CRiSM workshop on

Estimating Constants

20–22 April 2016, University of Warwick

Organisers: Nial Friel (UCD), Helen Ogden (Warwick), Christian Robert (Warwick)
Administrator: Olivia Garcia-Hernandez (Warwick)

Contents

1 Administrative Details **2**

2 Timetable **4**

3 Invited talks, in order of appearance **5**

4 Posters **12**

5 Participant List **19**

Map of Campus **21**

1 Administrative Details

Webpage

Any further information and announcements will be placed on the workshop webpage `www.warwick.ac.uk/estimatingconstants`

Getting Here

- Information on getting to the University of Warwick can be found at `www.warwick.ac.uk/about/visiting`
- Parking permits can be acquired at no further cost at `https://carparking.warwick.ac.uk/events/statistics-estimating-constants`

Registration and Locations

- **Registration** is open 11:00–11:45am, Wednesday 20th April, in the main atrium of the Zeeman building. Tea and coffee will be available.
- **Talks** will be held in room MS.01, Zeeman Building. MS.01 is on the ground floor, next to the atrium.
- **Breakfast** is provided for those with on-campus accommodation.
- **Lunch** is provided on Wednesday and Thursday in the atrium of the Zeeman building.
- **Dinner** is not provided (with the exception of the workshop dinner on Thursday, for those who have registered for this). There are several restaurants on campus, see the facilities section below.
- The **poster session and wine reception** is on Wednesday from 18:00, in the atrium of the Zeeman building.
- The **workshop dinner** is on Thursday at 19:00, at Arden restaurant, for those who have registered.
- The **workshop ends** at 11:45am, Friday 22nd April.

Accommodation

Accommodation is in en-suite rooms on campus. Keys can be collected from the conference reception in the student union atrium, with the exception of invited speakers who should collect their keys from Arden reception. All rooms have linen and toiletries. Rooms will be available after 15:00 for check in. All bedrooms must be vacated by 9:30am on the day of departure.

Internet Access

- **Campus:** Wireless access is most easily available via eduroam, which is supported across most of the Warwick campus. See `www.eduroam.org`.
- **Accommodation:** Wireless access is available, ask for log-in details whenever you check-in to your accommodation.

Facilities

- **Pubs and Resaurants:**
  - Xananas, Students' Union: `www.warwicksu.com/xananas`
  - Le Gusta, Arts Centre: `www.warwick.ac.uk/services/retail/legusta`
  - The Dirty Duck, Students' Union: `www.warwicksu.com/thedirtyduck`
  - For other options, see `www.warwick.ac.uk/services/retail/openingtimes`
- **Shop:** Rootes Grocery Store, next to the Students' Union. Open 8am - 8pm.
- **Arts Centre:** `www.warwickartscentre.co.uk`
- **Sports Centre:** `www.warwick.ac.uk/sport`
- **Health Centre:** `www.uwhc.org.uk`
- **Pharmacy:** Students Union Atrium. Open 9am - 6pm.

Telephone Numbers

- **Emergency:** Internal - 22222; External - 024 7652 2222
- **Security:** Internal - 22083; External - 024 7652 2083
- **Department of Statistics:** Internal - 574812; External - 024 7657 4812

Taxis

- **Swift Cabs** 024 7777 7777
- **Trinity Street Taxis** 024 7699 9999

2 Timetable

All talks will take place in room MS.01 in the Mathematics & Statistics Building.

Wednesday, 20th April

11:00 – 11:45:	*Registration and coffee*
11:45 – 12:30:	Adam Johansen
12:30 – 14:00:	*Lunch*
14:00 – 14:45:	Anne-Marie Lyne
14:45 – 15:30:	Pierre Jacob
15:30 – 16:00:	*Coffee break*
16:00 – 16:45:	Roberto Trotta
16:45 – 18:00:	'Elevator' talks
18:00 – 20:00:	Poster session, cheese and wine

Thursday, 21st April

9:00 – 9:45:	Michael Betancourt
9:45 – 10:30:	Nicolas Chopin
10:30 – 11:00:	*Coffee break*
11:00 – 11:45:	Merrilee Hurn
11:45 – 12:30:	Jean-Michel Marin
12:30 – 14:00:	*Lunch*
14:00 – 14:45:	Sumit Mukherjee
14:45 – 15:30:	Yves Atchadé
15:30 – 16:00:	*Coffee break*
16:00 – 16:45:	Michael Gutmann
16:45 – 17:30:	Panayiota Touloupou
19:00 – 22:00:	*Workshop dinner, Arden restaurant*

Friday, 22nd April

9:00 – 9:45:	Chris Sherlock
9:45 – 10:30:	Christophe Andrieu
10:30 – 11:00:	*Coffee break*
11:00 – 11:45:	Antonietta Mira

3 Invited talks, in order of appearance

Some Perspectives on Sequential Monte Carlo and Normalising Constants

Adam Johansen, University of Warwick

I will discuss the use of sequential Monte Carlo (SMC) methods to "estimate" (ratios of) normalising constants. I will begin with an introduction to SMC and its relationship to the approximation of normalising constants and move on to discuss some more recent ideas including some personal perspectives on interesting features of this approach and some open problems.

