MathSys Annual Retreat 2018

Programme and abstracts

Contents

Schedule

Tuesday 1 May

Time	Event	Location
08:00	Coach departs from campus	
12:30	Arrival at YHA Hawkshead	
12:30 - 13:30	Lunch	Dining room
13:30 - 14:00	Welcome and move in	
14:00 - 15:00	MSc poster session	Multiple
15:00 - 15:15	Coffee break	
15:15 - 15:45	6x 5 minute PhD talks (2 streams)	Main room, classroom
15:45 - 16:00	Coffee break	
16:00 - 16:30	6x 5 minute PhD talks (2 streams)	Main room, classroom
16:30 - 18:00	WARPs	
18:00 - 19:30	Dinner	Dining room
19:30 onward	Free time	

Wednesday 2 May

Time	Event	Location
7:30 - 9:00	Breakfast	Dining room
9:00 - 9:30	6x 5 minute PhD talks (2 streams)	Main room, classroom
9:30 - 9:40	Coffee	
9:40 - 10:00	4x 5 minute PhD talks (2 streams)	Main room, classroom
10:00 - 11:00	WARPs	
11:00 - 12:00	Workshop, choice of: 1. **Starting your PhD**- Gareth Alexander and Colm Connaughton 2. **Mid-PhD advice**- Federico Botta and Elizabeth Buckingham-Jeffery 3. **"So, you have your PhD, what now?"**- Martine Barons	 Lounge Main room Courtyard classroom
12:00 - 13:00	Lunch	Dining room
13:00 - 14:00	**Topology and dynamics in neural networks** Sam Johnson, Lecturer in Applied Mathematics, University of Birmingham	Main room

Wednesday 2 May (continued)

Time	Event	Location
14:00 - 19:00	Free time	
19:00 - 20:30	Dinner	Dining room
20:30 - 21:30	**Capacity of communication systems** Eddie Wilson, Chair in Intelligent Transport Systems/Head of Department, University of Bristol	Main room

Thursday 3 May

Time	Event	Location
7:30 - 9:00	Breakfast	Dining room
9:00 - 9:30	SSLC meeting	TBC
9:30 - 10:30	**The corridors of power: a complexists journey through the Civil Service** Peter Dawson, Senior Scientific Officer, UK Home Office	Main room
10:30 - 11:00	Coffee	
11:00 - 12:00	**Are Monte Carlo methods dead?** Chris Nemeth, Lecturer in Statistical Learning, University of Lancaster	Main room
12:00 - 13:00	Lunch	Dining room
13:00 - 14:00	Director's talk	Main room
14:00 - 14:30	Prize giving and group photo	Main room
14:30 - 18:00	WARPs	
18:00 - 19:30	Dinner	Dining room
19:30 onward	Free time	

Friday 4 May

Time	Event	Location
7:30 - 9:00	Breakfast	Dining room
9:00	Coach departs YHA Hawkshead	

PhD student talk schedule

Tuesday 1 May

Time	Stream 1 (main room)	Stream 2 (classroom)
15:15 - 15:20	Peter Byfield	Nada Jankovicova
15:20 - 15:25	Charlotte Roman	Ayman Boustati
15:25 - 15:30	Gian Lorenzo Spisso	Ian Atkinson
15:30 - 15:35	Henry Charlesworth	Bhavan Chahal
15:35 - 15:40	Aditi Shenvi	Iliana Peneva
15:40 - 15:45		
16:00 - 16:05	Connor White	Benjamin Atkins
16:05 - 16:10	Joe Hilton	Christopher Norman
16:10 - 16:15	Jessie Liu	Christopher Davis
16:15 - 16:20	Trystan Leng	Robert Gowers
16:20 - 16:25		Timothy Pollington
16:25 - 16:30		

Wednesday 2 May

Time	Stream 1 (main room)	Stream 2 (classroom)
09:00 - 09:05	Jevgenij Gamper	Joseph Pollard
09:05 - 09:10	Sophie Meakin	Guillem Mosquera Donate
09:10 - 09:15	Katherine Broadfoot	Jack Binysh
09:15 - 09:20	Giovanni Mizzi	Laura Guzman Rincon
09:20 - 09:25	Simon Graham	Sami Al-Izzi
09:25 - 09:30		
09:40 - 09:45	Arthur King	Roger Hill
09:45 - 09:50	Andrew Hilditch	Cameron Lack
09:50 - 09:55	Kieran Kalair	Alvaro Cabrejas Egea
09:55 - 10:00	Kutlwano Bashe	

Student abstracts (in alphabetical order)

Sami Al-Izzi

Fluid structure interactions

I will talk briefly about fluid structure interactions at low Reynolds number and how it relates to my research (plus possible demos).

