multiple comparisons:
a generic approach via Laplace approximation.

15.40 – 16.10 **Break**

16.10 – 17.00 **Eric Moulines, Institut Télécom / Télécom ParisTech (ENST)**
The Island Particle model.

The approximation of the Feynman-Kac semigroups by systems of interacting particles is a very active research field, with applications in many different areas. In this paper, we study the parallelization of such approximations. The total population of particles is divided into sub-populations, referred to as \emph{islands}. The particles within each island follow the usual selection / mutation dynamics.

The key result of this work is that it is possible to cast the dynamic at the island level into the Feynman-Kac formalism. We show that the evolution of each island is also driven by a Feynman-Kac semigroup, whose transition and potential can be explicitly related to ones of the original problem. Therefore, the same genetic type approximation of the Feynman-Kac semi-

group may be used at the island level; each island might undergo selection / mutation algorithm. We investigate the impact of the size of the population within each island and the number of the islands, and study different type of interactions within and across islands. We find conditions under which introducing interactions between islands is beneficial. The theoretical results are supported by Monte-Carlo experiments in different settings.

THURSDAY, 16 MAY 2013 – ROOM MS.03, ZEEMAN BUILDING

09.00 – 09.50 **Jose Blanchet, Columbia University**

09.50 – 10.40 **Cristiano Varin, Università Ca' Foscari Venezia**
"The Ranking Lasso"
Ranking a vector of alternatives on the basis of a series of paired comparisons is a relevant topic in many instances. A popular example is ranking contestants in sport tournaments. To this purpose, paired comparison models such as the Bradley-Terry model are often used. In this talk, I will discuss fitting paired comparison models with a lasso-type procedure that forces contestants with similar abilities to be classified into the same group. Benefits of the proposed method are easier interpretation of rankings and a significant improvement of the quality of predictions with respect to the standard maximum likelihood fitting. The proposed fitting method poses non-trivial computational difficulties that will be discussed in detail. The methodology is illustrated through ranking of teams in sport competitions and ranking of statistical journals based on citations exchange.

This talk is based on joint works with Guido Masarotto and with Manuela Cattelan and David Firth.

10.40 – 11.10 **Coffee Break**

11.10 – 12.00 **Matti Vihola**
Convergence Properties of Pseudo-Marginal Markov Chain Monte Carlo Algorithms
Pseudo-marginal Markov chain Monte Carlo (MCMC) is a generic emerging class of algorithms for computationally challenging Bayesian inference, where a density of interest cannot be evaluated (or is computationally expensive to evaluate), but it can be estimated at any point in an unbiased manner. Our work on the convergence properties of the pseudo-marginal MCMC helps to understand when the methods may be useful and what are their fundamental limitations.

Our first finding is an expected limitation: the pseudo-marginal MCMC is always worse in terms of asymptotic variance than the corresponding (ideal) marginal algorithm, where the density is available. On the other hand, when the estimators of the density are made more accurate, the resulting pseudo-marginal Markov chain will approximate the marginal algorithm in arbitrary precision in terms of the asymptotic variance. We consider also various sufficient conditions which guarantee certain rates of convergence of a pseudo-marginal algorithm in terms of the properties of the normalised density estimates and the properties of the marginal algorithm. Our results on

convergence rates imply central limit theorems.

12.00 – 14.00 **Lunch**

14.00 – 14.50 **Iain Murray, University of Edinburgh**
Flexible density estimation applied to intractable priors
I will give a case study of an inference problem with a simulation-based prior. We can draw samples (although they are computationally expensive), but we cannot evaluate the density of these samples. To improve on importance sampling from the prior, I decided to fit a density model to the simulation samples, a form of emulation. The textbook baseline density estimators, such as mixture models, underfit the data. I will argue for increased use of flexible autoregressive models as a useful family of density estimators with broad use.

