Christmas Workshop on sequential Monte Carlo and related methods

---

Monday 21st of December

Sylvain Rubenthaler

Title : Stability of the optimal filter in continuous time: beyond the Benes filter.

Abstract: We exhibit a case in which the optimal filter is stable with respect to its initial condition. The computations are such that this result will allow to prove the stability of an associated particle filter. The set of assumptions we propose is a neighbourhood of the Benes filter. Observations in continuous time could be understood as "arbitrarily frequent observations". The proof is based on a decomposition which takes us back to optimal filters with observations which are not continuous in time.

Kari Heine

Title: Fluctuations of a distributed particle filter with local exchange

Abstract: We study a distributed particle filter proposed by Bolic et al.~(2005). This algorithm involves m groups of M particles, with interaction between groups occurring through a ``local exchange'' mechanism. We establish a central limit theorem in the regime where the group size M is fixed and the number of groups tends to infinity. A formula we obtain for the asymptotic variance can be interpreted in terms of colliding Markov chains, enabling analytic and numerical evaluations of how the asymptotic variance behaves over time, with comparison to a benchmark algorithm consisting of m independent particle filters.

This is a joint work with Nick Whiteley (University of Bristol).

Joris Bierkens

Title: Zig Zag Sampling

Abstract: In Turitsyn, Chertkov and Vucelja (2011), a non-reversible Markov Chain Monte Carlo (MCMC) method on an augmented state space was introduced, here referred to as Lifted Metropolis-Hastings (LMH). A scaling limit of the magnetization process in the Curie-Weiss model is derived for LMH, as well as for Metropolis-Hastings (MH). These scaling limits perfectly explain the empirical results of Turitsyn et al. (2011). The scaling limit of LMH turns out to be a non-reversible piecewise deterministic exponentially ergodic `zig zag' Markov process.

As it turns out, a multidimensional extension of the zig zag process can be formulated, which allows one to sample from general target densities in R^n. It seems that the new sampling method ('Zig Zag Sampling') is ergodic but a proof of this is not entirely trivial. It also seems that Zig Zag sampling allows for a simple but very efficient implementation which makes it comparable or even more efficient than e.g.

Metropolis Adjusted Langevin (MALA), without the need for any tuning. This is research in progress.

Omiros Papaspiliopoulos

TBA

Nicolas Chopin

Title: Leave Pima Indians alone: binary regression as a benchmark for Bayesian computation

Abstract: Whenever a new approach to perform Bayesian computation is introduced, a common practice is to showcase this approach on a binary regression model and datasets of moderate size. This paper discusses to which extent this practice is sound. It also reviews the current state of the art of Bayesian computation, using binary regression as a running example. Both sampling-based algorithms (importance sampling, MCMC and SMC) and fast approximations (Laplace and EP) are covered. Extensive numerical results are provided, some of which might go against conventional wisdom regarding the effectiveness of certain algorithms. Implications for other problems (variable selection) and other models are also discussed.

joint work with James Ridgway
http://fr.arxiv.org/abs/1506.08640

Sumeetpal Singh

Title: Blocking Strategies and Stability of Particle Gibbs Samplers

Abstract: Sampling from the conditional (or posterior) probability distribution of the latent states of a Hidden Markov Model, given the realization of the observed process, is a non-trivial problem in the context of Markov Chain Monte Carlo. To do this Andrieu et al. (2010) constructed a Markov kernel which leaves this conditional distribution invariant using a Particle Filter. From a practitioner's point of view, this Markov kernel attempts to mimic the act of sampling all the latent state variables as one block from the posterior distribution but for models where exact simulation is not possible. There are some recent theoretical results that establish the uniform ergodicity of this Markov kernel and that the mixing rate does not diminish provided

the number of particles grows at least linearly with the number of latent states in the posterior. This gives rise to a cost, per application of the kernel, that is quadratic in the number of latent states which could be prohibitive for long observation sequences. We seek to answer an obvious but important question: is there a different implementation with a cost per-iteration that grows linearly with the number of latent states, but which is still stable in the sense that its mixing rate does not deteriorate? We address this problem using blocking strategies, which are easily parallelizable, and prove stability of the resulting sampler.

