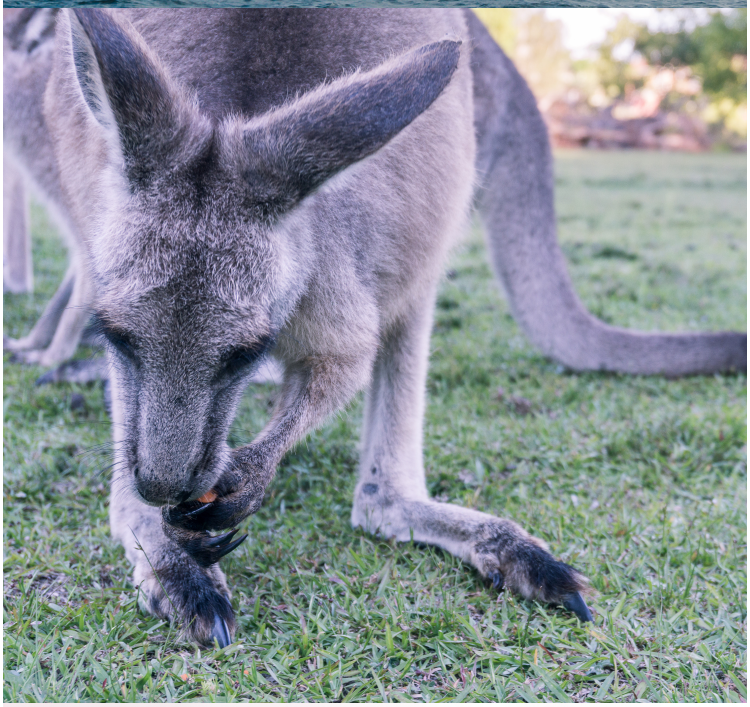
12th

INTERNATIONAL CONFERENCE ON MONTE CARLO METHODS AND APPLICATIONS



MCM 2019

CONFERENCE PROGRAM AND INFORMATION

JULY 8-12
SYDNEY
AUSTRALIA

Contents

Welcome to Sydney for MCM 2019

We are delighted to host you in Sydney for the *12th International Conference on Monte Carlo Methods and Applications*, MCM 2019. We hope that you will have a memorable week, enjoying both the scientific program and the beautiful coastal city of Sydney.

We wish to thank all the Program Committee members for organizing their proposed sessions. We are also very pleased with the lineup of plenary speakers and are thankful for their participation in making this conference a Monte Carlo highlight for 2019. We hope you will have the opportunity to both learn new ideas and build new collaborations and friendships.

Our scientific program features eight one-hour invited plenary talks. We also have approximately 140 invited and contributed talks of 30 minutes each (including questions). These talks are given in sessions of 1-2 hours consisting of 2-4 talks. We will have an award for the best student presentation.

Sydney is the capital of New South Wales and Australia's largest city with a population of around 5 million. It is best known for its harbour front featuring the iconic Sydney Opera House, which is one of the 20-th century's most famous and distinctive buildings. The harbour front also features Sydney Harbour bridge, a heritage-listed steel bridge that carries traffic between Sydney's central business district and the North Shore. The dramatic view of the bridge, the harbour, and the Sydney Opera House is an iconic image of Sydney and Australia itself. Sydney is also home to some of the most famous and iconic beaches in the world, such as Bondi beach and Manly beach. A stunning coastal walk from Bondi beach to Coogee beach and a ferry trip to Manly beach are some of the most scenic tourist attractions near the city. Other well-known tourist attractions are the Royal botanical gardens and the vibrant Darling Harbour, which is very close to the conference venue and has many restaurants.

We hope that you will enjoy your time in Sydney and have a productive and interesting conference. Please do not hesitate to contact any of the organizers if you have any problems or special requests during the conference.

Once again, a warm welcome to Sydney!

The Organizers,

Zdravko Botev, Shev MacNamara, Matias Quiroz, Quoc Thong Le Gia, Josef Dick, Libo Li, Dorota Toczydlowska, Christopher Poulton, Anthony Dooley, Mary Coupland, Nadima El-Hassan

The MCM Conference Series

The biennial *International Conference on Monte Carlo Methods and Applications* (MCM) (formerly *IMACS Seminar on Monte Carlo Methods*) is one of the most prominent conference series devoted to research on the mathematical aspects of stochastic simulation and Monte Carlo methods,

The conference is held on odd years and alternates with the International Conference on Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing (MCQMC) which is held on even years. These two conference series cover similar topics and their sets of participants have a significant intersection. Many participants of the highly successful MCQMC 2012 held in Sydney (http://www.mcqmc2012.unsw.edu.au/) will return for MCM 2019.

The locations of the 12 conferences are set out below.

Year	Location
1997	Brussels, Belgium
1999	Varna, Bulgaria
2001	Salzburg, Austria
2003	Berlin, Germany
2005	Tallahassee, FL USA
2007	Reading, UK
2009	Brussels, Belgium
2011	Borovets, Bulgaria
2013	Annecy-le-Vieux, France
2015	Linz, Austria
2017	Montréal, Canada
2019	Sydney, Australia
2021	???

Committees

Organizing Committee

Zdravko Botev (UNSW chair)	(UNSW, Australia)
Shev MacNamara (UTS chair)	(UTS, Australia)
Matias Quiroz	(UTS, Australia)
Quoc Thong Le Gia	(UNSW, Australia)
Josef Dick	(UNSW, Australia)
Libo Li	(UNSW, Australia)
Dorota Toczydlowska	(UTS, Australia)
Christopher Poulton	(UTS, Australia)
Anthony Dooley	(UTS, Australia)
Mary Coupland	(UTS, Australia)
Nadima El-Hassan	(UTS, Australia)

Steering Committee

Ronald Cools	(KU Leuven, Belgium)
Ivan Dimov	(Academy of Science, Bulgaria)
Thomas Müller-Gronbach	(Universität Passau, Germany)
Christian Lécot	(Université de Savoie, France)
Pierre L'Ecuyer (chair)	(University of MontrÃ©al, Canada)
Karl Sabelfeld	(Russian Academy of Science, Russia)
Wolfgang Schmid	(Universität Salzburg, Austria)
Paula Whitlock	(City University of New York, USA)

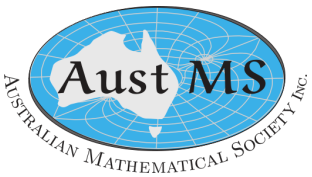
Program Committee

José Blanchet	(Stanford University, USA)
Nicolas Chopin	(ENSAE, France)
Jon Cockayne	(University of Warwick, UK)
Ronald Cools	(KU Leuven, Belgium)
Ivan Dimov	(Academy of Science, Bulgaria)
Christopher Drovandi	(Queensland U. of Technology)
Paul Dupuis	(Brown University, USA)
Stefan Geiss	(Jyväskylä University, Finland)
Mark Girolami	(Imperial College, UK)
Emmanuel Gobbet	(Ecole Polytechnique, France)
David Gunawan	(UNSW, Australia)
Stefan Heinrich	(U. Kaiserslautern, Germany)
Fred Hickernell	(IIT, USA)
Tony Karvonen	(Aalto University, Finland)
Robert Kohn	(UNSW, Australia)
Peter Krietzer	(JKU, Austria)
Frances Kuo	(UNSW, Australia)
Pierre L'Ecuyer	(University of MontrÃ©al)
Feng Li	(CUFE, China)
Christiane Lemieux	(U. of Waterloo, Canada)
Thomas Müller-Gronbach	(U. Passau, Germany)
Christopher Nemeth	(Lancaster University, UK)
Dirk Nuyens	(KU Leuven, Belgium)
Art Owen	(Stanford University, USA)
Raghu Pasupathy	(Purdue University, USA)
Gareth Peters	(Heriot-Watt University, UK)
Friedrich Pillichshammer	(JKU, Austria)
Nadhir Ben Rached	(RWTH Aachen University, Germany)
Matias Quiroz	(UTS, Australia)
Chang-Han Rhee	(Northwestern University, USA)
Ad Ridder	(Vrije University, Netherlands)
Christian Robert	(Universite Paris-Dauphine, France)
Wolfgang Schmid	(Universität Salzburg, Austria)
Ian Sloan	(UNSW, Australia)
Leah South	(Lancaster University, UK)
Raul Tempone	(KAUST, Saudia Arabia)
Dorota Toczydlowska	(UTS, Australia)
Gerardo Rubino	(INRIA Rennes–Bretagne Atlantique, France)
Liangliang Wang	(Simon Fraser University, Canada)
Henryk Wozniakowski	(Columbia U. and U. of Warsaw, USA/Poland)
Ming Xu	(Queensland U. of Technology, Australia)
Andrew Zammit-Mangion	(University of Wollongong, Australia)

Sponsors

We are very grateful for the financial support from the following sponsors

- University of Technology Sydney
- University of New South Wales
- The Australian Mathematical Society
- The Australian Mathematical Sciences Institute
- The Australian Research Council Center of Excellence for Mathematical & Statistical Frontiers
- The New South Wales Government



Practical information

Conference Venue

The 12-th International Conference on Monte Carlo Methods and Applications will take place in downtown Sydney, at the University of Technology Sydney Business School.

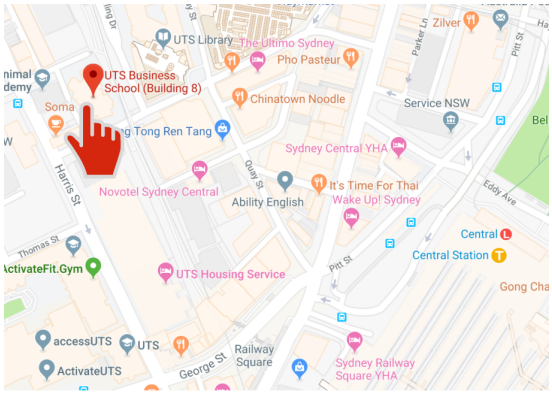
General maps showing the venues for the conference and banquet dinner are located on the back cover of this document. These can be accessed through Google maps, where you also easily can get the direction from your current location. We encourage the participants to use this feature and, if data roaming is a problem, please note that Google maps can be accessed in offline mode, see for example http://time.com/4203431/google-maps-use-offline-how-2/.

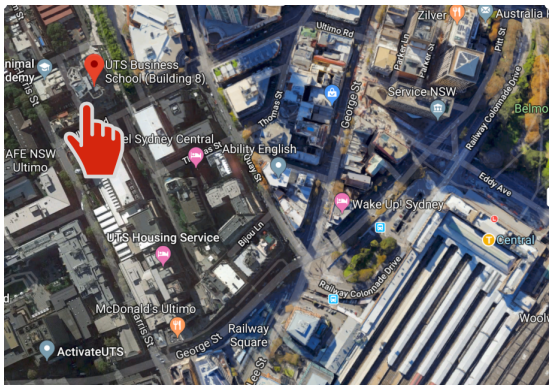
Main conference

The conference is held at *UTS Business School*. The exact address is **UTS Business School (Building 8), 14-28 Ultimo Rd, Ultimo NSW 2007**.