Russian roulette estimates for Bayesian inference of doubly-intractable models

Anne-Marie Lyne, Institut Curie

Doubly-intractable posterior distributions arise when the likelihood has an intractable normalising term which is a function of the unknown parameters. This occurs in a range of situations, but is most common when the data are viewed as realisations of a random graph, with the nodes representing random variables, and the edges representing a probabilistic interaction between nodes. It is difficult to carry out Bayesian parameter inference over such models, as the intractability of the normalising term means that standard sampling techniques such as the Metropolis-Hastings (MH) algorithm cannot be used.

We use Pseudo-marginal Markov chain Monte Carlo (MCMC) methodology - in which an unbiased estimate of the target can be used in place of the exact target in the MH acceptance ratio and remarkably the Markov chain converges to the same invariant distribution. To implement this approach we express the target distribution as an infinite series which is then truncated unbiasedly. As the positivity of these estimates cannot be guaranteed, we use the absolute value of the estimate in the MH acceptance ratio and afterwards correct the samples so that expectations with respect to the exact posterior are obtained. The technique is illustrated on a number of problems such as the 2-D Ising model and the Fisher-Bingham distribution.

Coupling Particle Systems

Pierre Jacob, Harvard University

In the state-space models, the normalizing constant refers to the likelihood at a given parameter value, of which particle filters give unbiased estimators. In many settings, the interest does not lie in the value of the constant itself, but in the comparison of the normalizing constants associated with different parameters. Such a comparison is facilitated by introducing positive correlations between the estimators produced by particle filters. We propose coupled resampling schemes that increase the correlation between two particle systems. The resulting algorithms improve the precision of finite-difference estimators of the score vector, and can be used in correlated pseudo-marginal algorithms. Furthermore, the coupled resampling schemes can be embedded into debiasing algorithms, yielding unbiased estimators of expectations with respect to the smoothing distribution. We will discuss the pros and cons compared to particle MCMC.

Recent advances in model likelihoods in cosmology

Roberto Trotta, Imperial College London

Bayesian model comparison in cosmology is often used as the statistical framework of choice to select among competing physical models for complex and sophisticated datasets, ranging from measurements of temperature differences in the relic radiation from the Big Bang to data on the location of hundreds of thousands of galaxies in the visible Universe.

In this talk I will review algorithmic solutions to the problem of estimating the Bayesian evidence, necessary for computing the Bayes factor, that have been developed in cosmology. I will focus in particular on nested sampling based techniques, like the MultiNest and PolyChord algorithms, and recent machine learning techniques to accelerate their computation. I will also present a computationally useful shortcut to the determination of Bayes factor for nested models, namely the Savage-Dickey density ratio.

Adiabatic Monte Carlo

Michael Betancourt, University of Warwick

By using local information to guide the exploration of a target distribution, Markov Chain Monte Carlo, in particular modern implementations like Hamiltonian Monte Carlo, has been a cornerstone of modern statistical computation. Unfortunately this local information is not generally sufficient to admit computations that require global information, such as estimating expectations with respect to multimodal distributions or marginal likelihoods. When coupled with an interpolation between the target distribution and a simpler auxiliary distribution, however, Markov Chain Monte Carlo can be an important component, for example in simulated annealing, simulated tempering, and their variants. Unfortunately, determining an effective interpolation is a challenging tuning problem that hampers these methods in practice.

In this talk I will show how the same differential geometry from which Hamiltonian Monte Carlo is built can also be used to construct an optimal interpolation dynamically, with no user intervention. I will then present the resulting Adiabatic Monte Carlo algorithm with discussion of its promise and some of the open problems in its general implementation.

The Poisson transform for unnormalised statistical models

Nicolas Chopin, ENSAE

(joint work with Simon Barthelm, GIPSA-LAB, Grenoble)
Paper available at: `http://arxiv.org/abs/1406.2839`

Contrary to standard statistical models, unnormalised statistical models only specify the likelihood function up to a constant. While such models are natural and popular, the lack of normalisation makes inference much more difficult. Here we show that inferring the parameters of a unnormalised model on a space Ω can be mapped onto an equivalent problem of estimating the intensity of a Poisson point process on Ω. The unnormalised statistical model now specifies an intensity function that does not need to be normalised. Effectively, the normalisation constant may now be inferred as just another parameter, at no loss of information. The result can be extended to cover non- IID models, which includes for example unnormalised models for sequences of graphs (dynamical graphs), or for sequences of binary vectors. As a consequence, we prove that unnormalised parametric inference in non-IID models can be turned into a semi- parametric estimation problem. Moreover, we show that the noise-contrastive divergence of Gutmann & Hyvarinen (2012) can be understood as an approximation of the Poisson transform, and extended to non-IID settings. We use our results to fit spatial Markov chain models of eye movements, where the Poisson transform allows us to turn a highly non-standard model into vanilla semi-parametric logistic regression.