Benjamin Atkins

Vote Ben for president!

How to win an election (without any of the necessary qualifications).

Ian Atkinson

Summary of the ARCHER MPI course

We all received an email from Heather forwarding the following ARCHER training course http://www.archer.ac.uk/training/courses/2018/01/OnlineMPI/index.php which ran through early term 2. This talk will be an introduction to the ARCHER training courses and a brief summary of the MPI course material.

Kutlwano Bashe

Network analysis of U.S. mutual fund investments

We explore how the structural properties underlying the network of investor portfolios and their assets are linked to the stability of the system using quarterly data from 2003 to 2016 of U.S. Mutual Fund Investment Holdings. In particular, we look at how the diversification of portfolios by investors as well as the similarities in investor portfolios has evolved over time especially before, during and after the 2008 Global Financial Crisis and the implications for regulators.

Jack Binysh

Maxwell's theory of solid angle and the construction of knotted vortex fields

When I look at an object (a book say) from different points in space, the object sweeps out some area on my retina, which changes depending on where I look from (face on, edge on etc) - this is the Solid Angle function for the object. It turns out that this function only depends on the boundary curve of the object (the rim of the book) and not on the surface spanning that boundary curve - Solid Angle is really a property of a curve.
I'll describe the properties of this function in the particular case the curve is knotted, with some historical detail - Gauss and Maxwell thought about this function in the context of magnetostatics. I may also discuss its application to the simulation of knotted vortices in fluids, superfluids etc. We'll see how much I actually manage.

Ayman Boustati

Artificial intelligence: utopia or dystopia

"AI is the new electricity" - Andrew Ng. "With AI we are summoning the demon" - Elon Musk. Which of these statements in more accurate? Maybe both! Or maybe neither! In this non-technical talk, I am going to examine the short, medium and long term impact that AI will bring about to society and whether it will create a better world for everyone or just doom and gloom.

Katherine Broadfoot

Measles and parameter estimation

I'll talk a little about measles and then focus on parameter estimation we've performed that is forming the bulk of my phd so far. Starting with some pretty basic maximum likelihood estimation and building through to MCMC. I'll also talk about how we've dealt with some missing data, and more generally some of the difficulties we've faced with finding ways to use the data usefully.

Peter Byfield

Non-stationary Gaussian processes

I will be talking about non-stationary Gaussian processes. Gaussian processes are powerful probabilistic models currently finding great use (and popularity) throughout the statistical and machine learning communities. I will provide some minimal background on the topic of simple Gaussian process regression and then move on to discuss the introduction of non-stationarity through careful kernel design. If time permits I will try to outline some approaches towards inference in such models and may discuss some connections/extensions with deep Gaussian processes.

Alvaro Cabrejas Egea

On cryptocurrency and self-sustaining Ponzi schemes

We all have heard ""blockchain"" and ""cryptocurrency"" thrown around as buzzwords recently, but what are they, how do they actually work and what can we do with them?
Here you will find a (very) brief introduction including examples of the most imaginative ""applications"", such as self-sustaining pyramid schemes.

Bhavan Chahal

Introduction to urvis

Mario Gutiérrez-Roig has been working on a python module for analysing urban visual landscapes, which will be released soon. I will give a short demonstration of what it can do (Doctor Who themed).

Henry Charlesworth

Future state maximization and collective motion

I will briefly introduce and discuss a model of collective motion we have been working on based on the principle of "future state maximization", whereby agents move in a way that maximizes the number of future states that are available to them in the future. The idea is that by maximally "keeping their options open" they are in some sense maintaining as much control over their future environment as they possibly can. We have found that an application of this principle to a group of agents with simple visual sensors leads to robust, highly ordered collective motion with features that are observed in real flocks of starlings.

Christopher Davis

Using mixture density networks to emulate complex models

Complex computationally-costly models are very common in mathematical biology. Often, these models need to be evaluated many times and if the simulation takes long to run then this severely limits model fitting and analysis.