The conference venue is next to the Novotel Sydney Central. The conference venue is a short walk from both Sydney Central (a major hub for trains) and Railway Square (a major hub for buses).

The following maps show the location of the conference venue (see also back of this book).





We do not recommend coming by car. Although there are plenty of parking spots nearby, they tend to be very expensive. The conference venue is centrally located and is therefore easily accessed by public transport.

Registration, information desk, coffee breaks, and conference rooms

The registration and information desk will be located at the conference venue. There will be signs that guide you to the registration location. Registration opens at 8:00 am on Monday morning. The coffee and lunch breaks will take place right next to the seminar and plenary (CB08.03.005) rooms. The parallel sessions will be in seminar room CB08.02.002, room CB08.03.004, room CB08.03.002.

Coffee/Breakfast/Lunch breaks

The conference will provide coffee and lunch/sandwiches right next to the plenary/seminar rooms. The lunch breaks will usually last 1 hour. If you can wish to have a more eclectic experience, there are plenty of lunch options in proximity to the conference venue. For example, you can try the trendy Cáfe Soma: https://goo.gl/maps/2qLvzisr5EtTxK39A.

You can find places to eat in nearby Chinatown here https://www.theurbanlist.com/sydney/a-list/where-to-eat-in-chinatown-sydney. Please note that some of these places might require a 15-20 minute walk. We recommend that you use Google maps for suggestions.

Wireless

If your home institution has *eduroam* then you will be able to connect freely to the internet at the conference (wherever there is wifi available). Please note that you need to set up *eduroam* at your home institution before you will be able to use it on the UTS campus.

There will be a wifi network available for the conference, with name Monte Carlo Conference, and password GodDoesNotPlayDice.

Links for tourist information

- The government of New South Wales provides tourist information https://www.cityofsydney.nsw.gov.au/learn/about-sydney/tourist-information.
- The lonely planet https://www.lonelyplanet.com/australia/sydney.
- Search for Sydney at Tripadvisor https://www.tripadvisor.com.au/.

Public transportation

Detailed information on Sydney's public transportation system can be found at https://transportnsw.info/. Note that an Opal card is required to use the public transportation system. This includes buses, trains and ferries. The Opal card may be purchased and charged at the airport or at Sydney Central. You can also do this in most convenience stores. The card is charged at a minimum of 10 Australian dollars (AUD). A single trip within the city limits is approximately 2.20-3.60 AUD. On Sundays, the maximum amount your card can be charged is 2.70 AUD, regardless of how many trips you do - a perfect opportunity to take the ferry and explore Manly beach!

Taxi is an alternative form of transportation. Uber operates in Sydney at favorable rates. To use Uber you need to install the app, see https://uber.com/ for more information.

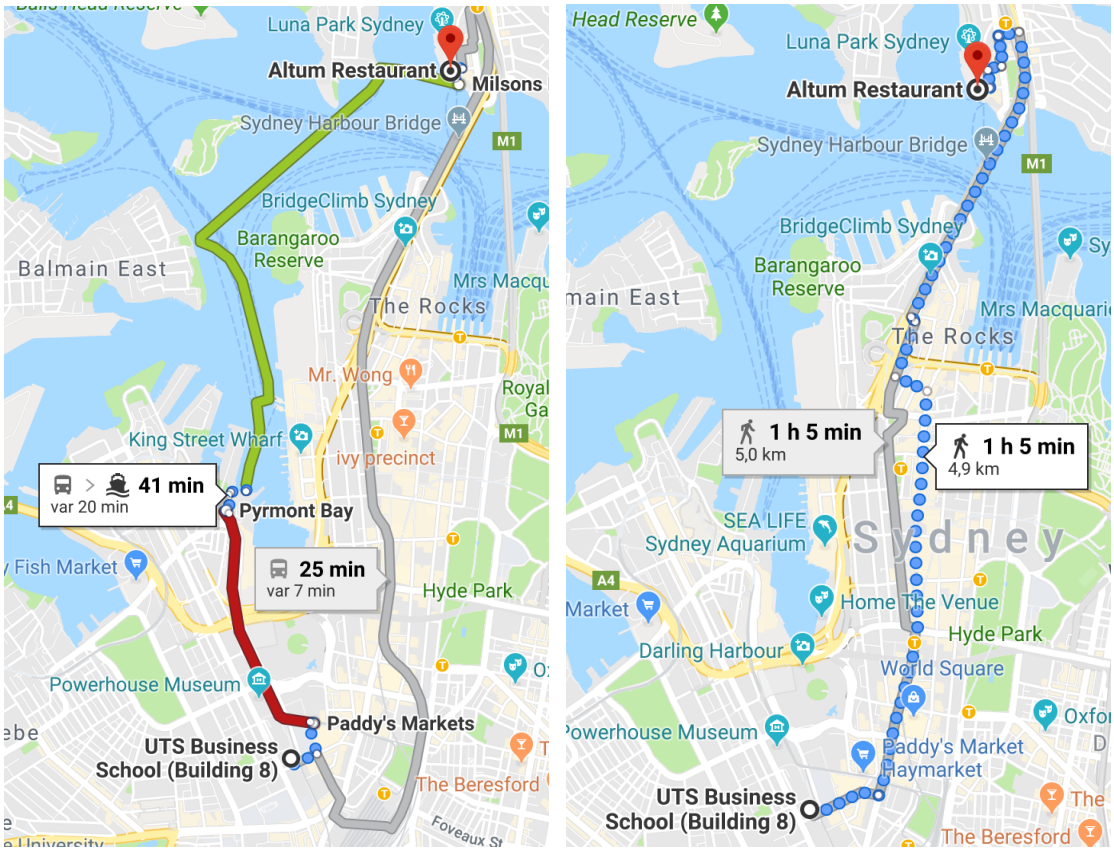
Conference Banquet Dinner

The conference banquet will be held 6:00pm on Wednesday evening (July 10th) at:

Altum Restaurant
1 Olympic Drive,
Milsons Point
NSW 2061

This location has a fantastic view over the harbour towards the Harbour Bridge and the Opera House. It is worthwhile arriving a little early to go for a short walk along the harbour side. See the restaurant website: www.altumrestaurant.com

Here are two different ways to get from the conference location to the restaurant:



The option to take the ferry (left map in the picture above) to the restaurant is recommended if you do not wish to walk for a long time. You need an OPAL card to catch the ferry, as with all public transport in Sydney. Opal Cards can be bought from most convenience shops, including the 7-Eleven on the corner of Quay St. and Ultimo Road (diagonally opposite the UTS Library, very close to the conference location). The Ferry Trip will be less than $10 each way. You can also pay by 'tapping on' with a Credit Card, or by buying a Single Trip Opal Ticket at the Ferry Wharf. You can catch the ferry from Pyrmont Bay ferry wharf, which is about 20 minutes walk from the conference location. Get off the Ferry at "Milsons Point / Luna Park." The restaurant is very close to that ferry stop, no more than 2 minutes walk.

Note that the choice to participate in the conference dinner was made when you registered. Only those who chose this option when they registered can attend.

Presentations

Best student presentation award

There will be an award for the best presentation by a student. Those who registered as a student will be eligible. The winner will be presented at the end of the conference during the closing remarks, on Friday.

Instructions for speakers

Plenary talks are 45-50 minutes plus 10-15 minutes for questions and discussion. All other talks are 20 minutes, plus 5 minutes for questions and discussion and 5 minutes to allow people to move between sessions. We ask the speakers to please make sure that you do not exceed your time. Focus on the essential of your message. Given the short time allowed to each speaker, it is generally not possible to give the full details of your work. You should concentrate on providing a clear explanation of your main results and their significance.

All speakers: please upload your presentation, before the session begins.

Instructions for session chairs

Session chairs have the responsibility to make sure the speakers adhere tightly to the schedule. The session chair introduces the speaker briefly. Some participants might want to switch between parallel sessions to attend specific talks. To make sure that this can be done smoothly, session chairs should enforce strict adherence to the schedule. We recommend that the chair signals the speaker indicating there are 10, 5 and 1 minutes of speaking time left.

Session chairs should also contact their session speakers ahead of time to verify their presence and inform the organizers of any potential no-shows. In case a speaker that is not the last of the session does not show up, please leave that slot empty and wait for the announced time in the program before allowing the next speaker.

It will help to keep to time if you encourage all speakers to upload their presentation, before the session begins.

Equipment

You can bring and use your own laptop for your presentation. VGA is NOT available. All computers have HDMI. We recommend for any MAC users to bring their own equipment and adaptors. All lecture rooms are equipped with a computer and a projector for displaying computer output. The computer in the room has a USB port, so you can simply upload your presentation, as either PDF or as power point. All AV lecterns are the same in all lecture theatres and classrooms and the instructions on how to use the lecterns are glued on the desk.

All speakers please upload your presentation before the session begins It is preferable if speakers use the computer already provided in the room. This helps to keep the session to time, because then there is no time wasted in the changeover between speakers.

Schedule

The abstracts can be found at the page numbers shown in red.

Monday morning, July 8

08^{00} **Registration opens**

$9^{00} - 9^{05}$ **Opening Remarks** Room: CB08.03.005
Zdravko Botev

$9^{05} - 10^{00}$ **Invited Plenary Talk** Room: CB08.03.005
Scott Sisson
The future of Monte Carlo methods research: a random walk?
Chair: *Matias Quiroz* *p. 24*

$10^{00} - 10^{30}$ Coffee break

	Room: CB08.02.002	Room: CB08.03.004	Room: CB08.03.002	Room: CB08.03.005
	Advances in sequential Monte Carlo methods: Theory, algorithms and applications	**Advances in Rare-event Simulation**	**Current Challenges in High-Dimensional Algorithms**	**Advances in Exact and approximate Bayesian computation**
	Chair: *Christopher Drovandi* *p. 34*	Chair: *Victor Elvira* *p. 36*	Chair: *Peter Kritzer* *p. 38*	Chair: *Robert Kohn* *p. 41*
$10^{30} - 11^{00}$	*Giorgos Vasdekis* On the Exponential Ergodicity of the ZigZag Process	*Chang-Han Rhee* Importance Sampling for Rare Catastrophic Events	*Fred Hickernell* An Optimal Adaptive Algorithm Based on the Decay Rate of the Series Coefficients of the Input Functions	*Matias Quiroz* The block-Poisson estimator for optimally tuned signed pseudo-marginal MCMC
$11^{00} - 11^{30}$	*Anthony Lee* Likelihood approximations for latent variable models: computational and statistical scalability	*Thomas Taimre* Exploiting Asymptotics and Polar Coordinates for Rare Tail Estimation	*Fred J. Hickernell* An Optimal Adaptive Algorithm Based on a Pilot Sample	*Imke Botha* Bayesian Parameter Inference for Stochastic Differential Equation Mixed Effects Models
$11^{30} - 12^{00}$	*Leah F. South* Unbiased and Consistent Nested Sampling via Sequential Monte Carlo	*Victor Elvira* Importance sampling for signal processing applications	*Josef Dick* Numerical methods for partial differential equations with random coefficients	*Chun Fung Kwok* Introducing high-dimensional sensitivity analysis of Bayesian MCMC inference
$12^{00} - 12^{30}$	*Axel Finke* Limit theorems for sequential MCMC methods		*Dirk Nuyens* Approximation in Chebyshev space using lattice rules	*Minh-Ngoc Tran* A long short-term memory stochastic volatility model