Power posteriors +

Merrilee Hurn, University of Bath

One of the approaches available for estimating marginal likelihoods is thermodynamic integration. This talk will consider the method of power posteriors and recent work by various authors to maximise its efficiency and accuracy.

Hidden Gibbs random fields model selection using Block Likelihood Information Criterion

Jean-Michel Marin, University of Montpellier

Performing model selection between Gibbs random fields is a very challenging task. Indeed, because of the Markovian dependence structure, the normalizing constant of the fields cannot be computed using standard analytical or numerical methods. Furthermore, such unobserved fields cannot be integrated out, and the likelihood evaluation is a doubly intractable problem. This forms a central issue to pick the model that best fits an observed data. We introduce a new approximate version of the Bayesian Information Criterion (BIC). We partition the lattice into contiguous rectangular blocks, and we approximate the probability measure of the hidden Gibbs field by the product of some Gibbs distributions over the blocks. On that basis, we estimate the likelihood and derive the Block Likelihood Information Criterion (BLIC) that answers model choice questions such as the selection of the dependence structure or the number of latent states. We study the performances of BLIC for those questions. In addition, we present a comparison with ABC algorithms to point out that the novel criterion offers a better trade-off between time efficiency and reliable results.

Mean field Ising models

Sumit Mukherjee, Columbia University

(joint work with Anirban Basak, Duke University)

In this talk we consider the asymptotics of the log partition function of an Ising model on a sequence of finite but growing graphs/matrices. We give a sufficient condition for the mean field prediction to the log partition function to be asymptotically tight, which in particular covers all regular graphs with degree going to infinity. We show via several examples that our condition is "almost necessary" as well.

As application of our result, we derive the asymptotics of the log partition function for approximately regular graphs, and bi-regular bi-partite graphs. We also re-derive analogous results for a sequence of graphs convering in cut metric.

A Scalable quasi-Bayesian framework for graphical models

Aguemon Atchadé, University of Michigan

Doubly-intractable posterior distributions can be handled either by specialized Markov Chain Monte Carlo algorithms, or by developing a quasi-likelihood approximation of the statistical model that is free of intractable normalizing constants. For high-dimensional problems, the latter approach is more tractable and is the focus of this talk. We discuss how this approach applies to high-dimensional graphical models. And we present some results on the contraction properties of the resulting quasi-posterior distributions. Computational aspects will also be discussed.

Noise-contrastive estimation and its generalizations

Michael Gutmann, University of Helsinki

Parametric statistical models are often not properly normalized, that is, they do not integrate to unity. While unnormalized models can, in principle, be normalized by dividing them by their integral, the cost of computing the integral is generally prohibitively large. This is an issue because without normalization, the likelihood function is not available for performing inference.

I present a method called "noise-contrastive estimation" where unnormalized models are estimated by solving a classification problem. I explain some of its properties and applications, and show that it is part of a general estimation framework based on the Bregman divergence.

Related papers:
`http://jmlr.csail.mit.edu/papers/v13/gutmann12a.html`
`http://arxiv.org/abs/1202.3727`

Bayesian model selection for partially observed epidemic models

Panayiota Touloupou, University of Warwick

(joint work with Simon Spencer, Bärbel Finkenstädt Rand, Peter Neal and TJ McKinley)

Bayesian model choice considers the evidence in favour of candidate models, where in this instance each model reflects an epidemiologically important hypothesis. Model selection for epidemic models is challenging due to the need to impute a large amount of missing data, in the form of unobserved infection and recovery times. The incompleteness of the data makes the computation of the marginal likelihood, which is used to measure the evidence in favour of each model, intractable and therefore we need to find an effective way of estimating it.

In this talk, we describe an algorithm which combines MCMC and importance sampling to obtain computationally efficient estimates of the marginal likelihood in the context of epidemiology. We compare the proposed approach with several alternative methods under various simulation setups. The method is used to further our understanding of transmission dynamics of Escherichia coli O157:H7 in cattle.