We introduce the idea of using emulation, where the simulation is approximated with a statistical model that is typically much quicker to evaluate, through the implementation of mixture density networks.

Jevgenij Gamper

On importance of explicit assumptions, and benefits of generative thinking in medical image analysis - a case of digital histopathology

In this talk I will outline some of the common problems tackled in digital pathology. In particular, using a generative point of view we will attempt to obtain insights about the nature of problems at hand, and will speculate on the validity of the solutions and modelling approaches currently used in the field. An objective treatment of the assumptions behind the current state of the art solutions and their implications is of crucial importance, as these methods are starting to be widely adopted in clinical applications.

Robert Gowers

Neural correlations in space and time

Each neuron typically has thousands of synapses across its dendrites which can receive input over time. The arrival these background synaptic inputs can be treated as a stochastic process. However, the arrival time and position of synaptic inputs are typically not independent of each other. I will demonstrate that many correlations in space and time are not only analytically tractable, but can be implemented computationally.

Simon Graham

Towards a Bayesian approach for gland instance segmentation in colon histology images

The analysis of glandular morphology within colon histopathology images is a crucial step in determining the stage of colon cancer. Despite the importance of this task, manual segmentation is laborious, time-consuming and can suffer from subjectivity among pathologists. The rise of computational pathology has led to the development of automated methods for gland segmentation that aim to overcome the challenges of manual segmentation. However, current methods do not take into account a measure of uncertainty that is vital in clinical practice. For example, ambiguous areas may require further examination from numerous pathologists. In this talk, I will discuss how current fully convolutional neural networks for gland segmentation can be developed to incorporate uncertainty and how this uncertainty can be leveraged in routine pathological practice.

Laura Guzman Rincon

A surprise

A surprise, by Laura Guzman Rincon.

Andrew Hilditch

Data cleaning and the construction of the rail network

The rail network in Great Britain is one of the most complicated in the world. Some of the network still utilises Victorian infrastructure and is blended with modern technology. The network is mostly covered by berth movements that can be obtained from the signalling network. From these movements I have created a network model.

Roger Hill

Short term missions

I will explain what the short term scientific mission system is and how to apply for funding to visit and collaborate with other European institutions.

Joe Hilton

Household models for endemic diseases

Many developing countries are undergoing demographic transition, with birth rates decreasing and life expectancy increasing, changing the age profile and household structure of these countries. This has the potential to profoundly influence infection dynamics and the actions of disease control measures. In this talk I will present a model which extends household models to endemic settings by incorporating demographic processes, and will demonstrate the importance of this extension by comparing results for two different diseases in two different countries.

Nada Jankovicova

Bayesian statistics and electronic health records

In my talk I will be talking about Bayesian statistics and why use it in electronic health record (EHR) mining.

Kieran Kalair

The simplest macroscopic traffic model

The Lighthill-Whitham-Richards model is the simplest but most widely used macroscopic traffic model. I will discuss our work combining this model with a data-driven approach, incorporating stochastic elements present in data, and seeing how it performs on modern road networks. Depending on time, the talk will broadly cover aspects of macroscopic modelling, big-data analysis, stochastic processes and distribution fitting.

Arthur King

The wrong crowd

Discussing simple models of complex systems.

Cameron Lack

A web based interface for epidemiology models

Play testing a web based tool for presenting mathematical models to make them more accessible to non-specialists.

Trystan Leng

Approximating epidemics on networks

The spread of infectious diseases can be conceptualised as stochastic processes on networks. One approach would to be straightforwardly simulate the process, but doing so tells us little about the causal determinants of the dynamics of the process. An alternative and widely used method is that of moment-closure, where an analytically tractable set of ordinary differential equations is obtained by approximating the behaviour of higher-order structures (e.g. triples) in terms of lower-order structures (e.g. pairs). However, the assumptions underlying such closures, and the consequent reasons for errors in approximation, are often unclear. This talk introduces the concept of moment-closure for epidemics on networks, and discusses three possible sources of error: 1) disease dynamics, 2) network topology, and 3) local to global dynamics.

Zhangdaihong Liu

Statistically controlling for confounding constructs Is harder than you think

In this talk, I will be introducing the concept of confounds and the importance of controlling confounds in statistical analysis. I will also show the effects of confounds in the application of neuroimaging and behaviours.

Gian Lorenzo Spisso

Bounds of the value function via coupling

I'll describe an application of coupling techniques to bound the value function of a stochastic optimization process.