$12^{30} - 13^{30}$ Lunch break

Monday afternoon, July 8

$13^{30}-14^{30}$ **Invited Plenary Talk** Room: CB08.03.005

Kerrie Mengersen

The Challenges Discoveries and Examples of ABC

Chair: *Christopher Drovandi* *p. 25*

$14^{30}-15^{00}$ Coffee break

	Room: CB08.02.002 **Advanced Monte Carlo in risk management** Chair: *Emmanuel Gobet* *p. 44*	Room: CB08.03.004 **Stochastic Computation and Complexity I** Chair: *Stefan Heinrich* *p. 46*	Room: CB08.03.005 **Efficient computations for Bayesian inference using Monte Carlo methods** Chair: *Matias Quiroz* *p. 49*	Room: CB08.03.002 **Importance Sampling for Robust Model Simulation** Chair: *Ad Ridder* *p. 51*
$15^{00}-15^{30}$	*Florian Bourgey* Multilevel Monte-Carlo method and lower/upper bounds in Initial Margin computations	*Andreas Neuenkirch* Smooth dependence of fractional Brownian motion on its Hurst parameter and the limits of rough paths theory	*Robert Kohn* Efficiently Combining Pseudo Marginal and Particle Gibbs Sampling	*Henry Lam* Safety evaluation of black-box prediction models via rare-event simulation
$15^{30}-16^{00}$	*Pierre Del Moral* Particle methods in risk analysis	*Paweł Przybyłowicz* Randomized Euler scheme for strong approximation of solutions of SDEs with time-irregularities	*Christopher Drovandi* Robust Approximate Bayesian Inference with Synthetic Likelihood	*Chang-Han Rhee* Nonlinear space-filling design and non-parametric importance sampling
$16^{00}-16^{30}$	*Benjamin Jourdain* Weak error analysis for some mean-field SDEs	*Winfried Sickel* Approximation Numbers of Embeddings of Anisotropic Sobolev Spaces of Dominating Mixed Smoothness	*Gael Martin* Focused Bayesian Prediction	*Ad Ridder* Variance reduction of quantile sensitivities
$16^{30}-17^{00}$	*Grangereau Maxime* McKean stochastic optimal control of an energy storage system to reduce demand variability	*Libo Li* Positivity preserving schemes for alpha-CIR process	*Daniel Simpson* Pictures and Fear: Some thoughts on convergence	*Richard Gerlach* Semi-parametric Dynamic Asymmetric Laplace Models for Tail Risk Forecasting, Incorporating Realized Measures

Tuesday morning, July 9

$9^{00} - 10^{00}$ **Invited Plenary Talk** Room: CB08.03.005

Takashi Goda

Richardson extrapolation and higher order QMC

Chair: *Fred Hickernell* *p. 26*

$10^{00} - 10^{30}$ Coffee break

	Room: CB08.03.005 **Stochastic Computation and Complexity II** Chair: *Sotirios Sabanis* *p. 53*	Room: CB08.02.002 **Monte Carlo Methods for Spatial Problems** Chair: *Shev MacNamara* *p. 55*	Room: CB08.03.004 **Simulation Optimization** Chair: *Raghu Pasupathy* *p. 58*	Room: CB08.03.002 **QMC Theory and Applications I** Chair: *Vesa Kaarnioja* *p. 60*
$10^{30} - 11^{00}$	*Larisa Yaroslavtseva* On the performance of the Euler-Maruyama scheme for SDEs with discontinuous drift coefficient	*Pavel Krupskiy* Spatial Cauchy Processes with Local Tail Dependence	*Jing Dong* Confidence Region for Model Parameters in Stochastic Gradient Descent via Batch Means	*Philipp Guth* A Quasi-Monte Carlo Method for an Optimal Control Problem
$11^{00} - 11^{30}$	*Thomas Müller-Gronbach* A strong order 3/4 method for SDEs with discontinuous drift coefficient	*Susanna Cramb* Computing the Australian Cancer Atlas: Getting it 'just right'	*José Blanchet* Efficient Iterative Schemes for Optimal Transport Distributionally Robust Optimization	*Hiroki Kajiura* An analogue of QMC integration over a finite group: difference sets and its generalization
$11^{30} - 12^{00}$	*Michaela Szölgyenyi* Convergence order of the Euler-Maruyama scheme in dependence of the Sobolev regularity of the drift	*Andrew Zammit-Mangion* Multi-Scale Process Modelling of Massive Spatial Datasets	*Harsha Honnappa* Risk-sensitive Inference	*Zhijian He* On the Use of Quasi-Monte Carlo in Quantile Estimation
$12^{00} - 12^{30}$	*Paweł Morkisz* Derivative-free Mistein scheme for strong approximation of solutions of SDEs in analytic noise model	*Matt Moores* Bayesian Indirect Likelihood for the Potts Model	*Raghu Pasupathy* Retrospective approximation for Stochastic Optimization	*Hozumi Morohosi* Simulating empirical copulas: Empirical comparison MC and QMC

$12^{30} - 13^{30}$ Lunch break

Tuesday afternoon, July 9

$13^{30}-14^{30}$	**Invited Plenary Talk** Room: CB08.03.005 *Peter Kritzer* **Construction Algorithms for (Polynomial) Lattice Points** Chair: *Ian Sloan* *p. 27*			
$14^{30}-15^{00}$	Coffee break			
	Room: CB08.02.002 **Monte Carlo Algorithms in Bayesian Statistics** Chair: *Ziwen An* *p. 63*	Room: CB08.03.002 **QMC Theory and Applications II** Chair: *Ian Sloan* *p. 66*	Room: CB08.03.004 **Applied Probability and Monte Carlo Methods I** Chair: *Chang-Han Rhee* *p. 68*	Room: CB08.03.005 **Stochastic Computation and Complexity III** Chair: *Thomas Müller-Gronbach* *p. 70*
$15^{00}-15^{30}$	*Dan Li* Efficient Bayesian Estimation for GARCH-type Models via Sequential Monte Carlo	*Jordan Van* A Numerical Study of the $R_d(\phi_d)$ Sequence	*Jose Blanchet* Wasserstein-based Distributionally Robust Performance Analysis with Wasserstein Martingale Constraints	*Noufel Frikha* Integration by parts formulae for killed processes and their unbiased Monte Carlo simulation.
$15^{30}-16^{00}$	*Krys Latuszynski* Barkers algorithm for Bayesian inference in intractable likelihood models	*Frances Kuo* Uncertainty Quantification using Periodic Random Variables and Lattice Quasi-Monte Carlo Rules	*Henry Lam* Validating Optimization with Probabilistic Constraints	*Stefan Heinrich* Complexity of stochastic integration in fractional Sobolev classes
$16^{00}-16^{30}$	*Khue-Dung Dang* Subsampling Sequential Monte Carlo for Static Bayesian Models	*Vesa Kaarnioja* Uncertainty Quantification for Capacitors with Uncertain Boundaries Using QMC	*Ad Ridder* On Counting Distributions for Mortality Projections	*Francisco Bernal* The role of the Eikonal equation in numerics for bounded SDEs
$16^{30}-17^{00}$	*Ziwen An* Robust Bayesian Synthetic Likelihood via a Semi-Parametric Approach	*Ian Sloan* Kernel-Based Lattice Point Interpolation for Uncertainty Quantification using Periodic Random Variables	*Ajay Jasra* Central Limit Theorems for Coupled Particle Filters	*David Krieg* How good is random information?

Wednesday morning, July 10

$9^{00} - 10^{00}$ **Invited Plenary Talk** Room: CB08.03.005

Arnulf Jentzen

Overcoming the curse of dimensionality: from nonlinear Monte Carlo to deep artificial neural networks

Chair: *Josef Dick* *p. 28*

$10^{00} - 10^{30}$ Coffee break

	Room: CB08.03.002	Room: CB08.03.005	Room: CB08.03.004	Room: CB08.02.002
	Stochastic Computation and Complexity IV	**Recent Advances in Monte Carlo Methods for Robustness, Long Memory Processes and Rough Differential Equations**	**Monte Carlo Estimation: Miscellaneous Applications**	**Scalable Bayesian Inference for Complex Statistical Models**
	Chair: *Larisa Yaroslavtseva* *p. 72*	Chair: *José Blanchet* *p. 75*	Chair: *Spiro Penev* *p. 77*	Chair: *Liangliang Wang* *p. 80*
$10^{30} - 11^{00}$	*Steffen Dereich* CLTs for stochastic approximation schemes under nonstandard assumptions	*Xinyun Chen* Perfect Sampling for Queues with Autoregressive Arrivals	*Willem D. Schutte* Choosing the Tuning Parameter for Goodness-of-Fit Tests that Employ Bootstrapped Critical Values by Using Estimated Sizes	*Mathieu Fourment* Effective Online Bayesian Phylogenetics Via Sequential Monte Carlo with Guided Proposals
$11^{00} - 11^{30}$	*Emmanuel Gobet* Central limit theorem for discretization errors based on stopping time sampling	*Henry Lam* Enhanced Balancing of Bias-Variance Tradeoff in Stochastic Simulation	*Anil Dolgun* Assessing Weighted Inter-rater Agreement Under Missing Data	*Saifuddin Syed* Parallel tempering: Scaling limits, and optimality
$11^{30} - 12^{00}$	*Sotirios Sabanis* SGLD with dependent data streams: Convex and Non-Convex case	*Jing Dong* ϵ-Strong Simulation of Fractional Brownian Motion and Related Stochastic Differential Equations	*Spiridon Penev* Robust index tracking	*Ming Xu* Variance Reduction Properties of the Reparameterization Trick
$12^{00} - 12^{30}$	*Timo Welti* Deep Artificial Neural Networks Overcome the Curse of Dimensionality in PDE Approximation	*Daniel Roth* Convergence of one-step survival Brownian bridge Monte Carlo for barrier options	*Natalya Tracheva* Projection Monte Carlo Methods For Estimation of the Bidirectional Scattering-Surface Reflectance Distribution Function	*Liangliang Wang* An Annealed Sequential Monte Carlo Method for Bayesian Phylogenetics

$12^{30} - 13^{30}$ Lunch break

Wednesday afternoon, July 10

	Room: CB08.02.002 **Probabilistic Numerical Methods** Chair: *Toni Karvonen* *p. 83*	Room: CB08.03.002 **Approximate Bayesian Computation** Chair: *Grégoire Claré* *p. 85*	Room: CB08.03.004 **Advances in Variational Inference** Chair: *Ming Xu* *p. 86*	Room: CB08.03.005 **Applied Probability and Monte Carlo methods II** Chair: *Chang-Han Rhee* *p. 88*
$13^{30}-14^{00}$	*Motonobu Kanagawa* Convergence rates of adaptive Bayesian quadrature	*Julien Stoehr* An attempt to make ABC cheaper	*Luca Maestrini* Streamlined Variational Inference for Inverse Problems Models	*Anand Deo* Limiting distributional fixed points in large banking networks
$14^{00}-14^{30}$	*Sho Sonoda* Numerical Integration Method for Training Neural Networks	*Anthony Ebert* Approximate Bayesian computation to model passenger flow within airport terminals	*Nghia Nguyen* Stochastic Variational Bayes with Particle Filter	*Raghu Pasupathy* Distributed Stochastic Optimization
$14^{30}-15^{00}$	*Toni Karvonen* Methods for Large-Scale and High-Dimensional Probabilistic Integration	*Grégoire Claré* ABC within Gibbs sampling	*Alex Cooper* Variational forecasts of observation-driven models	*Alessandro Zocca* The Skipping Random Walk Metropolis Sampler
$15^{00}-15^{15}$	group photo of all participants			
$18^{00}-21^{00}$	Gala dinner at Altum Restaurant, nearby to Luna Park (registration required for all except plenary speakers)			