Pseudo-marginal MCMC using averages of unbiased estimators

Chris Sherlock, Lancaster University

(joint work with Alexandre Thiery, National University of Singapore)

We consider pseudo-marginal MCMC where the unbiased estimator of the posterior is constructed using an average of exchangeable unbiased estimators, and compare the efficiency of a chain which uses the average of m estimators to that of a chain which uses just one of the estimators. Recent theory has shown that the chain that uses all m estimators mixes better than the chain that uses only one of them. We provide theoretical bounds on the improvement in mixing efficiency obtainable by averaging the m estimators and, motivated by this theory and by simulation studies, we discuss the translation to a choice of m for optimal computational efficiency. Staying with averages, we then consider the recent innovation of correlated pseudo-marginal MCMC.

Estimating likelihood ratios in latent variable models and its application in MCMC

Christophe Andrieu, University of Bristol

The probabilistic modelling of observed phenomena sometimes require the introduction of (unobserved) latent variables, which may or may not be of direct interest. This is for example the case when a realisation of a Markov chain is observed in noise and one is interested in inferring its transition matrix from the data. In such models inferring the parameters of interest (e.g. the transition matrix above) requires one to incorporate the latent variables in the inference procedure, resulting in practical difficulties. The standard approach to carry out inference in such models consists of integrating the latent variables numerically, most often using Monte Carlo methods. In the toy example above there are as many latent variables as there are observations, making the problem high-dimensional and potentially difficult.

We will show how recent advances in Markov chain Monte Carlo methods, in particular the development of "exact approximations" of the Metropolis-Hastings algorithm (which will be reviewed), can lead to algorithms which scale better than existing solutions.

Reduced-Variance Estimation with Intractable Likelihoods

Antonietta Mira, USI Lugano, Switzerland and Insubria University, Como, Italy

(joint work with N. Friel and C. Oates)

Many popular statistical models for complex phenomena are intractable, in the sense that the likelihood function cannot easily be evaluated. Bayesian estimation in this setting remains challenging, with a lack of computational methodology to fully exploit modern processing capabilities.We introduce novel control variates for intractable likelihoods that can reduce the Monte Carlo variance of Bayesian estimators, in some cases dramatically. We prove that these control variates are well-defined and provide a positive variance reduction. Furthermore we derive optimal tuning parameters that are targeted at optimising this variance reduction. The methodology is highly parallel and offers a route to exploit multi-core processing architectures for Bayesian estimation that complements recent research in this direction. Results presented on the Ising model, exponential random graphs and non-linear stochastic dierential equations are consistent with our theoretical findings.

4 Posters

An Adaptive MCMC Method for Multiple Changepoint Analysis with applications to Large Datasets

Alan Benson, University College Dublin

Poster board 1

We consider the problem of Bayesian inference for changepoints where the number and position of the changepoints are both unknown. In particular, we consider product partition models where it is possible to integrate out model parameters for the regime between each change point, leaving a posterior distribution over a latent binary vector indicating the presence or not of a change point at each observation. This problem has been considered by Fearnhead (2006) where one can use a filtering recursion algorithm to make exact inference. However the complexity of this algorithm depends quadratically on the number of observations. Our approach relies on an adaptive Markov Chain Monte Carlo (MCMC) method for finite discrete state spaces. We develop an adaptive algorithm which can learn from the past state of the Markov chain in order to build proposal distributions which can quickly discover where change point are likely to be positioned. We prove that our algorithm leaves the posterior distribution ergodic. Crucially, we demonstrate that our adaptive MCMC algorithm is viable for large datasets for which the exact filtering recursion approach is not. Moreover, we show that inference is possible in a reasonable time.

Bayesian inference for misspecified exponential random graph models

Lampros Bouranis, University College Dublin

Poster board 2

Exponential Random Graph models are an important tool in network analysis for describing complicated dependency structures. However, Bayesian parameter estimation for these models is extremely challenging, since evaluation of the posterior distribution typically involves the calculation of an intractable normalizing constant. This barrier motivates the consideration of tractable approximations to the likelihood function, such as pseudolikelihoods, which offer a principled approach to constructing such an approximation. Naive implementation of a posterior from a misspecified model is likely to give misleading inferences. We provide practical guidelines to calibrate in a quick and efficient manner samples coming from an approximated posterior and discuss the efficiency of this approach. The exposition of the methodology is accompanied by the analysis of real-world graphs. Comparisons against the Approximate Exchange algorithm of Caimo and Friel (2011) are provided, followed by concluding remarks.

