Abstracts

Dan Crisan, Imperial College London

Talk Monday: **Data Assimilation using Particle Filters**

The purpose of this talk is twofold: First I will attempt to provide a common platform of communication between two sister areas of research, data assimilation and stochastic filtering, particularly its numerical component. Second, I will describe the major difficulties encountered when solving the filtering problem in high dimensions and suggest a number of possible remedies for overcoming these difficulties.

Talk Tuesday: **Data Assimilation for Stochastic Transport Models**

I will present a new approach of introducing stochasticity in a large class of models for fluid dynamical systems that preserve some of the natural conservations laws satisfied by the deterministic models. I will discuss the statistical calibration of the stochastic noise, the immediate applications to the uncertainly quantification of the models and an application of Particle Filters to data assimilation for such systems. I will show how one can apply the approach to dynamical systems modelled by the 2d Euler equation and by the 2d multilayer quasi-geostrophic equation. This is joint work with Colin Cotter, Darryl Holm, Wei Pan and Igor Shevchenko.

Aretha Teckentrup (University of Edinburgh)

Talk on Monday: **Surrogate Models in Large-Scale Bayesian Inverse Problems**

We are interested in the inverse problem of estimating unknown parameters in a mathematical model from observed data. We follow the Bayesian approach, in which the solution to the inverse problem is the distribution of the unknown parameters conditioned on the observed data, the so-called posterior distribution. We are particularly interested in the case where the mathematical model is non-linear and expensive to simulate, for example given by a partial differential equation.

In this first talk, we consider the use of surrogate models to approximate the Bayesian posterior distribution. We present a general framework for the analysis of the error introduced in the posterior distribution, and discuss particular examples of surrogate models such as Gaussian process emulators and randomised misfit approaches.:

Talk on Tuesday: **Multilevel Sampling Methods for Large-Scale Bayesian Inverse Problems**

We are interested in the inverse problem of estimating unknown parameters in a mathematical model from observed data. We follow the Bayesian approach, in which the solution to the inverse problem is the distribution of the unknown parameters conditioned on the observed data, the so-called posterior distribution. We are particularly interested in the case where the mathematical model is non-linear and expensive to simulate, for example given by a partial differential equation.

In this second talk, we consider the use of multilevel strategies to significantly reduce the computational cost of Monte Carlo type sampling methods for computing expectations with respect to the posterior distribution. We study two approaches to introducing multilevel methods in the context of Bayesian inverse problems, one based on ratio estimators and one based on Markov chain Monte Carlo.

Mark Girolami (Imperial College London)

Talk on Monday: **Diffusions and Dynamics on Statistical Manifolds for Statistical Inference**

Abstract: The use of Differential Geometry in Statistical Science dates back to the early work of C.R.Rao in the 1940's when he sought to assess the natural distance between population distributions. The Fisher-Rao metric tensor defined the Riemannian manifold structure of probability measures and from this local manifold geodesic distances between probability measures could be properly defined. This early work was then taken up by many authors within the statistical sciences with an emphasis on the study of the efficiency of statistical estimators. The area of Information Geometry has developed substantially and has had major impact in areas of applied statistics such as Machine Learning and Statistical Signal Processing.

A different perspective on the Riemannian structure of statistical manifolds can be taken to make breakthroughs in the contemporary statistical modelling problems. Langevin diffusions and Hamiltonian dynamics on the manifold of probability measures are defined to obtain Markov transition kernels for Monte Carlo based inference. This work was motivated by the many challenges presented by contemporary problems of statistical inference, such as for example inference over partial differential equations describing complex physical engineering systems. This lecture aims to provide an accessible introduction to the Langevin and Hamiltonian Monte Carlo methods.

Talk on Tuesday: **Probabilistic Numerical Computation: a Role for Statisticians in Numerical Analysis?**

Abstract: Consider the consequences of an alternative history. What if Leonhard Euler had happened to read the posthumous publication of the paper by Thomas Bayes on "An Essay towards solving a Problem in the Doctrine of Chances"? This paper was published in 1763 in the Philosophical Transactions of the Royal Society, so if Euler had read this article, we can wonder whether the section in his three volume book Institutionum calculi integralis, published in 1768, on numerical solution of differential equations might have been quite different.

