Report from the Organisers: Advanced Monte Carlo Methods for Complex Inference Problems

The availability of data sets which are analytically challenging is ever increasing in areas such as Ecology, Epidemiology and Genetics. The stochastic models that represent the data generating processes are high dimensional and the only computationally feasible and accurate way to perform statistical inference is with Monte Carlo. Over the past several years a few techniques have emerged as being effective and have genuinely enthused Statisticians: Approximate Bayesian Computation, Sequential Monte Carlo (SMC) and Markov Chain Monte Carlo (MCMC).

The challenges these complex models pose have led to the development of algorithms of increased technical sophistication. For example, recently MCMC and SMC have been combined to give rise to a class of algorithms that accurately approximate "idealised" algorithms that themselves cannot be implemented but are favoured due to the accurate results they would yield. A major challenge is to articulate a general theory for such constructions. The programme's aim was to assemble the leading experts, covering theory and practice, in order to bridge gaps in understanding of the mathematical properties of these methods and to address the fundamental challenge of designing effective Monte Carlo methods for complex statistical problems.

A workshop covering the major themes started the programme and was a catalyst for the remaining three research intensive weeks. The talks included introductory elements aimed at acquainting new researchers with the subject area. It was cross-disciplinary with contributions from Engineers, Mathematical Biologists and Statisticians, but they were coherent in their message: *real problems involve complex models with many random variables, and these are costly to evaluate. Practical inference techniques must scale well with the data without compromising accuracy and efficiency.*

The workshop was well attended, swelled by daily attendees, and seeded new ideas for the research intensive period that followed in which participants delivered seminars and engaged in roundtable discussions. Problems discussed included practical SMC for high-dimensional models, MCMC for large data sets, Information Geometry in MCMC and Stochastic Approximation in

high dimensions.

The participants embraced opportunities for collaboration with enthusiasm: some involved their respective research teams overseas in collaborative discussions via video conferencing, some organised visits to the Institute for their Industrial collaborators, some travelled within the UK to give seminars and set up new links. There were surprises too: two groups working independently on SMC for Graphical Models came to know about each others work and combined forces. Grant proposals have arisen and a proposal to bring a large international Engineering conference to Cambridge has been submitted.

Other than new collaborations, an important measure of success is the new publications that are generated. It is early days but there were announcements of new results, e.g. Dr. Lee's new insights into exact Simulation using SMC, Dr. Heine's new parallelizable resampling strategies for SMC.

Word count: 469