Thursday morning, July 11

$9^{00} - 10^{00}$ **Invited Plenary Talk** Room: CB08.03.005

Gersende Fort

Monte Carlo methods and Optimization: Intertwinings

Chair: *Emmanuel Gobet* *p. 29*

$10^{00} - 10^{30}$ Coffee break

	Room: CB08.02.002	Room: CB08.03.002	Room: CB08.03.004	Room: CB08.03.005
	Multilevel Monte Carlo methods	**Monte Carlo Methods for Large Dependent Data**	**Diffusion Processes: Theory and Applications**	**Monte Carlo Methods in Environmental Applications**
	Chair: *Tomohiko Hironaka* *p. 90*	Chair: *Feng Li* *p. 93*	Chair: *Alexander Shkolnik* *p. 96*	Chair: *Andrew Zammit-Mangion* *p. 98*
$10^{30} - 11^{00}$	*Sarat Babu Moka* Unbiased Estimation of the Reciprocal Mean for non-negative Random Variables	*Clara Grazian* Approximate Bayesian Conditional Copula	*Sumei Zhang* Efficient simulation of a two-factor stochastic volatility jump-diffusion model with non-zero correlation between variance factors	*Michael Bertolacci* Climate inference on Australian daily rainfall using distributed MCMC
$11^{00} - 11^{30}$	*David Warne* Multilevel Monte Carlo for simulation and inference of biochemical reaction networks	*Rubén Loaiza-Maya* Variational Bayes Estimation of Discrete-Margined Copula Models with Application to Time Series	*Celia García-Pareja* Exact Simulation of Coupled Wright-Fisher Diffusions	*Petra Kuhnert* Vision and Decision: creating actionable insights from Bayesian Hierarchical Models for the Great Barrier Reef
$11^{30} - 12^{00}$	*Andreas Van Barel* Multilevel quasi Monte Carlo for optimization of PDEs	*Yanfei Kang* Predicting forecasting algorithm performance with diverse time series features	*Nicolas Massin* Simulation of the time needed by a diffusion process in order to exit from a given interval (WOMS algorithm).	*Noel Cressie* False Discovery Rates to Detect Signals From Incomplete Spatially Aggregated Data
$12^{00} - 12^{30}$	*Tomohiko Hironaka* Multilevel Monte Carlo Methods for Estimating the Expected Value of Sample Information	*Feng Li* Bayesian high-dimensional covariate-dependent copula modeling with application to stocks and text sentiments	*Alexander Shkolnik* Unbiased Simulation of Multivariate Jump-Diffusions	*Laura Cartwright* Bayesian Atmospheric Tomography for Detection and Estimation of Methane Emissions

$12^{30} - 13^{30}$ Lunch break

Thursday afternoon, July 11

$13^{30}-14^{30}$ **Invited Plenary Talk** Room: CB08.03.005

Bruno Tuffin

Rare-Event Simulation of Regenerative Systems: Estimation of the Mean and Distribution of Hitting Times

Chair: *Pierre L'Ecuyer* *p. 30*

$14^{30}-14^{35}$ **MCM2021** Preview and Invitation

$14^{35}-15^{00}$ Coffee break

	Room: CB08.03.002	Room: CB08.03.004	Room: CB08.02.002	Room: CB08.03.005
	Recent advances on Bayesian computation	**Random Number Generation**	**Advances in Scalable Monte Carlo**	**Applied Probability and Monte Carlo methods III**
	Chair: *David Gunawan* *p. 101*	Chair: *Pierre L'Ecuyer* *p. 103*	Chair: *Leah South* *p. 106*	Chair: *Thomas Taimre* *p. 109*
$15^{00}-15^{30}$	*Michael Stanley Smith* High-dimensional Copula Variational Approximation through Transformation	*Hiroshi Haramoto* Study on upper bounds of sample sizes for two-level tests in NIST SP800-22	*Cheng Li* Distributed Bayesian Inference for Varying Coefficient Spatiotemporal Models	*Matthew Varble* Estimating Tails of IID Sums without the MGF
$15^{30}-16^{00}$	*Reza Hajargasht* Approximation Properties of Variational Bayes for Vector Autoregressions	*Makoto Matsumoto* Tiny Mersenne Twister with Parameterization	*Joris Bierkens* Dimensional Scaling of Piecewise Deterministic Monte Carlo	*Nicholas Johnson* Efficient simulation of tail probabilities for subexponential sums with dependent random weights
$16^{00}-16^{30}$	*Robert Salomone* Implicit Langevin Algorithms for Sampling From Log-concave Densities	*Marc-Antoine Savard* A Software Tool to Analyze the Lattice Structure of RNGs	*Andi Wang* Quasi-stationary Monte Carlo Methods	*José Blanchet* Efficient Rare Event for Heavy-tailed Chance Constrained Optimization
$16^{30}-17^{00}$	*David Gunawan* Flexible Density Tempering Approaches for State Space Models with an Application to Factor Stochastic Volatility Models	*Asaki Saito* Generating High-Quality Pseudorandom Sequences Using Chaotic True Orbits on Algebraic Integers	*Joshua J Bon* Scalable Samplers for High-Dimensional Models with Stochastic Nets	*Emmanuel Gobet* Uncertainty Quantification for Stochastic Approximation Limits and applications to risk/performance metrics in finance

$18^{00}-21^{00}$ Business dinner for members of steering committee

Friday morning, July 12

	Room: CB08.02.002 **Nets, Sequences, and Discrepancy** Chair: *Wolfgang Schmid* *p. 111*	Room: CB08.03.004 **Network Reliability and the Splitting Method** Chair: *Gerardo Rubino* *p. 113*	Room: CB08.03.002 **Methods for Acceleration and Variance Reduction** Chair: *Florian Puchhammer* *p. 115*	Room: CB08.03.005 **Monte Carlo methods in change-point detection** Chair: *Georgy Sofronov* *p. 117*
$09^{00}-09^{30}$	*Tetiana Stepaniuk* Hyperuniform point sets on flat tori	*Javiera Barrera* Limit distributions of the upper order statistics for the Lévy-frailty Marshall-Olkin distribution	*Khai Xiang Au* Constrained Hamiltonian Monte Carlo for PDE Inverse Problems	*Nishanthi Raveendran* Monte Carlo Methods for Spatial Clustering
$09^{30}-10^{00}$	*Kosuke Suzuki* Characterization of Digital $(0,2)$-Sequences in Base 2 and Its Application	*Gerardo Rubino* Splitting–based method for network reliability estimation	*Yu Guang Wang* Fast Haar Convolution for Geometric Deep Learning	*Lijing Ma* A Monte Carlo Method for Multiple Change-Point Detection in a Segmented ARMA Model
$10^{00}-10^{30}$	*Jaspar Wiart* On The Dependence Structure Of Scrambled (t,m,s)-Nets	*Daniel Mackinlay* The Multilevel Splitting Method for Static Problems: Applications to Wireless Communications Systems Performances	*Florian Puchhammer* Variance Reduction for Chemical Reaction Networks with Array-RQMC	*Irene Hudson* Change-Point Detection via the Cross-Entropy Method
$10^{30}-11^{00}$	Coffee break			
$11^{00}-12^{00}$	**Invited Plenary Talk** Room: CB08.03.005 *Youssef Marzouk* **Sampling via Transport: Preconditioning and Low-Dimensional Structure** Chair: *Frances Kuo* *p. 31*			
$12^{00}-12^{10}$	**Closing Remarks** Room: CB08.03.005 *Zdravko Botev*			

Abstracts of Plenary Talks

Monday 9:00–10:00 **Room: CB08.03.005**

The future of Monte Carlo methods research: a random walk?

Scott Sisson
UNSW Sydney, Australia, scott.sisson@unsw.edu.au

Theory and methods research for Monte Carlo techniques has come a long way in the past 25 years. For those not willing to wait another 25 years to see what happens next, it is reasonable to ask what the drivers of Monte Carlo research will be in the coming decades, so that, should we choose to, we can move to those areas ahead of time, either individually or as a discipline. This talk will explore some past and current research trends, from the biased perspective of my own research, and then unconvincingly speculate about what we may be working on in the next 25 years.

Monday 13:30–14:30 **Room: CB08.03.005**

The Challenges Discoveries and Examples of ABC

Kerrie Mengersen
Queensland University of Technology, Australia, k.mengersen@qut.edu.au

Approximate Bayesian Computation (ABC) has become an established computational approach to Bayesian analysis. The past two decades has seen an explosion in theoretical and methodological research on ABC approaches, as well as a wealth of substantive applications that have employed ABC to gain important insights into complex problems. In this presentation, I will reprise the state of the art in ABC, touch on ongoing challenges and highlight some recent discoveries. In doing so, I will describe some specific examples to illustrate ways in which ABC has been extended to address novel questions in Bayesian statistics. This discussion of the Challenges, Discoveries and Examples (CDE) of ABC will be embedded in a number of practical case studies. A number of authors and colleagues are responsible for the research described in this presentation and will be acknowledged during the presentation.

[1] Chen, C.C.M., Drovandi, C.C., Keith, J.M., Anthony, K., Caley, M.J., Mengersen, K. (2017) *Bayesian semi-individual based model with approximate Bayesian computation for parameter calibration: modelling crown-of-thorns populations on the Great Barrier Reef.* Ecological Modelling 364, 113-123.

[2] Clarte, G., Robert, C.P., Ryder, R., Stoehr, J. (2019) *Component-wise approximate Bayesian computation via Gibbs-like steps.* arXiv preprint arXiv:1905.135.99

[3] Drovandi, C.C., Grazian, C., Mengersen, K., Robert, C. (2018) Approximating the likelihood in ABC. Handbook of Approximate Bayesian Computation, 321-368.

[4] Ebert, A., Dutta, R., Mengersen, K., Mira, A., Ruggeri, F., Wu, P. (2018) *Likelihood-free parameter estimation for dynamic queuing networks: case study of passenger flow in an international airport terminal.* arXiv preprint arXiv:1804.02526

[5] Nott, D.J., Drovandi, C.C., Mengersen, K., Evans, M. (2018) *Approximation of Bayesian predictive p-values with regression ABC.* Bayesian Analysis 13(1), 59-83.

[6] Pudlo, P. Marin, J.-M., Estoup, A., Cornuet, J.-M., Gautier, M., Robert, C.P. (2016) *Reliable ABC model choice via random forests.* Bioinformatics 32(6), 859-866.