14.50 – 15.40 **Nicolas Chopin, ENSAE, France**
Properties of the particle Gibbs sampler (Joint work with Sumeetpal S. Singh)
The particle Gibbs sampler is a Markov chain algorithm which operates on the extended space of the auxiliary variables generated by a interacting particle system. In particular, it samples from the discrete variables that determines the particle genealogy. We show that the corresponding Markov kernel converges in some sense to some limit kernel as the number of particles goes to infinity. We establish the ergodicity of the Particle Gibbs Markov kernel, for any number of particles, under certain assumptions. We discuss several algorithmic variations, either proposed in the literature or original. For some of these variations, we are able to prove that they strictly dominate the original algorithm in terms of either Peskun ordering, or efficiency ordering, while for the others, we provide counter-examples that they do not.

15.40 – 16.10 **Coffee Break**

16.10 – 17.00 **Christian Robert, Paris-Dauphine**
"ABC as the new empirical Bayes approach?"
Approximate Bayesian computation (ABC) has now become an essential tool for the analysis of complex stochastic models when the likelihood function is unavailable. The well-established statistical method of empirical likelihood however provides another route to such settings that bypasses simulations from the model and the choices of the ABC parameters (summary statistics, distance, tolerance), while being provably convergent in the number of observations. Furthermore, avoiding model simulations leads to significant time savings in complex models, as those used in population genetics. The ABCel algorithm we present in this talk provides in addition an evaluation of its own performances through an associated effective sample size. The method is illustrated on several realistic examples.
(Joint work with K.L. Mengersen and P. Pudlo)

FRIDAY, 17 MAY 2013, ROOM MS.03, ZEEMAN BUILDING

10.00 – 10.50 **Henry Wynn, London School of Economics**
Monomial ideals in probability and statistics
Monomial ideals are algebraic ideals generated by monomials, the building bricks multidimensional polynomials. They have a special order ideal property, which

make them useful for coding the failure (or working) sets of multi-state coherent systems. This is done my mapping minimal cuts (or paths) into minimal generators of the ideal. The so-called minimal free resolutions of the ideal leads to inclusion-exclusion bounds for system failure which are substantially tighter than classical Bonferroni bounds, at every truncation depth. The key entities are the multi-grade Hilbert series and the multi graded Betti numbers, which can be found via new fas computational algebra functions. The methods are competitive with asymptotic methods and point to applications in k-out-of-n theory, scan statistics, the design o robust systems and other areas. The statistical use of monomial ideals stems from the fact that hierarchical conditional probability models also have an associated ideal, the Stanley-Reisner ideal and there is cross-fertilization with the coherent system theory.

10.50 – 11.15 **Coffee Break**

11.15 – 12.05 **Jim Berger, Duke University**
Adaptive Importance Sampling and Exoplanet Discovery
Discovery of exoplanets (planets circling other stars) is a model selection problem, with models corresponding to the number of circling planets. Model M_q has (2+5q) parameters, where q is the number of planets. Computing the marginal likelihoods, as required in Bayesian model selection, is complicated by extreme multimodality of the likelihood function. An adaptive importance sampling scheme was developed to address this problem, incorporating a number of innovations.

12.05 – 12.55 **Mark Girolami, UCL**
Playing Russian Roulette with Intractable Likelihoods

A general scheme to exploit Exact-Approximate MCMC methodology for intractable likelihoods is suggested. By representing the intractable likelihood as an infinite Maclaurin or Geometric series expansion unbiased estimates of the likelihood can be obtained by finite time stochastic truncations of the series via Russian Roulette sampling. Whilst the estimates of the intractable likelihood are unbiased they induce a signed measure in the exact-approximate MCMC procedure which may introduce bias in the invariant distribution of the chain. By exploiting results from the Quantum Chromodynamics literature the signed measures can be employed in an exact-approximate sampling scheme in such a way that expectations with respect to the desired target distribution are preserved. This provides a general methodology to construct exact-approximate sampling schemes for a wide range of models and the methodology will be demonstrated on well known examples such as posterior inference of coupling parameters in Ising models and defining the posterior for Fisher-Bingham distributions defined on the D-Sphere. A large scale example will be provided for a GMRF model, with fine scale mesh refinement, describing the Ozone Column data, to our knowledge this is the first time that fully Bayesian inference over a model of this size has been feasible without the need to resort to any approximations. Finally a critical assessment of the strengths and weaknesses of the methodology will be considered with pointers to ongoing research. This is joint work with Anne-Marie Lyne, Heiko Strathmann, Daniel Simpson, and Yves Atchade.