Joint work with F. Lindsten and E. Moulines

Randal Douc

Title: Posterior consistency for partially-observed Markov models

Abstract: In this work, we establish the posterior consistency property for a parametrized family of partially observed fully-dominated Markov models. The main assumption is expressed in terms of the Kullback neighborhood associated to one transition of the complete Markov chain, which should be of positive probability with respect to the prior distribution. We show that under some additional mild assumptions, the posterior consistency can be derived in straight line from the consistency of the Approximate Maximum Likelihood estimator. The result is then extended to possibly non-compact parameter sets and non-stationary observations. We check our assumptions on examples including stochastic volatility models and non-linear switching autoregressions with underlying Markov chains of order m.

Joint work with Jimmy Olsson and Francois Roueff

Nick Whiteley

Title: Variance estimation and asymptotically optimal allocation in the particle filter

Abstract: Particle filters provide approximations of marginal likelihoods and filtering expectations in hidden Markov models. However, estimating the Monte Carlo variance of these approximations, without generating multiple independent realizations of the approximations themselves, is not straightforward. We present an unbiased estimator of the variance of the marginal likelihood approximation, and consistent estimators of the asymptotic variance of the approximations of the marginal likelihood and filtering expectations. These estimators are byproducts of a single run of a particle filter and have no added computational complexity or storage requirements. With additional storage requirements, one can also consistently estimate higher-order terms in the non-asymptotic variance. When the number of particles is allowed to vary over time, we show how one can approximate their asymptotically optimal allocation.

Alex Mijatovic

Title: On the Poisson equation for Metropolis-Hastings chains

Abstract: We define an approximation scheme for a solution of the Poisson equation of a geometrically ergodic Metropolis-Hastings chain Φ. The approximations give rise to a natural sequence of control variates for the ergodic average $S_k(F)=(1/k)\sum_{i=1}^{k}F(\Phi_i)$, where F is the force function in the Poisson equation. The main result of the paper shows that the sequence of the asymptotic variances (in the CLTs for the control-variate estimators) converges to zero. We apply the algorithm to geometrically and non-geometrically ergodic chains and present numerical evidence for a significant variance reduction in both cases.

Tuesday 22nd December

Sebastian Reich

Title: Ensemble transform particle filters for continuous-time filtering

Abstract: I will discuss an extension of ensemble transform particle filters, as proposed recently for sequential data assimilation,
to the continuous-time filtering context. This extension can be viewed as a mesh-free/integral free alternative to the Mckean-Vlasov approach proposed by Crisan and Xiong and the feedback particle filter of Yang, Mehta, and Meyn, which avoids an undesirable division by the unknown marginal PDF. I will discuss algorithmic approaches, such as hybrid filters, Gibbs flow, and localisation, for tackling the curse of dimensionality and for reducing the computational complexity of continuous-time ensemble transform particle filters.

Ajay Jasra

Title: Multilevel Sequential Monte Carlo Samplers.

Abstract: The approximation of expectations w.r.t. probability distributions associated to the solution of partial differential equations (PDEs) is considered herein; this scenario appears routinely in Bayesian inverse problems. In practice, one often has to solve the associated PDE numerically, using, for instance finite element methods and leading to a discretisation bias, with step-size level h_L. In addition, the expectation cannot be computed analytically and one often resorts to Monte Carlo methods. In the context of this problem, it is known that the introduction of the multi-level Monte Carlo (MLMC) method can reduce the amount of computational effort to estimate expectations, for a given level of error. This is achieved via a telescoping identity associated to a Monte Carlo approximation of a sequence of probability distributions with discretisation levels $\infty>h_0>h_1\cdots>h_L$. In many practical problems of interest, one cannot achieve an i.i.d. sampling of the associated sequence of probability distributions. A sequential Monte Carlo (SMC) version of the MLMC method is introduced to deal with this problem. It is shown that under appropriate assumptions, the attractive property of a reduction of the amount of computational effort to estimate expectations, for a given level of error, can be maintained in the SMC context. The approach is numerically illustrated on a Bayesian inverse problem. This is a joint work with Yan Zhou (NUS), Kody Law (ORNL), Raul Tempone (KAUST) and Alex Beskos (UCL).

Eric Moulines

Title: The pros and cons of the unadjusted Langevin algorithm to sample over a high-dimensional space.

Abstract: Sampling over high-dimensional space has become a prerequisite in the applications of Bayesian statistics to machine learning problem. In many situations of interest, the log-posterior distribution is concave. The likelihood part is generally smooth and gradient Lipshitz while the prior is concave but typically not smooth (the archetypical problem is the LASSO or the elastic-net penalty, bu many other problems can be cast into this framework).