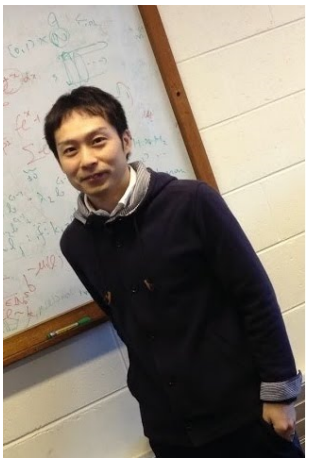
[7] Sisson, S.A., Fan, Y., Beaumont, M. (2018) *Handbook of Approximate Bayesian Computation.* Chapman and Hall.

[8] Wang, X., Nott, D.J., Drovandi, C.C., Mengersen, K., Evans, M. (2018) *Using history matching for prior choice.* Technometrics 60(4), 445-460.

Tuesday 9:00–10:00 **Room: CB08.03.005**

Richardson extrapolation and higher order QMC

Takashi Goda
University of Tokyo, goda@frcer.t.u-tokyo.ac.jp



In his seminal paper [1], Dick introduced an explicit construction of quasi-Monte Carlo (QMC) point sets based on *digit interlacing* such that the integration error decays faster than $O(1/N)$ for smooth integrands, which was the beginning of higher order QMC. Later on, it was shown in [4] and some other papers that his digit interlacing approach fits well with component-by-component construction of polynomial lattice point sets, which gave birth to *interlaced polynomial lattice rules* that have nowadays applications to partial differential equations with random coefficients and Bayesian inversion/estimation problems.

Recently, non-trivial connections between Richardson extrapolation and higher order QMC have been found. Firstly it was shown in [2] that Richardson extrapolation applied to (standard) polynomial lattice rules with geometric spacing of N achieves essentially the same error rate with interlaced polynomial lattice rules. This roughly means Richardson extrapolation can play an alternative role to digit interlacing. We call the resulting quadrature algorithm *extrapolated polynomial lattice rules*. Secondly it was shown in [3] that Richardson extrapolation allows for truncation of higher order QMC point sets which are constructed based on digit interlacing, as is almost clear from the title of the paper. This way, Richardson extrapolation, a classical numerical algorithm more than 100 years old, conveys novel insights into (relatively much younger) higher order QMC. The aim of this talk is to present these recent developments of higher order QMC.

Part of this talk is based on joint work with Josef Dick and Takehito Yoshiki.

[1] J. Dick. Walsh spaces containing smooth functions and quasi-Monte Carlo rules of arbitrary high order. *SIAM J. Numer. Anal.* 46(3):1519–1553, 2008.

[2] J. Dick, T. Goda, and T. Yoshiki. Richardson extrapolation of polynomial lattice rules. *SIAM J. Numer. Anal.* 57(1):44–69, 2019.

[3] T. Goda. Richardson extrapolation allows truncation of higher order digital nets and sequences. *IMA J. Numer. Anal.* 2019 (to appear).

[4] T. Goda, and J. Dick. Construction of interlaced scrambled polynomial lattice rules of arbitrary high order. *Found. Comput. Math.* 15(5):1245–1278, 2015.

Tuesday 13:30–14:30 **Room: CB08.03.005**

Construction Algorithms for (Polynomial) Lattice Points

Peter Kritzer
RICAM, Austria, peter.kritzer@oeaw.ac.at

The (fast) component-by-component construction of lattice point sets and polynomial lattice point sets is a powerful method to obtain quadrature rules for approximating integrals over the d-dimensional unit cube.

In this talk, we give a survey of recent developments regarding the component-by-component construction, and present improvements in the construction cost for lattice rules and polynomial lattice rules in weighted function spaces. Moreover, we discuss other algorithms to construct (polynomial) lattice rules, such as the successive coordinate search algorithm and digit-by-digit constructions.

We show bounds on the errors of the resulting quasi-Monte Carlo rules, and present numerical results.

This work is in collaboration with the following colleagues:

Josef Dick
UNSW Sydney, Australia, josef.dick@unsw.edu.au

Adrian Ebert
KU Leuven, Belgium, adrian.ebert@cs.kuleuven.be

Gunther Leobacher
KFU Graz, Austria, gunther.leobacher@uni-graz.at

Dirk Nuyens
KU Leuven, Belgium, dirk.nuyens@cs.kuleuven.be

Onyekachi Osisiogu
RICAM, Austria, onyekachi.osisiogu@oeaw.ac.at

Friedrich Pillichshammer
JKU Linz, Austria, friedrich.pillichshammer@jku.at

Wednesday 9:00–10:00 **Room: CB08.03.005**

Overcoming the curse of dimensionality: from nonlinear Monte Carlo to deep artificial neural networks

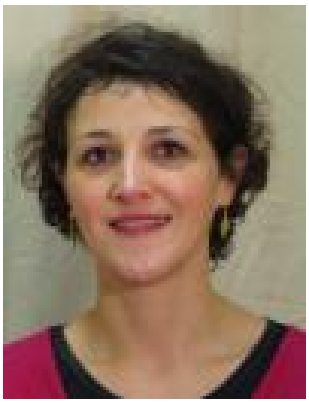
Arnulf Jentzen
ETH Zürich, Switzerland, arnulf.jentzen@sam.math.ethz.ch

Partial differential equations (PDEs) are among the most universal tools used in modelling problems in nature and man-made complex systems. For example, stochastic PDEs are a fundamental ingredient in models for nonlinear filtering problems in chemical engineering and weather forecasting, deterministic Schroedinger PDEs describe the wave function in a quantum physical system, deterministic Hamiltonian-Jacobi-Bellman PDEs are employed in operations research to describe optimal control problems where companys aim to minimise their costs, and deterministic Black-Scholes-type PDEs are highly employed in portfolio optimization models as well as in state-of-the-art pricing and hedging models for financial derivatives. The PDEs appearing in such models are often high-dimensional as the number of dimensions, roughly speaking, corresponds to the number of all involved interacting substances, particles, resources, agents, or assets in the model. For instance, in the case of the above mentioned financial engineering models the dimensionality of the PDE often corresponds to the number of financial assets in the involved hedging portfolio. Such PDEs can typically not be solved explicitly and it is one of the most challenging tasks in applied mathematics to develop approximation algorithms which are able to approximately compute solutions of high-dimensional PDEs. Nearly all approximation algorithms for PDEs in the literature suffer from the so-called "curse of dimensionality" in the sense that the number of required computational operations of the approximation algorithm to achieve a given approximation accuracy grows exponentially in the dimension of the considered PDE. With such algorithms it is impossible to approximately compute solutions of high-dimensional PDEs even when the fastest currently available computers are used. In the case of linear parabolic PDEs and approximations at a fixed space-time point, the curse of dimensionality can be overcome by means of Monte Carlo approximation algorithms and the Feynman-Kac formula. In this talk we introduce new nonlinear Monte Carlo algorithms for high-dimensional nonlinear PDEs. We prove that such algorithms do indeed overcome the curse of dimensionality in the case of a general class of semilinear parabolic PDEs and we thereby prove, for the first time, that a general semilinear parabolic PDE with a nonlinearity depending on the PDE solution can be solved approximately without the curse of dimensionality.

Thursday 9:00–10:00 **Room: CB08.03.005**

Monte Carlo methods and Optimization: Intertwinings

Gersendne Fort
CNRS & IMT, France, Gersende.Fort@math.univ-toulouse.fr



Coupling Monte Carlo methods and Optimization techniques revealed to be a powerful tool to answer intractable computations.

On one hand, new Monte Carlo algorithms are designed to be more robust to some difficult sampling situations, by learning on the fly how to be more efficient: the sampling mecanism evolves at each iteration, since parameterized by a quantity which is seen as the current estimate of the optimal strategy. Conversely, optimization techniques such as Stochastic Approximation algorithms are, by definition, coupled with Monte Carlo methods: when the objective function is unknown and the function or its derivatives are expressed as an expectation, it can be replaced by a Monte Carlo sum even reduced to a single point. In both situations, efficient implementations of the algorithms rely on iterating two steps: a simulation step and an optimization step. The output of the simulation step governs the optimization step; and the output of the optimization step governs the simulation step.

Motivated by Statistical Learning applications, the talk will derive sufficient conditions for the convergence analysis of these intertwined Monte Carlo and Optimization algorithms, with an emphasis on the case when the sampling step relies on Markov chain Monte Carlo algorithms. Therefore, the sampling mecanism is biased. For example, when plugged in a stochastic gradient-based algorithm, it means that at each iteration, the gradient is replaced with a biased estimate. We will see that even if this bias does not vanish, the intertwined algorithms converges.

Thursday 13:30–14:30 **Room: CB08.03.005**

Rare-Event Simulation of Regenerative Systems: Estimation of the Mean and Distribution of Hitting Times

Bruno Tuffin
INRIA Rennes Bretagne-Atlantique, France, Bruno.Tuffin@inria.fr

Rare events occur with a very small probability but are important to analyze because of potential catastrophic consequences. During this talk, we will focus on rare event for regenerative processes, that is, processes such that sections of the process are statistically independent of each other. For many complex and/or large models, simulation is the only tool at hand but requires specific implementations to get an accurate answer in a reasonable time. There are two main families of rare-event simulation techniques: importance sampling (IS) and splitting.

In part one, we will briefly review them and then compare their respective advantages. Later (somewhat arbitrarily) we will devote most of the talk to IS.

In part two, we focus on the estimation of the mean hitting time of a rarely visited set. A natural and direct estimator consists in averaging independent and identically distributed copies of simulated hitting times, but an alternative standard estimator uses the regenerative structure allowing to represent the mean as a ratio of quantities. We will see that in the setting of crude simulation, the two estimators are actually asymptotically identical in a rare-event context, but inefficient for different, even if related, reasons: the direct estimator requires a large average computational time of a single run whereas the ratio estimator faces a small probability computation. We then explain how the ratio estimator is improved when using IS.

In the third part of the talk, we will discuss the estimation of the distribution, not just the mean, of the hitting time to a rarely visited set of states. We will exploit the property that the distribution of the hitting time divided by its expectation converges weakly to an exponential distribution, as the target set probability decreases to zero. The problem then reduces to the extensively studied estimation of the mean described previously. It leads to simple estimators of a quantile and conditional tail expectation of the hitting time. Some variants will be presented and the accuracy of the estimators illustrated on numerical examples.

This talk is mostly based on collaborative works with Peter W. Glynn and Marvin K. Nakayama.

Friday 11:00–12:00 **Room: CB08.03.005**

Sampling via Transport: Preconditioning and Low-Dimensional Structure

Youssef Marzouk
Massachusetts Institute of Technology, ymarz@mit.edu

Integration against an intractable probability measure is a fundamental challenge in Bayesian inference and well beyond. A useful approach to this problem seeks a deterministic coupling of the measure of interest with a tractable "reference" measure (e.g., a standard Gaussian). Such couplings are induced by transport maps, and enable direct simulation from the desired measure simply by evaluating the transport map at samples from the reference. In recent years, an enormous variety of representations and constructions for such transport maps have been proposed—ranging from monotone polynomials or invertible neural networks to the flows of ODEs. Approximate transports can also be used to "precondition" and accelerate standard Monte Carlo schemes. Within this framework, one can describe many useful notions of low-dimensional structure: for instance, sparse or decomposable transports underpin modeling and computation with non-Gaussian Markov random fields, and low-rank transports arise frequently in inverse problems.