Sophie Meakin

Inferring interaction parameters from incidence data

How can we quantify the interaction between a pair of populations, such as spatially separated populations or different risk groups? The between-population interactions are complex and high-quality data on relevant interactions are often unavailable. In this talk I present a method that gets round this problem by estimating the interaction using widely-available data on disease incidence.

Giovanni Mizzi

Using Google and Twitter to make short term predictions of dengue case count in Rio de Janeiro

Dengue is a major threat to public health in Brazil, the world's sixth biggest country by population, with nearly 1.5 million cases recorded in 2016 alone and case counts continuing to grow. However, official data on the current number of dengue cases can often be severely delayed, with incremental delivery of data and a wait of up to six months for full case count information. Previous studies have sought to exploit rapidly available data on dengue-related Google searches or Twitter messages to deliver improved estimates of dengue incidence, but have not accounted for the true nature of the delays in dengue data across Brazil, creating barriers for the application of these approaches in practice. My previous research showed that it is possible to deliver improved weekly estimates of dengue incidence in Rio de Janeiro by drawing on Google and Twitter data together, whilst explicitly accounting for the structure of the delays in incoming dengue case count data. More recently we have been exploring how to use daily online data in order to generate short term forecasts. We find that it is indeed possible to produce estimates at the beginning of the week that are not significantly

different from those that we previously generated at the end of the week. Online data play a crucial role in this, and again better results are obtained when Google and Twitter data are used together.

Guillem Mosquera
A surprise
A surprise, by Guillem Mosquera.

Christopher Norman
Designing experimental protocols to improve model parameter identifiability
The healthy function of neuronal systems relies on the rapid signal transmission permitted by ion channels. These structures mediate the flow of ions across cell membranes in response to some stimulation (e.g. voltage) and their dysfunction can result in severe neurological conditions including epilepsy, ataxia, and migraine. Markov models have been used with success to model the fast dynamics of ion channels, however the estimation of model parameters from experimental data (e.g. voltage-clamp recordings) remains a challenge. Experimental protocols can however be structured to better distinguish model parameters, here a maximum likelihood approach is applied to inform experimental protocols which improve parameter identifiability.

Iliana Peneva
Using machine learning to understand cancer
Machine learning has found applications in many different areas: robotics (e.g. the work of Boston Dynamics), art (transforming a painting using DeepDream), finance (fraud detection), and entertainment (the recommendation systems in Amazon, Netflix and Spotify). One area in which machine learning has contributed to a lot of promising results is healthcare. It has helped identify biomarkers, improved the accuracy of diagnosis, and led to more efficient treatments.
In this talk, I will introduce the concept of latent variable models (LVM) and how they have been used in these discoveries. I will describe as well a data integration LVM, which can offer insight into aggressive diseases, and present results from its application to datasets from The Cancer Genome Atlas project.

Joseph Pollard
Singularity theory in physics
The mathematics of singularity theory has been applied to the study of various physical phenomena, including many problems in optics. Recent experiments on defects in cholesteric liquid crystals provide a realisation of all kinds of interesting singularities, including some that have never been seen in any other physical system. I'll talk about my work on understanding these defects, and also show lots of pretty pictures.

Timothy Pollington
How many samples?
When generating a sampling distribution from any function f, often the number of times we sample from it is adhoc. One million sounds about right, right? Inspired by a recent Mathematics Today article by Briggs & Ying, I trial their method with my own simulator that estimates the cases averted by a disease control programme to a pre-defined precision. This isn't an epi talk but I'm not a statistician either.

Charlotte Roman
Path finding
How robots move to avoid obstacles.

Aditi Shenvi
Model selection in chain event graphs

A Chain Event Graph (CEG) is a probabilistic modelling tool which was developed to deal with asymmetric state spaces. Context-specific independence relations can be read directly from the topology of a CEG. These kinds of relations are often found while modelling public health interventions and can be critical to the analysis of the intervention, thus making CEGs a strong choice for modelling such problems. I will briefly talk about the choice of CEGs over Bayesian Networks and the need for developing efficient model selection algorithms for CEGs.

Connor White
The major histocompatibility complex
Going over what the major histocompatibility complex (MHC) is and its role in the immune system. Why is the MHC interesting and what do we believe the selective pressures are on it.