We will describe methods to sample such distributions, which are adapted from the state-of-the-art optimization procedures which have been developed in this context. We will also provide convergence in total variation and Wasserstein distance to the equilibrium, showing explicitly the dependence in the dimension of the parameter space and the sparsity (effective dimension of the model). We will also show how to obtain non-asymptotic bounds for additive functionals of these chains which provide "sensible" estimates. We will discuss several applications and possible extensions to the non-convex case and / or the non-smooth case.

Adam Johansen

Title: On "Particle Filters" for Parameter Estimation

Abstract: Sequential Monte Carlo methods have been very widely applied to approximate the optimal filtering and smoothing problems. When the objective is offline parameter estimation within a simulated likelihood or particle Markov chain Monte Carlo context, it is possible to relax many of the constraints which exist in the online filtering setting. This talk will discuss three simple approaches to improving estimation of parameters in this context by employing Sequential Monte Carlo algorithms which target sequences of distributions other tan the natural filtering distributions within this context.

The work presented includes joint work with Arnaud Doucet, Pieralberto Guarniero and Anthony Lee.

George Deligiannidis

Title: A correlated pseudo marginal algorithm.

Abstract: The pseudo-marginal algorithm is a popular variant of the Metropolis-Hastings which allows one to handle intractable likelihoods functions, by replacing them by an unbiased estimator often arising from importance sampling or particle filters. Recent theoretical results have established that the cost per iteration of such schemes is for many common applications quadratic in the number of observations T. This cost may be prohibitive for large datasets. We present a new method which significantly reduces this cost in various applications, by introducing correlation between successive values of the estimated likelihood.

Yvo Pokern

Title: Gibbs Flow for approximate transport with application to Bayesian computation.

Abstract: A measurable function T mapping R^d to itself such that $T(X)\sim\pi_{1}$ if $X\sim\pi_{0}$ is called a transport map transporting the measure π_{0}, e.g. a prior distribution, to π_{1}, e.g. a posterior distribution. If one could obtain an analytical expression for a transport map from any π_{0} to any π_{1} then this could be straightforwardly applied to sample from any distribution. One would map draws from an easy-to-sample distribution π_{0} to the target distribution π_{1} using this transport map. Although it is usually impossible to obtain an explicit transport map for complex target distributions, we show here how to build a tractable approximation of a novel transport map. This is achieved by moving samples from π_{0} using an ordinary differential equation whose drift is a function of the full conditional distributions of the target. Even when this ordinary differential equation is time-discretized and the full conditional distributions are approximated numerically, the resulting distribution of the mapped samples can be evaluated and used as a proposal within Markov chain Monte Carlo and sequential Monte Carlo schemes. We show experimentally that it can improve significantly performance of state-of-the-art sequential Monte Carlo methods for a fixed computational
complexity.

Carles Breto

Title: Plug-and-play SMC methodology applied to statistical modeling of infectious disease dynamics

Abstract: Plug-and-play inference algorithms based on SMC methods facilitate scientific model development for dynamic systems. State-space models for such systems may be characterized by measurement and transition densities. This transition density is often used by inference algorithms for either simulating from it (i.e., plug-and-play) or evaluating it. However, such evaluation requires closed-form expressions that may be intractable for models of scientific interest. In this context, statistical analysis has been facilitated by recent advances in plug-and-play approaches that will be presented, focusing on the R software package “pomp” (partially observed Markov processes) and on the use of iterated filtering algorithms for developing new classes of models and for state-of-the-art analysis of infectious disease data.

Axel Finke

Title: Pseudo-Marginal Monte Carlo Optimisation

Abstract: We extend existing Monte Carlo schemes for performing optimisation in latent-variable settings, i.e. in situations in which the objective function is intractable. This often occurs when performing (marginal) maximum-likelihood or (marginal) maximum-a-posteriori estimation.

To this end, we present a flexible framework for combining the SAME algorithm from Doucet, Godsill & Robert (2002) with state-of-the-art MCMC kernels such as pseudo-marginal MCMC or conditional SMC kernels. We also construct population-based approaches by incorporating these ideas into SMC samplers.

This is joint work with Adam M. Johansen.