Probabilistic Integration

Francois-Xavier Briol, University of Warwick
Poster board 3

Probabilistic numerical methods aim to model numerical error as a source of epistemic uncertainty that is subject to probabilistic analysis and reasoning, enabling the principled propagation of numerical uncertainty through a computational pipeline. The poster will present probabilistic numerical integrators based on Markov chain and Quasi Monte Carlo and prove asymptotic results on the coverage of the associated probability models for numerical integration error. The performance of probabilistic integrators is guaranteed to be no worse than non-probabilistic integrators and is, in many cases, asymptotically superior. These probabilistic integrators therefore enjoy the ”best of both worlds”, leveraging the sampling efficiency of advanced Monte Carlo methods whilst being equipped with valid probabilistic models for uncertainty quantification. Several applications and illustrations will be provided, including examples from computer vision and system modelling using non-linear differential equations.

Bayesian model comparison with un-normalised likelihoods

Richard Everitt, University of Reading
Poster board 4

Markov random field models are used widely in computer science, statistical physics and spatial statistics and network analysis. However, Bayesian analysis of these models using standard Monte Carlo methods is not possible due to their intractable likelihood functions. Several methods have been developed that permit exact, or close to exact, simulation from the posterior distribution. However, estimating the evidence and Bayes' factors (BFs) for these models remains challenging in general. This paper describes new random weight importance sampling and sequential Monte Carlo methods for estimating BFs that use simulation to circumvent the evaluation of the intractable likelihood, and compares them to existing methods. In some cases we observe an advantage in the use of biased weight estimates; an initial investigation into the theoretical and empirical properties of this class of methods is presented.

An Application of Reversible Jump MCMC and Stochastic Approximation to Molecular Design

Patrick Grinaway, Memorial Sloan Kettering Cancer Center

Poster board 5

Despite the existence of useful models for atomic-scale interactions, designing novel molecules (such as drugs) using this prior information has been extremely difficult. This difficulty results from the nature of the model–a high-dimensional markov random field over configurations of the molecule–and the desired traits, which are expectations under this complex distribution. While in prior work, MCMC would be used to sample the MRF corresponding to different molecules, in this work, we sample molecules as well. This introduces the requirement for reversible jump MCMC, as a change in molecular identity results in a change in dimensionality of the configurations. We then combine this approach with the Self-Adjusted Mixture Sampling (SAMS) technique developed by Tan, to achieve a consistent estimates of the ratios of normalizing constants for each MRF conditioned on a chemical identity. We then sought to bias sampling of molecules to prefer various properties. However, this introduces a ratio of normalizing constants into the acceptance ratio. To resolve this, we resort to running a separate MCMC sampler in parallel, using SAMS to generate consistent on-line estimates of the required ratios of partition functions. Several important challenges remain, such as improving acceptance rates for the reversible jump step, as well as improving mixing in molecule space.

A Look-Ahead Approach for Sequential Monte Carlo Methods: the iAPF

Pieralberto Guarniero, University of Warwick

Poster board 6

The poster illustrates the use of look-ahead functions in a hidden Markov model setting and present an original iterative look-ahead particle filter scheme, based on subsequent waves of particles gradually improving their path exploration efficiency. The algorithm, possibly starting with no information at all regarding the aforementioned look-ahead functions, proceeds in a forwards/backwards iterative fashion estimating the look-ahead functions, gradually improving the precision of their estimate and using them to get estimates of the models normalising constant, that corresponds to the marginal likelihood of the registered sequence of observations. Some simulation results from the algorithm implementation, showing some promising potential will be included.

Efficient sequential Monte Carlo sampling of rare trajectories in reverse time

Jere Koskela, University of Warwick

Poster board 7

Rare event simulation seeks estimate probabilities of unlikely but significant events, such as extreme weather, market crashes, or failure rates in communications networks. In complex models the probabilities of such events are often intractable, and naive simulation fails because of the low probability of the event of interest. Sequential Monte Carlo provides a practical method for sampling rare events by biasing probability mass toward the event of interest, though as always the design of good proposal distributions is difficult but crucial. The typical approach for sampling rare trajectories of a stochastic process is an exponential twisting of the forward dynamics, motivated by approximating a large deviation principle. I present an alternative, based on the observation that a forwards dynamics conditioned on ending in a rare state coincide with unconditional reverse-time dynamics started from the rare state. This observation has led to very efficient simulation methods in coalescent-based population genetics. I will introduce reverse-time SMC as a generic algorithm, discuss settings in which it is advantageous, and present some novel applications both for coalescents and other stochastic processes.

Light and Widely Applicable MCMC: an ABC perspective on MCMC for big data

Florian Maire, University College Dublin

Poster board 8

MCMC methods offer a flexible framework to estimate intractable conditional expectations that typically arise in Bayesian inference. In this work, we consider the timely topic of sampling a Markov chain targeting a posterior distribution given a ”prohibitively” large number of observations. In such a situation, off the shelf MCMC samplers such as the Metropolis-Hastings algorithm can be computationally inefficient since the likelihood function has to be evaluated at each iteration. We propose Light and Widely Applicable (LWA) MCMC, a novel approximation of the Metropolis-Hastings kernel to address this issue. Inspired by Approximate Bayesian Computation, we design a Markov chain whose transition makes use of an unknown but fixed, and arbitrary small fraction of the available data, where the random choice of sub-sample is guided by the fidelity of this sub-sample to the observed data, as measured by summary (or sufficient) statistics. We investigate the theoretical behavior of this ”noisy” but computationally efficient MCMC kernel and illustrations on diverse set of examples show how generic and flexible LWA-MCMC is. In each case LWA-MCMC yields excellent performance and in some cases a dramatic improvement compared to existing methodologies.