Would the awareness by Euler of the "Bayesian" proposition of characterising uncertainty due to unknown quantities using the probability calculus have changed the development of numerical methods and their analysis to one that is more inherently statistical?

Fast forward the clock two centuries to the late 1960s in America, when the mathematician F.M. Larkin published a series of papers on the definition of Gaussian Measures in infinite dimensional Hilbert spaces, culminating in the 1972 work on "Gaussian Measure on Hilbert Space and Applications in Numerical Analysis". In that work the formal definition of the mathematical tools required to consider average case errors in Hilbert spaces for numerical analysis were laid down and methods such as Bayesian Quadrature or Bayesian Monte Carlo were developed in full, long before their independent reinvention in the 1990s and 2000s

brought them to a wider audience.

Now in 2017 the question of viewing numerical analysis as a problem of Statistical Inference in many ways seems natural and is being demanded by applied mathematicians, engineers and physicists who need to carefully and fully account for all sources of uncertainty in mathematical modelling and numerical simulation.

Now we have a research frontier that has emerged in scientific computation founded on the principle that error in numerical methods, which for example solves differential equations, entails uncertainty that ought to be subjected to statistical analysis. This viewpoint raises exciting challenges for contemporary statistical and numerical analysis, including the design of statistical methods that enable the coherent propagation of probability measures through a computational and inferential pipeline.

Alexandra Carpentier (University of Potsdam/University of Magdeburg)

Title: **Sequential Learning under dependant data: from stochastic to adversarial bandits**

The field of sequential learning focuses on learning problems where the data are not available in batch form at the beginning of the learning process, but have to be collected online in an active manner by the learner. A specific instance of this problem is the bandit problem. Efficient sampling strategies in this context depend heavily on the underlying mechanism that generate the data. On the one end of the spectrum, the data is generated in an independent fashion (stochastic setting), while in the other hand of the spectrum, the underlying mechanism is arbitrary (adversarial setting). Project A03 focuses in interpolating between these two extremes.

Wilhelm Stannat (TU Berlin)

Title: **Long-time stability and accuracy of the ensemble Kalman-Bucy filter**

We discuss the main mathematical features of the ensemble Kalman filter applied to filtering problems in continuous time. Results concerning stability and accuray will be presented in the case of fully observed processes and small measurement noise. To assess consistency, we also derive the dynamical mean-field equation for the empirical distribution in the infinite ensemble limit and compare it with the Kushner-Stratonovich equation for the posterior distribution of the optimal filter. Major open questions in the partially observed case as well as generalizations to second order accurate transform filters will be discussed.

Yuri Shprits (GFZ Potsdam)

Title: **Reconstructing evolution of the near-Earth radiation environment from single point measurements**

Energetic particles in the near-Earth space that are commonly referred to as Van Allen radiation belts, or trapped radiation, pose a significant risk to Earth-orbiting satellites and humans in space. During geomagnetic storms, the radiation in the near-Earth space can dramatically increase, and numerous anomalies are often reported by satellite operators. Satellite measurements are usually limited to a single location in space, which complicates the data analysis geared towards reproducing the global state of the radiation belts. Up until recently, most of the research concentrated on the analysis of data from individual spacecraft, which does not allow for inferring the global evolution of the radiation environment.

Analysis is complicated by the fact that measurements are given at individual satellite locations, have different instrumental errors, and often vary by orders of magnitude. We discuss newly developed data assimilation methods and outline a proposal to develop new methods that will enable efficient data assimilation from multiple satellite missions into complex physics-based models for the evolution of energetic and relativistic particles. The developed data assimilation methods for systems where quantities may vary by several orders of magnitude may be used in the future for other applications in space sciences and potentially other areas of science and engineering.

Ralf Engbert (University of Potsdam)

Title: **Likelihood-based model evaluation in cognitive science**

Dynamical models are increasingly important in cognitive science. Based on two examples, I will demonstrate current goals and problems of likelihood-based model evaluation in the field. First, in models of eye-movement control during scene viewing, we are investigating dynamical processes of attention and oculomotor functioning to explain statistical patterns of spatial point processes. Second, in eye guidance during reading, spatial and temporal aspects of eye-movement control are interacting at many different levels. I will provide an overview on the viability of a likelihood-based approach for an underlying dynamical model.