Rodrigo Targino

Title: Sequential Monte Carlo Samplers for capital allocation under copula-dependent risk models

Abstract: In this talk we assume a multivariate risk model has been developed for a portfolio and its capital derived as a homogeneous risk measure. The Euler (or gradient) principle, then, states that the capital to be allocated to each component of the portfolio has to be calculated as an expectation conditional to a rare event, which can be challenging to evaluate in practice. We exploit the copula-dependence within the portfolio risks to design a Sequential Monte Carlo Samplers based estimate to the marginal conditional expectations involved in the problem, showing its efficiency through a series of computational examples.

Wednesday 23rd December

Axel Gandy

Title: The chopthin algorithm for resampling

Abstract: Resampling is a standard and important step in particle filters and more generally sequential Monte Carlo methods. In this presentation, we introduce a new method, called chopthin, for resampling weighted particles. In contrast to standard resampling methods the algorithm does not produce a set of equally weighted particles; instead it merely enforces an upper bound on the ratio between the weights. Simulation studies show that the chopthin algorithm consistently outperforms standard resampling methods. The algorithms chops up particles with large weight and thins out particles with low weight, hence its name. It implicitly guarantees a lower bound on the effective sample size. The algorithm can be implemented very efficiently, making it practically useful. Implementations for C++, R (on CRAN), Python and for Matlab are available.

Kostas Kalogeropoulos

Title: Sequential Monte Carlo for stochastic volatility models with memory

Abstract: We consider continuous-time stochastic volatility models driven by fractional Brownian motion. Due to the non-Markovianity and high-dimensionality of the latent paths, estimating posterior expectations is a computationally challenging undertaking. We present a reparameterization framework based on the Davies and Harte method for sampling stationary Gaussian processes and use this framework to construct a hybrid Monte Carlo algorithm that allows computationally efficient

Bayesian inference despite the high-dimensional latent variables arising in this context. The hybrid Monte Carlo algorithm can either be used on its own or undertake the task of the rejuvenation step in a sequential Monte Carlo algorithm. The latter not only allows a more robust computational scheme, but also offers a model choice framework that can be based on either i) the marginal likelihood estimates or ii) model predictive performance through the sequences of posterior predictive distributions. The methodology is illustrated on both simulated data and the S&P500 time series.

Alexandros Beskos

TBA

Joaquin Miguez

Title: Nested particle filtering for recursive parameter estimation in state space models

Abstract: We address the problem of approximating the probability measure of the fixed parameters of a state-space dynamic system using a sequential Monte Carlo method. The proposed approach relies on a nested structure that employs two layers of particle filters to approximate the posterior probability law of the static parameters and the dynamic variables of the system of interest. The proposed algorithm operates in a purely sequential and recursive manner. In particular, the computational complexity of the recursive steps of the proposed method is constant over time. We analyse the approximation of integrals of real bounded functions with respect to the posterior distribution of the system parameters computed via the proposed scheme. For a finite time horizon, we prove that the approximation errors converge in L_p under mild assumptions and with rate proportional to $\frac{1}{\sqrt{N}}$, where N is the number of Monte Carlo samples in the parameter space. Under a set of stronger assumptions related to the stability of the optimal filter for the model, we prove that the convergence of the L_p norms of the approximation errors is uniform over time. The uniform convergence result has some relevant consequences. One of them is that the proposed scheme can asymptotically ``identify'' (estimate exactly) the parameter values for a class of state-space models. A subset of the assumptions that yield uniform convergence also leads to a positive lower bound, uniform over time and the number of particles, on the normalised effective sample size of the filter. We will show some simple numerical examples that illustrate some of the theoretical findings.

Jochen Broecker

Title: On a strong notion of reliability for ensemble forecasts

Abstract: An important notion in the context of ensemble forecasts (e.g. generated by particle filters or ``ensemble'' versions of the Kalman filter) is forecast reliability. Definitions of reliability (as used in operational weather and climate forecasts for instance) basically state that the verification should be statistically indistinguishable from the ensemble members. The main deficit with these definitions is that no statement is made about the joint distribution of a series of verification--ensemble pairs. This precludes a precise analysis of reliability tests (both existing and to be devised). In the context of particle filters and sequential monte carlo methods, the reliability of forecasts is often analysed in terms of a few moments only, while the reliability of the forecast as a whole is not assessed.

In this talk, a new notion of reliability, called strong reliability, will be presented which alleviates these problems to some extent. Strong reliability is a natural extension of existing notions of reliability yet it provides a basis for rigorous treatment of reliability tests. Both existing and new reliability tests will be discussed in this light.