Approximate MCMC using Geometric Averages

Felipe Medina Aguayo, University of Warwick
Poster board 9

We present an inexact MCMC algorithm based on unbiased estimates of the target density coming from importance sampling. The main difference with respect to well-known algorithms in this context, e.g. Grouped Independence Metropolis-Hastings (GIMH) and Monte Carlo within Metropolis (MCWM), is the use of geometric averages of the importance sampling estimates instead of arithmetic averages. By doing this a bias is produced, but it can be diminished after introducing a correction based on previous work on exact algorithms with randomised acceptance. This new method may perform much better in cases where the stickiness of GIMH and the bias of MCWM are appalling. Joint work with Anthony Lee and Gareth Roberts.

Efficient Monte Carlo Methods for the Potts Model at Low Temperature

Mehdi Molkaraie, ETH Zurich
Poster board 10

We consider the problem of estimating the partition function of the ferromagnetic q-state Potts model. We propose an importance sampling algorithm in the dual of the normal/Forney factor graph representing the model. The algorithm can efficiently compute an estimate of the partition function when the coupling parameters of the model are strong (corresponding to models at low temperature) or when the model contains a mixture of strong and weak couplings. We show that, in this setting, the proposed algorithm significantly outperforms the state-of-the-art methods.

Scalable Inference for the Inverse Temperature of a Hidden Potts Model

Matthew Moores, University of Warwick
Poster board 11

The Potts model is a discrete Markov random field that can be used to label the pixels in an image according to an unobserved classification. The strength of spatial dependence between neighbouring labels is governed by the inverse temperature parameter. This parameter is difficult to estimate, due to its dependence on an intractable normalising constant. Several approaches have been proposed, including the exchange algorithm and approximate Bayesian computation (ABC), but these algorithms do not scale well for images with a million or more pixels. We introduce a precomputed binding function, which improves the elapsed runtime of these algorithms by two orders of magnitude. Our method enables fast, approximate Bayesian inference for computed tomography (CT) scans and satellite imagery.

This is joint work with Kerrie Mengersen, Tony Pettitt and Chris Drovandi at Queensland University of Technology, and Christian Robert at the University of Warwick and Université Paris Dauphine.

Optimal Bayes decision rules in cluster analysis via greedy optimisation

Riccardo Rastelli, University College Dublin
Poster board 12

In cluster analysis interest lies in capturing the partitioning of individuals into groups, whereby those belonging to the same group share similar attributes or relational profiles. Trans-dimensional Gibbs samplers have been recently applied to a variety of contexts including Finite Mixtures, Bayesian nonparametric Mixtures, Block Models for Networks and Hidden Markov Models. The output of these tools includes a posterior sample of partitions, characterising the posterior probabilities for the allocations. However, mainly due to the categorical nature of the clustering variables, creating summaries for this sample is particularly challenging. In this work we adopt a Bayesian theoretic decision approach and introduce a greedy algorithm capable of finding an optimal clustering solution with a low computational demand. The main advantage of our algorithm is its generality and the fact that it can accomodate many different loss functions and hence optimality criteria. We propose applications to Gaussian Finite Mixtures, Stochastic Block Models and Latent Block Models.

Pseudo-Marginal MCMC for parameter estimation in α-Stable distribution

Marina Riabiz, University of Cambridge
Poster board 13

The α-stable distribution describes the behaviour of normalized sums of random variables, when the finite variance hypothesis required by the Central Limit Theorem is relaxed. The distribution is governed by two key parameters, tail thickness $\alpha \in (0, 2]$ and skewness $\beta \in [-1, 1]$, in addition to scale $\sigma > 0$, and location $\mu \in \mathbb{R}$, the first two being determinant in representing heavy tailedness (the lower α is, the more probable are extreme events) and asymmetry. We focus on the parameter inference problem for $\theta = (\alpha, \beta) \in \Theta$. This task is made difficult, due to the lack of a closed form expression of the probability density $p_\theta(\mathbf{y})$, causing frequentist procedures to give approximate estimates. We discuss a Bayesian method, based on the pseudo-marginal Markov chain Monte Carlo (MCMC) approach, that requires only unbiased estimates of the intractable likelihood.