External speakers (in order of appearance)

Federico Botta
Research Fellow in Data Science, Warwick Business School
Talk (joint with Elizabeth Buckingham-Jeffery): Mid-PhD advice
Aimed at first and second year PhD students

Elizabeth Buckingham-Jeffery
EPSRC PDRA in Applied Mathematics, University of Manchester
Talk (joint with Federico Botta): Mid-PhD advice
Aimed at first and second year PhD students

Martine Barons
Director of the Applied Statistics & Risk Unit, University of Warwick
Talk: "So, you have your PhD, what now?"
What are the routes into industry or academic work? What is a Postdoc and how useful is it for an industry career? How can I apply for a grant? This talk will cover a broad range of post-PhD career options.
Aimed at second and final year PhD students

Sam Johnson
Lecturer in Applied Mathematics, University of Birmingham
Topology and dynamics in neural networks
The brain is a paradigmatic complex system: individual neurons are relatively well understood, yet when even just a few hundred of them interact, collective behaviour emerges which we are only beginning to understand. I will discuss some recent results derived from simple mathematical models which highlight the crucial role that network topology can play in shaping neural phenomena, and muse on some broader implications.

Eddie Wilson
Chair in Intelligent Transport Systems/Head of Department, University of Bristol
Talk: Capacity of communication systems

Peter Dawson
Senior Scientific Officer, UK Home Office
Talk: The corridors of power: a complexists journey through the Civil Service
Analysis in the civil service and a shameless recruitment plug.

Chris Nemeth
Lecturer in Statistical Learning, University of Lancaster
Talk: Are Monte Carlo methods dead?
Monte Carlo methods, and in particular Markov chain Monte Carlo techniques, have been the gold standard computational tool for Bayesian modelling over the past 30 years. These algorithms can be applied in general settings, from identifying traits in phylogenetic trees to detecting Earth-like planets in distant solar systems, their supporting theoretical guarantees have led them to be widely

used by scientists and industry practitioners alike. However, a significant drawback is that traditional Monte Carlo algorithms scale poorly with large datasets, leading to a computational cost that grows at least proportionally with the data size. This leads to a prohibitive cost for modern-day machine learning and data science applications and has led practitioners towards scalable approximate alternatives, such as variational methods, which have no theoretical guarantees on the resulting approximation error. In this talk, I'll discuss some recent advances in scalable Monte Carlo methods which maintain the favourable theoretical properties of standard Monte Carlo methods, but which are generalisable.

Room allocations

Staff

Unit 1 - Francisco Rodrigues and family
Unit 2 - Robin Ball
Unit 3 - Colm Connaughton
Unit 4 - Gareth Alexander
Unit 5 - Eddie Wilson
Unit 6 - Sam Johnson
Unit 7 - Stefan Grosskinsky
Unit 8 - Martine Barons (Tuesday); Liz Buckingham-Jeffery (Wednesday, Thursday)
Room 9 (in main house) - Peter Dawson
Room 11 (in main house) - Fede Botta

Students (rooms in main house)

Room 2	Room 3	Room 4	Room 5
Gian Lorenzo Spisso	Alvaro Cabrejas Egea	Joseph Pollard	Kieran Kalair
Guillem Mosquera	Arthur King	Juan Ungredda	Kutlwano Bashe
Harvey Devereux	Benjamin Miller	Son Le	Monty Sharma
Henry Charlesworth	Giovanni Mizzi	Timothy Pollington	Robert Gowers
Steve Tseng	Jevgenij Gamper		Roger Hill
Ian Atkinson	Joe Hilton		Simon Graham
Jack Binysh			
Jonathan Skipp			

Room 6	Room 7	Room 8	Room 10
Alex Holmes	Andrew Hilditch	Aditi Shenvi	Charlotte Roman
Mohammad Noorbakhsh	Ayman Boustati	Aliya Jangabylova	Xiaoyue Xi
Peter Byfield	Cameron Lack	Annika Stechemesser	Zhangdaihong Liu
Sami Al-Izzi	Connah Johnson	Bhavan Chahal	
Samuel Forbes	Michael Luya	Duleabom An	
		Iliana Peneva	
		Katherine Broadfoot	

Room 12	Room 14
Benjamin Atkins	
Christopher Davis	Emma Southall
Christopher Norman	Laura Guzman Rincon
Edoardo Barp	Nada Jankovicova
Trystan Leng	Sophie Meakin
	Susanna Cant