I will present a broad overview of this framework, describing how to construct suitable classes of transport maps, and then focus on two recent developments: adaptive MCMC schemes that use transport to create more favorable target geometry, and greedy variational methods that build high-dimensional maps by composing multiple *low-dimensional* maps or flows.

This work is in collaboration with the following colleagues:

Daniele Bigoni
Massachusetts Institute of Technology, dabi@mit.edu

Matthew Parno
CRREL, matthew.d.parno@erdc.dren.mil

Alessio Spantini
Massachusetts Institute of Technology, alessio.spantini@gmail.com

Olivier Zahm
INRIA, olivier.zahm@inria.fr

Abstracts of Invited and Contributed Talks

Monday 10:30–12:30, **Room: CB08.02.002**

Advances in sequential Monte Carlo methods: Theory, algorithms and applications

Chair: Christopher Drovandi

On the Exponential Ergodicity of the ZigZag Process

Gareth Roberts
University of Warwick, UK, Gareth.O.Roberts@warwick.ac.uk
Giorgos Vasdekis
University of Warwick, UK, G.Vasdekis@warwick.ac.uk

Piecewise Deterministic Markov Processes have recently drawn the attention of the Markov Chain Monte Carlo community. The first reason for this is that, in general, one can simulate exactly the entire path of such a process. The second is that these processes are non-reversible, which sometimes leads to faster mixing [3]. Most of the Piecewise Deterministic Algorithms constructed nowadays are based on two processes: The ZigZag and the Bouncing Particle Sampler Processes. The ZigZag Process moves linearly in the space $\mathbf{R}^d$ in directions taken from the set $\{-1,+1\}^d$, for a random period of time, changing direction on one coordinate at a time. An important question related to these samplers is the existence of a Central Limit Theorem which is closely connected to the property of Exponential Ergodicity. For both of the processes this property has been verified in [2] and [1] for target distributions that have light tails (amongst other properties). In this talk we will prove that the ZigZag Sampler is not Exponentially Ergodic in a heavy tails scenario and will briefly explain how one could try to correct this by allowing the ZigZag process to take more velocities than ± 1.

[1] Joris Bierkens, Gareth O. Roberts, and Pierre-André Zitt. *Ergodicity of the zigzag process.* `to appear in` Annals of Applied Probability.

[2] George Deligiannidis, Alexandre Bouchard-CÃ´tÃ©, Arnaud Doucet. *Exponential Ergodicity of the Bouncy Particle Sampler.* `to appear in` Annals of Statistics.

[3] Persi Diaconis, Susan Holmes, and Radford M. Neal. *Analysis of a nonreversible Markov chain sampler.* `to appear in` Annals of Applied Probability, 2000.

Likelihood approximations for latent variable models: computational and statistical scalability

Anthony Lee
University of Bristol, anthony.lee@bristol.ac.uk
Lawrence Middleton
University of Oxford, lawrence.middleton@spc.ox.ac.uk
Arnaud Doucet
University of Oxford, doucet@stats.ox.ac.uk

A popular statistical modelling technique is to model data as a partial and noisy observation of a random process. This allows, in principle, one to fit sophisticated domain-specific models with interpretable parameters. However, the likelihood function in such models is typically intractable, and so likelihood-based inference techniques must deal with this intractability in some way. I will briefly discuss two simple likelihood-based approaches to intractability, pseudo-marginal Markov chain Monte Carlo and simulated maximum likelihood, and comment on statistical and computational scalability in some example settings.

Unbiased and Consistent Nested Sampling via Sequential Monte Carlo

Robert Salomone
UNSW, Australia, robert.salomone@uqconnect.edu.au

Leah F. South
Lancaster University, U.K., leah.south@hdr.qut.edu.au

Christopher C. Drovandi
QUT, Australia, c.drovandi@qut.edu.au

Dirk P. Kroese
UQ, Australia, kroese@maths.uq.edu.au

We introduce a new class of sequential Monte Carlo methods called Nested Sampling via Sequential Monte Carlo (NS-SMC), which reframes the Nested Sampling method of [2] in terms of sequential Monte Carlo techniques. This new framework allows one to obtain provably consistent estimates of the marginal likelihood and posterior inferences when Markov chain Monte Carlo (MCMC) is used to produce new samples. An additional benefit is that marginal likelihood estimates are also unbiased. For applications of NS–SMC, we give advice on tuning MCMC kernels in an automated manner via a preliminary pilot run, and present a new method for appropriately choosing the number of MCMC repeats at each iteration. A numerical study is conducted where the performance of NS–SMC and temperature–annealed SMC is compared on several challenging and realistic problems.

[1] Robert Salomone, Leah F. South, Christopher C. Drovandi, and Dirk P. Kroese. Unbiased and consistent nested sampling via sequential Monte Carlo. *arXiv:1805.03924*, 2018.

[2] John Skilling. Nested sampling for general Bayesian computation. *Bayesian Analysis*, 1(4):833–859, 2006.

Limit theorems for sequential MCMC methods

Axel Finke
National University of Singapore, Singapore, axel.finke@nus.edu.sg

Arnaud Doucet
University of Oxford, UK, doucet@stats.ox.ac.uk

Adam M. Johansen
University of Warwick & The Alan Turing Institute, UK, a.m.johansen@warwick.ac.uk

Both *sequential Monte Carlo (SMC)* methods (a.k.a. "particle filters") as well as *sequential MCMC* methods methods constitute classes of algorithms which can be used to approximate expectations with respect to (a sequence of) probability distributions and their normalising constants. While SMC methods sample particles conditionally independently at each time step, sequential MCMC methods sample particles according to an MCMC kernel. Introduced over twenty years ago in [1], sequential MCMC methods have attracted renewed interest recently as they empirically outperform SMC methods in some applications. We establish a strong law of large numbers and a central limit theorem for sequential MCMC methods and provide conditions under which errors can be controlled uniformly in time. In the context of state-space models, we also provide conditions under which sequential MCMC methods can indeed outperform standard SMC methods in terms of asymptotic variance of the corresponding Monte Carlo estimators.

[1] C. Berzuini, N. G. Best, W. R. Gilks and L. Cristiana Dynamic conditional independence models and Markov chain Monte Carlo methods. *Journal of the American Statistical Association,* 92(440):1403–1412, 1997.

Monday 10:30–12:00, **Room: CB08.03.004**

Advances in Rare-event Simulation

Chair: Victor Elvira

Importance Sampling for Rare Catastrophic Events

Chang-Han Rhee
Northwestern University, USA, chang-han.rhee@northwestern.edu

Bohan Chen
Centrum Wiskunde & Informatica, Netherlands, b.chen@cwi.nl

Jose Blanchet
Stanford University, USA, jose.blanchet@stanford.edu

Bert Zwart
Centrum Wiskunde & Informatica, Netherlands, bert.zwart@cwi.nl

We propose a class of strongly efficient rare event simulation estimators for random walks and compound Poisson processes with a regularly varying increment distribution in a general large deviations regime [1]. Our estimator is based on an importance sampling strategy that hinges on the heavy-tailed sample path large deviations result recently established in [2]. The new estimators are straightforward to implement and can be used to systematically evaluate the probability of a wide range of rare events with bounded relative errors. They are *universal* in the sense that a single importance sampling scheme applies to a very general class of rare events that arise in heavy-tailed systems. In particular, our estimators can deal with rare events that are caused by multiple big jumps (therefore, beyond the principle of a single big jump) as well as multidimensional processes such as the buffer content process of a queueing network. We illustrate the versatility of our approach with several applications that arise in the context of mathematical finance, actuarial science, and queueing theory.

[1] B. Chen, J. Blanchet, C.-H. Rhee, B. Zwart. Efficient rare-event simulation for multiple jump events in regularly varying random walks and compound Poisson processes *To Appear in Mathematics of Operations Research*. arXiv:1706.03981.

[2] C.-H. Rhee, J. Blanchet, B. Zwart. Sample path large deviations for LÃ©vy processes and random walks with regularly varying increments. *To Appear in Annals of Probability*. arXiv preprint arXiv:1606.02795.

Exploiting Asymptotics and Polar Coordinates for Rare Tail Estimation

Thomas Taimre
School of Mathematics and Physics, The University of Queensland, Australia, t.taimre@uq.edu.au

Patrick Laub
School of Mathematics and Physics, The University of Queensland, Australia and Department of Mathematics, Aarhus University, Denmark, p.laub@uq.edu.au|p.laub@math.au.dk

We consider the problem of estimating the right tail probability of a sum of random variables when the density of the sum is not known explicitly, but whose asymptotic behaviour is known. We embed this asymptotic structure into a simple importance sampling estimator, in which we consider the radial and angular components of the distribution separately. The estimator and procedure are applicable in both the heavy- and light-tailed settings, as well as for dependent and independent summands. We illustrate the approach with a series of examples.

Importance sampling for signal processing applications

Victor Elvira
IMT Lilles Douai, France, victor.elvira@imt-lille-douai.fr

Ignacio Santamaría
Universidad de Cantabria, Spain, i.santamaria@unican.es

Art Owen
Stanford University, USA, owen@stanford.edu

Digital communications are based on the transmission of symbols that belong to a finite alphabet, each of them carrying one or several bits of information. The receiver estimates the symbol that was transmitted, and in the case of perfect communication without errors, the original sequence of bits is reconstructed. However, real-world communication systems (e.g., in wireless communications) introduce random distortions in the symbols, including additive Gaussian noise, provoking errors in the detected symbols at the receiver. The characterization of the symbol error rate (SER) of the system is of major interest in communications engineering. However, in many systems of interest, the integrals required to evaluate the symbol error rate (SER) in the presence of Gaussian noise are impossible to compute in closed-form, and therefore Monte Carlo simulation is typically used to estimate the SER. Naive Monte Carlo simulation has been traditionally used in the communications literature, even if it can be very inefficient and require very long simulation runs, especially at high signal-to-noise-ratio (SNR) scenarios [1]. At high SNR, the variance of the additive Gaussian noise is small, and hence the rate of errors is very low, which yields raw Monte Carlo impracticable for this rare event estimation problem. In this talk, we start describing (for non-experts) the problem of SER estimation of communication system. Then, we adapt a recently proposed multiple importance sampling (MIS) technique, called ALOE (for "At Least One rare Event" [2]), to this problem [3]. Conditioned to a transmitted symbol, an error (or rare event) occurs when the observation falls in a union of half-spaces or, equivalently, outside a given polytope. The proposal distribution for ALOE samples the system conditionally on an error taking place, which makes it more efficient than other importance sampling techniques. ALOE provides unbiased SER estimates with simulation times orders of magnitude shorter than conventional Monte Carlo. Then, we discuss the challenges of SER estimation in multiple-input multiple-output (MIMO) communications, where the rare-event estimation problem requires solving a large number of integrals in a higher-dimensional space. We propose a novel MIS-based approach exploiting the strengths of the ALOE estimator.