Flyweight Evidence Estimates

Ingmar Schuster, Universite Paris Dauphine
Poster board 14

We look at methods for evidence estimation that use only negligible computational effort compared to posterior estimation. This includes postprocessing of MCMC output with the Gelfand and Dey estimator, Adaptive MCMC with an Importance Sampling evidence estimator as well as advanced Importance Sampling methods for both posterior and evidence estimation. As an application we study new inference methods for nonparametric mixture models.

Kernel Adaptive Kernel Sequential Monte Carlo

Heiko Strathmann, Gatsby Unit, University College London

Poster board 15

Bayesian posterior inference with Monte Carlo methods has a fundamental role in statistics and probabilistic machine learning. Target posterior distributions arising in increasingly complex models often exhibit high degrees of nonlinearity and multimodality and pose substantial challenges to traditional samplers. We propose the Kernel Sequential Monte Carlo (KSMC) framework for building emulator models of the current particle system in a Reproducing Kernel Hilbert Space and use the emulator's geometry to inform local proposals. KSMC is applicable when gradients are unknown or prohibitively expensive and inherits the superior performance of SMC on multi-modal targets and its ability to estimate model evidence. Strengths of the proposed methodology are demonstrated on a series of challenging synthetic and real-world examples.

Quasi- and Multilevel Monte Carlo Methods for Computing Posterior Expectations in Bayesian Inverse Problems

Aretha Teckentrup, University of Warwick

Poster board 16

We consider the Bayesian inverse problem of determining the permeability in a Darcy flow model from noisy observations of functionals of the pressure head. Using Bayes' Theorem, we rewrite the posterior expectation of a quantity of interest as a ratio of prior expectations, and then compute these prior expectations using Monte Carlo type methods. We give a convergence analysis of the resulting ratio estimators for Monte Carlo, Quasi-Monte Carlo and multilevel Monte Carlo, and demonstrate the performance of the method on a simple model problem in subsurface flow. This is joint work with Rob Scheichl (Bath) and Andrew Stuart (Warwick).

On the design of informed Metropolis-Hastings proposals

Giacomo Zanella, University of Warwick

Poster board 17

We consider the problem of designing appropriate Markov chain Monte Carlo (MCMC) informed proposals in discrete spaces. In particular: assuming perfect knowledge of the target measure, what is the optimal Metropolis-Hastings proposal given a fixed set of allowed moves? Under regularity assumptions on the target, we derive the class of asymptotically optimal proposal distributions, which we call Balanced Proposals (BPs). Such proposals are asymptotically maximal elements, in terms of Peskun ordering, among proposals obtained as pointwise transformation of the target density. In continuous frameworks, this class of proposals includes the Langevin MCMC scheme and can be seen as a generalization of gradient-based methods to discrete frameworks. We discuss asymptotic analysis, applications to discrete frameworks and connections to Multiple-Try schemes.

5 Participant List

Name	Institution	Email
Timileyin Akindele Alao	Osun State University, Osogbo	timileyin.akindele@yahoo.com
Christophe Andrieu	University of Bristol	C.Andrieu@bristol.ac.uk
Aguemon Atchadé	University of Michigan	yvesa@umich.edu
Alejandra Avalos Pacheco	University of Warwick	A.Avalos-Pacheco@warwick.ac.uk
Kehinde Adekunle Bashiru	Osun State University, Osogbo	kehindebashiru@uniosun.edu.ng
Alan Benson	University College Dublin	alan.benson@insight-centre.org
Michael Betancourt	University of Warwick	betanalpha@gmail.com
Lampros Bouranis	University College Dublin	lampros.bouranis@ucdconnect.ie
Francois-Xavier Briol	University of Warwick	F-X.Briol@warwick.ac.uk
Jake Carson	University of Warwick	Jake.Carson@warwick.ac.uk
Cyril Chimisov	University of Warwick	K.Chimisov@warwick.ac.uk
Nicolas Chopin	ENSAE	nicolas.chopin@ensae.fr
Theo Damoulas	University of Warwick	td47@nyu.edu
Richard Everitt	University of Reading	r.g.everitt@reading.ac.uk
Josh Fass	Memorial Sloan Kettering Cancer Center	joshua.fass@gmail.com
Nial Friel	University College Dublin	nial.friel@ucd.ie
Habib Ganjgahi	University of Warwick	H.Ganjgahi@warwick.ac.uk
Mark Girolami	University of Warwick	M.Girolami@warwick.ac.uk
Patrick Grinaway	Memorial Sloan Kettering Cancer Center	patrick.grinaway@choderalab.org
Pieralberto Guarniero	University of Warwick	P.Guarniero@warwick.ac.uk
Michael Gutmann	University of Helsinki	michael.gutmann@helsinki.fi
Merrilee Hurn	University of Bath	masmah@bath.ac.uk
Pierre Jacob	Harvard University	pjacob@fas.harvard.edu
Paul Jenkins	University of Warwick	P.Jenkins@warwick.ac.uk
Tom Jin	University of Warwick	S.Jin@warwick.ac.uk
Adam Johansen	University of Warwick	a.m.johansen@warwick.ac.uk
Jere Koskela	University of Warwick	jere.koskela@gmail.com
Divakar Kumar	University of Warwick	D.Kumar@warwick.ac.uk
Krzysztof Latuszynski	University of Warwick	k.latuszynski@gmail.com
Anne-Marie Lyne	Institut Curie	lyne.annemarie3@gmail.com
Florian Maire	University College Dublin	florian.maire@ucd.ie

Continued on next page

Name	Institution	Email
Jean-Michel Marin	University of Montpellier	jean-michel.marin@umontpellier.fr
Felipe Medina Aguayo	University of Warwick	F.J.Medina-Aguayo@warwick.ac.uk
Anne Miloschewski	University of Warwick	A.Miloschewski@warwick.ac.uk
Antonietta Mira	Lugano, Switzerland and Como, Italy	antonietta.mira@usi.ch
Mehdi Molkaraie	ETHZ	mehdi.molkaraie@alumni.ethz.ch
Matt Moores	University of Warwick	M.T.Moores@warwick.ac.uk
Sumit Mukherjee	Columbia University	sm3949@columbia.edu
Helen Ogden	University of Warwick	h.ogden@warwick.ac.uk
Murray Pollock	University of Warwick	m.pollock@warwick.ac.uk
Riccardo Rastelli	University College Dublin	riccardo.rastelli@ucdconnect.ie
Lewis Rendell	University of Warwick	L.Rendell@warwick.ac.uk
Marina Riabiz	University of Cambridge	mr622@cam.ac.uk
Christian Robert	University of Warwick	bayesianstatistics@gmail.com
Ingmar Schuster	Université Paris Dauphine	ingmar.schuster@dauphine.fr
Chris Sherlock	Lancaster University	c.sherlock@lancaster.ac.uk
Simon Spencer	University of Warwick	S.E.F.Spencer@warwick.ac.uk
Julien Stoehr	University College Dublin	julien.stoehr@ucd.ie
Heiko Strathmann	University College London	heiko.strathmann@gmail.com
Nick Tawn	University of Warwick	N.Tawn@warwick.ac.uk
Aretha Teckentrup	University of Warwick	A.Teckentrup@warwick.ac.uk
Yuttapong Thawornwattana	University College London	yuttapong.thawornwattana.09@ucl.ac.uk
Panayiota Touloupou	University of Warwick	P.Touloupou@warwick.ac.uk
Roberto Trotta	Imperial College London	r.trotta@imperial.ac.uk
Simon White	MRC Biostatistics Unit	simon.white@mrc-bsu.cam.ac.uk
Giacomo Zanella	University of Warwick	zanella.gcm@gmail.com
Anyi Zou	University of Warwick	A.Zou@warwick.ac.uk


Scarman
WESTWOOD WAY
SCARMAN ROAD
GIBBET HILL ROAD
Radcliffe
Warwick Business School (WBS)
Gatehouse
UNIVERSITY ROAD
A45, M42, M5, M6, BIRMINGHAM
M6, M69, M1 LONDON
KIRBY CORNER ROAD
Tennis Centre
Games Hall
University House
Arden
EMB
Millburn House
Cryfield Sports Pavilion
LEIGHFIELD ROAD
Ramphal
Physics
Social Sciences
Bus Interchange
LIBRARY ROAD
Chemistry
Physical Sciences
Sciences
IARC
Library
Arts Centre
Chaplaincy
Modern Records Centre
Engineering
IMC
SERVICE ROAD ONLY
Rootes Building
Rootes
Students' Union
ATM
Senate House
Humanities
Computer Science
Mathematics
Zeeman
IDL
IIPSI
Psychology
Bluebell
CENTRAL CAMPUS
Sports Centre
HEALTH CENTRE ROAD
Health Centre
Jack Martin
Tocil
Claycroft
Arthur Vick
Argent Court
LYNCHGATE ROAD
SIR WILLIAM LYONS ROAD
MILLBURN HILL ROAD
A46, M40 WARWICK,
KENILWORTH, LEAMINGTON SPA
KEY TO SYMBOLS
Accommodation
Bus stop
Construction zones
Information totem
Male, female, disabled
Male, female, disabled (inc. baby change)
Pay & display parking
Multi-storey pay & display parking
Pay & display parking (Without accessibility parking)
Shopping on campus
Sustrans cycle route
Taxi rank