[1] D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo methods.* John Wiley & Sons, 2011.

[2] A. B. Owen, Y. Maximov, and M. Chertkov, *Importance sampling the union of rare events with an application to power systems analysis.* Electronic Journal of Statistics, Volume 13, pp. 231-254. 2019.

[3] V. Elvira and I. Santamaría. *Multiple Importance Sampling for Efficient Symbol Error Rate Estimation.* IEEE Signal Processing Letters, Volume 26, no. 3, pp. 420-424, March, 2019.

Monday 10:30–12:30, **Room: CB08.03.002**

Current Challenges in High-Dimensional Algorithms

Chair: Peter Kritzer

An Optimal Adaptive Algorithm Based on the Decay Rate of the Series Coefficients of the Input Functions

Yuhan Ding
Misericordia University, USA, yding@misericordia.edu

Fred Hickernell
Illinois Institute of Technology, USA, hickernell@iit.edu

Peter Kritzer
Austrian Academy of Sciences, Austria, peter.kritzer@oeaw.ac.at

Simon Mak
Georgia Institute of Technology, USA, smak6@gatech.edu

Automatic algorithms attempt to provide approximate solutions that differ from exact solutions by no more than a user-specified error tolerance. This talk describes an automatic adaptive algorithm by tracking the decay rate of the series coefficients of the input functions. We assume that the Fourier coefficients of the input function to be approximated decay sufficiently fast, but do not require the decay rate to be known a priori. We also assume that the Fourier coefficients decay steadily, although not necessarily monotonically. Under these assumptions, our adaptive algorithm is shown to achieve an approximation to the function satisfying the desired error tolerance, without prior knowledge of the norm of the function to be approximated. Moreover, the computational cost of our algorithm and the lower complexity bound of this kind of problem are presented in the talk. With this, our algorithm is shown to be essentially optimal. In addition to that, the tractablity of this kind of problem is discussed.

An Optimal Adaptive Algorithm Based on a Pilot Sample

Yuhan Ding
Misericordia University, USA, yding@misericordia.edu

Fred J. Hickernell
Illinois Institute of Technology, USA, hickernell@iit.edu

Peter Kritzer
Austrian Academy of Sciences, Austria, peter.kritzer@oeaw.ac.at

Simon Mak
Georgia Institute of Technology, USA, smak6@gatech.edu

Adaptive algorithms are convenient for the practitioner because they automatically determine the computational effort required to satisfy the error criterion. The function data acquired for constructing the approximate solution are also used to compute a data-based error bound for the approximate solution. Computation proceeds until this error bound becomes small enough. If the set of allowed input functions is convex, adaptive algorithms may offer no advantage to non-adaptive algorithms. We construct an adaptive algorithm for solving a general, linear problem where the input functions lie in a *non-convex cone*. The stopping criterion is based on theory, not heuristics. The cone of input functions is defined

so that sampling the most important Fourier series coefficients is sufficient to bound the magnitude of the unsampled Fourier coefficients. We show that our adaptive algorithm is optimal. We also determine conditions under which the problem is tractable. This work is related to the adaptive algorithms developed in [1, 2, 3].

[1] F. J. Hickernell, L. Jiang, Y. Liu, and A. B. Owen. Guaranteed conservative fixed width confidence intervals via Monte Carlo sampling. In J. Dick, F. Y. Kuo, G. W. Peters, and I. H. Sloan, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2012*, volume 65 of *Springer Proceedings in Mathematics and Statistics*, pages 105–128. Springer-Verlag, Berlin, 2013.

[2] F. J. Hickernell, Ll. A. Jiménez Rugama, and D. Li. Adaptive quasi-Monte Carlo methods for cubature. In J. Dick, F. Y. Kuo, and H. Woźniakowski, editors, *Contemporary Computational Mathematics — a celebration of the 80th birthday of Ian Sloan*, pages 597–619. Springer-Verlag, 2018.

[3] R. J. Kunsch, E. Novak, and D. Rudolf. Solvable integration problems and optimal sample size selection. *Journal of Complexity*, 2019. To appear.

Numerical methods for partial differential equations with random coefficients

Michael Feischl
TU Wien, Austria, michael.feischl@tuwien.ac.at

Christoph Schwab
ETH Zürich, Switzerland, christoph.schwab@sam.math.ethz.ch

Mathematical models often contain uncertainty in parameters and measurements. In this talk we focus on partial differential equations where some parameters are modeled by random variables. The main example comes for the diffusion equation where the diffusion parameters is modeled as a random field which randomly fluctuates around a given mean. We discuss recent progress on numerical methods in quantifying this uncertainty, in particular, we discuss the extension of the multi-level method to the multi-index method.

Approximation in Chebyshev space using lattice rules

Frances Kuo
UNSW Sydney, Australia, f.kuo@unsw.edu.au

Giovanni Migliorati
Sorbonne Université, France, migliorati@ljll.math.upmc.fr

Fabio Nobile
EPFL, Switzerland, fabio.nobile@epfl.ch

Dirk Nuyens
KU Leuven, Belgium, dirk.nuyens@cs.kuleuven.be

Lattice rules are mainly applied to periodic functions from Fourier space. Here we study how to use lattice rules for approximation of non-periodic functions from high-dimensional Chebyshev space. We map from Chebyshev space into cosine space and from cosine space into Fourier space. This enables us to use lattice rules for function approximation in Chebyshev space, partially using modifications to construction algorithms for lattice rules in Fourier space. "Chebyshev lattices" have previously been considered by [1] and [2, 3], their definitions correspond to ours when the number of points of the original

lattice rule is even. Fast DCTs can be used to map point values to frequency representation and visa versa. We provide new theory and efficient algorithms for CBC style construction of good lattice rules for approximation.

[1] R. Cools, K. Poppe. Chebyshev lattices, a unifying framework for cubature with Chebyshev weight function, BIT Numer. Math., 51:275-288, 2011.

[2] D. Potts, T. Volkmer. Fast and exact reconstruction of arbitrary multivariate algebraic polynomials in Chebyshev form, 2015 International Conference on Sampling Theory and Applications (SampTA), IEEE, 2015.

[3] D. Potts, T. Volkmer. Sparse high-dimensional FFT based on rank-1 lattice sampling, Appl. Comput. Harmon. Anal. 41:713-748, 2016.

Monday 10:30–12:30, **Room: CB08.03.005**

Advances in Exact and approximate Bayesian computation

Chair: Robert Kohn

The block-Poisson estimator for optimally tuned signed pseudo-marginal MCMC

Matias Quiroz
University of Technology, Australia, quiroz.matias@gmail.com

Minh-Ngoc Tran
University of Sydney, Australia,

Mattias Villani
Linköping University & Stockholm University, Sweden,

Robert Kohn
University of New South Wales, Australia,

Khue-Dung Dang
University of New South Wales, Australia,

We propose a pseudo-marginal Markov Chain Monte Carlo (MCMC) method that estimates the likelihood using a block-Poisson estimator. The estimator is a product of Poisson estimators, each based on an independent set of random numbers used to construct the estimator. The construction allows us to update the random numbers in a subset of the blocks in each MCMC iteration, thereby inducing a controllable correlation between the estimates at the current and proposed draw in the Metropolis-Hastings ratio. This makes it possible to use highly variable likelihood estimators (which are computationally much faster) without adversely affecting the sampling efficiency. Poisson estimators are unbiased but not necessarily positive. We therefore follow [1] and run the MCMC on the absolute value of the estimator, which we term signed pseudo-marginal MCMC, and use an importance sampling correction for occasionally negative likelihood estimates to estimate expectations of any function of the parameters consistently. We provide analytically derived guidelines to select the optimal tuning parameters for the block-Poisson estimator by minimizing the variance of the importance sampling corrected estimator per unit of computing time.

[1] A-M. Lyne, M. Girolami, Y. Atchadé, H. Strathmann and D. Simpson. On Russian roulette estimates for Bayesian inference with doubly-intractable likelihoods *Statistical Science*. Institute of Mathematical Statistics, Volume 30, 2018.

Bayesian Parameter Inference for Stochastic Differential Equation Mixed Effects Models

Imke Botha
QUT Brisbane, Australia, imke.botha@hdr.qut.edu.au

Chris Drovandi
QUT Brisbane, Australia, c.drovandi@qut.edu.au

Robert Kohn
UNSW, Sydney, Australia, r.kohn@unsw.edu.au

Parameter inference for stochastic differential equation mixed effects models (SDEMEMs) is a challenging problem. Analytical solutions for these models are rarely available, which means that the likelihood is also intractable. [1] proved that exact inference is possible for state space models with intractable likelihoods, provided the transition density can be simulated from. These pseudo-marginal methods replace the intractable likelihood with a nonnegative unbiased estimate, calculated with the use of auxiliary variables. A useful application of this idea is particle MCMC, where the unbiased estimate is supplied by a particle filter. While the exact posterior is targetted by these methods, a naive implementation for SDEMEMs can be highly inefficient. We present three extensions to the naive approach which exploits specific aspects of SDEMEMs and other advances such as correlated pseudo-marginal methods. We compare these methods to data from a tumour xenography study on mice. This is joint work with Chris Drovandi and Robert Kohn.

[1] Christophe Andrieu and Gareth O Roberts. *The pseudo-marginal approach for efficient Monte Carlo computations.* The Annals of Statistics, 37(2):697–725, 2009.

Introducing high-dimensional sensitivity analysis of Bayesian MCMC inference

Liana Jacobi
University of Melbourne,

Chun Fung Kwok
University of Melbourne,

Andres Ramirez-Hassan
Monash University,

Nhung Nghiem
University of Otago,

Prior specification has been the subject of a long-standing and lively debate in Bayesian inference, as it turns out to be difficult to assess in most problems due to the computational and analytical complexity and the large prior and posterior parameter spaces. While initially global robustness analysis was defined over classes of prior distributions, the focus has shifted to local sensitivity with respect to hyperparameters as a result of computational feasibility and the rise of conjugate priors.

Existing approaches for Bayesian sensitivity are generally limited to low-dimensional and local analysis (in a form of derivatives or grid-point evaluations) due to the computational and analytical complexity of the large (hyper-)parameter spaces.

In this paper we introduce a new computational approach that extends prior sensitivity analysis of posterior Markov chain Monte Carlo (MCMC) inference from the common one-dimensional point-wise perspective to high-dimensional over the hyper-parameter space by generating and analyzing manifolds of posterior measures. The approach builds on Gaussian Processes and recent advances in Automatic Differentiation for MCMC local sensitivity analysis. It provides a tool to assess more global and robust measures of prior sensitivity of posterior output via graphical and numerical robustness measures defined over the hyper-parameter space.

A long short-term memory stochastic volatility model

Nghia Nguyen
University of Sydney,

Minh-Ngoc Tran
University of Sydney,

David Gunawan
University of New South Wales,

Robert Kohn
University of New South Wales,

Stochastic Volatility (SV) models are widely used in the financial sector while Long Short-Term Memory (LSTM) models have been successfully used in many large-scale industrial applications of Deep Learning. Our article combines these two methods in a non-trivial way and proposes a model for capturing the dynamics of financial volatility process, which we call the LSTM-SV model.

The proposed model overcomes the short-term memory problem in conventional SV models,

is able to capture non-linear dependence in the latent volatility process, and often has a better out-of-sample forecast performance than SV models.

These conclusions are illustrated through simulation studies and applications to three financial time series datasets: US stock market index SP500, Australian stock index ASX200 and Australian-US dollar exchange rates.

We demonstrate how recent advances in particle MCMC can be successfully used for Bayesian inference in highly sophisticated state-space models like the proposed LSTM-SV model.

A user-friendly software package together with the examples reported are available at https://github.com/vbayeslab.

Monday 15:00–17:00, **Room: CB08.02.002**

Advanced Monte Carlo in risk management

Chair: Emmanuel Gobet

Multilevel Monte-Carlo method and lower/upper bounds in Initial Margin computations

Florian Bourgey
Ecole polytechnique, France, florian.bourgey@polytechnique.edu

Stefano De Marco
Ecole polytechnique, France, stefano.de.marco@cmap.polytechnique.fr

Emmanuel Gobet
Ecole polytechnique, France, emmanuel.gobet@polytechnique.edu

Alexandre Zhou
CERMICS - Ecole des Ponts ParisTech, France, alexandre.zhou@icloud.com

The Multi-level Monte-Carlo (MLMC) method developed by Giles [2008] has been successfully applied in innumerous fields of stochastic simulation. Quoting Giles [2015], MLMC "reduces the computational cost [*with respect to standard Monte-Carlo*] by performing most simulations with low accuracy at a correspondingly low cost, with relatively few simulations being performed at high accuracy and a high cost". A natural application of this method is the evaluation of *nested* expectation of the form $E\big[g(E[f(X,Y)|X])\big]$, where f, g are given functions and (X, Y) a couple of independent random variables. Apart from the pricing of American-type derivatives, such computations arise in a large variety of risk valuations (VaR or CVaR of a portfolio, CVA), or in the assessment of margin costs of centrally cleared portfolios. In this work, we focus on the computation of Initial margins. We analyze the properties and asymptotically optimal choices of MLMC estimators in practical situations of limited regularity of the outer function g (with singularities in the first derivative). In parallel, we investigate upper and lower bounds for nested expectations as above, in the spirit of primal/dual algorithms for stochastic control problems.

Particle methods in risk analysis

Pierre Del Moral
INRIA, France, pierre.del-moral@inria.fr

We present an overview of particle methods (a.k.a. Sequential Monte Carlo) in rare event simulation and risk analysis. We provide a catalog of branching type and interacting particle methods for sampling rare events and estimating sensitivy measures, including multilevel splitting techniques as well as island type particle methodologies. We illustrate these particle samplers in the context of black box and uncertainty propagation problems.

Weak error analysis for some mean-field SDEs

Benjamin Jourdain
CERMICS, France, jourdain@cermics.enpc.fr

The weak error between a stochastic differential equation with nonlinearity in the sense of McKean given by moments and its approximation by the Euler discretization with time-step h of a system of N interacting particles is proved to be of order $\mathcal{O}(N^{-1}+h)$. Numerical experiments confirm this behaviour and show that it extends to more general mean-field interaction and in particular to the Atlas mean-field model.

McKean stochastic optimal control of an energy storage system to reduce demand variability

Maxime Grangereau
EDF and CMAP, Ecole polytechnique, maxime.grangereau@edf.fr

Emmanuel Gobet
CMAP, Ecole polytechnique, emmanuel.gobet@polytechnique.edu

We consider a system, e.g., a micro-grid, characterized by its own exogenous production and consumption, connected to the network and to a controllable energy storage system, like a battery for instance. We control this energy storage system and aim at limiting simultaneously the power demand peaks on the network, the storage system aging and the fluctuations of the power supplied by the network. While the first two components of the cost are related to the trajectory of the system, and therefore to the standard stochastic optimal control paradigm, the last one is more related to the probability distribution of the system, and therefore to the McKean stochastic optimal control paradigm.

We propose here to model the problem of energy storage control as a McKean stochastic optimal control problem with scalar interactions. For this class of problems, we derive necessary and sufficient optimality conditions, using the stochastic Pontryagyn principle. This gives rise to a particular McKean-Forward Backward Stochastic Differential Equation. Some existence results for a solution to this class of problems are obtained in the general case. In the Linear-Quadratic case, some quasi-explicit formulas are available. In the non-quadratic case, a first order expansion of the optimal control with respect to the non-quadratic terms can be efficiently computed. The performance of our approach is demonstrated through numerical examples in both the linear-quadratic and quasi linear-quadratic cases.

Monday 15:00–17:00, **Room: CB08.03.004**

Stochastic Computation and Complexity I

Chair: Stefan Heinrich

Smooth dependence of fractional Brownian motion on its Hurst parameter and the limits of rough paths theory

Stefan Koch
Universität Mannheim, stefan.koch@uni-mannheim.de

Andreas Neuenkirch
Universität Mannheim, neuenkirch@math.uni-mannheim.de

We show that Mandelbrot-van Ness fractional Brownian motion (MvN-fBm) depends pathwise smoothly on its Hurst parameter $H \in (0,1)$. Moreover, this smooth dependence carries over to stochastic differential equations (SDEs) driven by MvN-fBm, if the driving fBm is scalar or satisfies $H > 1/2$.

However, for $H \leq 1/2$ the existence of the Lévy area for the joint process of MvN-fBm and its derivative with respect to H is unknown. In particular, the classical approach of Coutin and Qian (2002) to define this area as the limit of the area of the piecewise linearly interpolated process fails. Thus, the pathwise smoothness of such SDEs with respect to H remains an open problem in this case.

Randomized Euler scheme for strong approximation of solutions of SDEs with time-irregularities

Paweł Przybyłowicz
AGH University of Science and Technology, Krakow, Poland, pprzybyl@agh.edu.pl

We consider approximate solving of the following scalar SDE

$$\begin{cases} dX(t) = a(t, X(t))dt + b(t)dW(t), \quad t \in [0,T], \\ X(0) = \eta, \end{cases} \tag{7.1}$$

driven by a standard one-dimensional Wiener process $W = (W(t))_{t\in[0,T]}$. We assume that $a = a(t,y)$ and $b = b(t)$ are only measurable with respect to the time variable t, and a is globally Lipschitz with respect to the space variable y.

It is well-known that under such assumptions lack of convergence occurs, as long as we sample deterministically coefficients a and b with respect to the time variable t. Hence, we investigate behavior of the randomized Euler scheme X_n^{RE}, which evaluates a and b at randomly chosen points. By using Information-Based Complexity framework we show that randomized Euler scheme converges to the solution X of the underlying SDE but the convergence of X_n^{RE} to X may be arbitrarily slow. ([5]).

In order to get positive results we assume that b belongs to the Sobolev-Slobodeckij space $W^{\sigma,p}$, $\sigma \in (0,1)$, $p > 2$. In this case we show that the $L^2(\Omega)$-error of the algorithm X_n^{RE} is $O(n^{-\min\{\frac{1}{2}-\frac{1}{p},\sigma\}})$ ([1]). In particular, this extends the results, obtained for the randomized Euler scheme, from [1], [2], [4], and [6].

This is a joint work with Raphael Kruse (TU Berlin, Germany, kruse@math.tu-berlin.de)

[1] S. Heinrich, and B. Milla *The randomized complexity of initial value problems* J. Complexity, Volume 24, 2008, pages 77–88.

[2] A. Jentzen, and A. Neuenkirch *A random Euler scheme for Carathéodory differential equations* J. Comput. Appl. Math., Volume 224, 2009, pages 346–359.

[3] R. Kruse, and P. Przybyłowicz *Approximation of solutions of SDEs with fractional Sobolev regularity.* in preparation

[4] Raphael Kruse, and Yue Wu *Error analysis of randomized Runge-Kutta methods for differential equations with time-irregular coefficients* Comput. Methods Appl. Math., Volume 17, 2017, pages 479–498.

[5] P. Przybyłowicz *On arbitrary slow rate of convergence for randomized Euler scheme.* in preparation

[6] P. Przybyłowicz, and P. Morkisz *Strong approximation of solutions of stochastic differential equations with time-irregular coefficients via randomized Euler algorithm.* Appl. Numer. Math., Volume 78, 2014, pages 80–94.

Approximation Numbers of Embeddings of Anisotropic Sobolev Spaces of Dominating Mixed Smoothness

Winfried Sickel
FSU Jena, Germany, winfried.sickel@uni-jena.de

Thomas Kühn
University of Leipzig, Germany, Thomas.Kuehn@math.uni-leipzig.de

Tino Ullrich
TU Chemnitz, Germany, tino.ullrich@math.tu-chemnitz.de

We investigate the approximation of d-variate periodic functions in anisotropic Sobolev spaces of dominating mixed (fractional) smoothness $\vec{s}$ on the d-dimensional torus, where the approximation error is measured in the L_2-norm.
As it is well-known, in high dimensions functions from isotropic Sobolev spaces $H^s(T^d)$ can not be approximated sufficiently fast (in the sense of approximation numbers of corresponding embeddings). One needs to switch to smaller spaces. Since the beginning of the sixties it is known that Sobolev spaces of dominating mixed smoothness $H^s_{\rm mix}(T^d)$ may help. However, for very large dimensions even these classes are oversized. A way out is to sort the variables in dependence of there importance (in our case in dependence of the smoothness). We associate to each variable different smoothness assumptions. As smoother the function is with respect to the variable x_ℓ as weaker is the influence of this variable. This philosophy is reflected in the choice of the function space $H^{\vec{s}}_{\rm mix}(T^d)$ characterized by the norm

$$\| f \,|H^{\vec{s}}_{\rm mix}(T^d)\| := \Big[\sum_{k\in Z^d} |c_k(f)|^2 \prod_{j=1}^d \big(1+|k_j|\big)^{2s_j}\Big]^{1/2} < \infty\,.$$

We assume $\vec{s}=(s_1,s_2,\dots,s_d)$ and

$$s_1 = s_2 = \ldots = s_\nu < s_{\nu+1} \le s_{\nu+2} \ldots \le s_d$$

for some number $1 \le \nu < d$. It will be the main aim of my talk to describe the behaviour of the approximation numbers

$$a_n(I_d:\ H^{\vec{s}}_{\rm mix}(T^d) \to L_2(T^d))$$

in dependence of $n, \vec{s}, \nu$ and d. Almost all of our results will be based on an elementary lemma, well-known in the literature, simplyfying also our earlier results with respect to this topic. This is joined work with Thomas Kühn (Leipzig) and Tino Ullrich (Bonn).

Positivity preserving schemes for alpha-CIR process

Libo Li
UNSW Sydney, Australia, libo.li@unsw.edu.au

Dai Taguchi
Osaka University, Japan, dai.taguchi.dai@gmail.com

We propose a positivity preserving implicit Euler-Maruyama scheme for a jump-extended Cox-Ingersoll-Ross (CIR) process where the jumps are governed by a compensated spectrally positive α-stable process for $\alpha \in (1, 2)$. Different to the existing positivity preserving numerical schemes for jump-extended CIR or CEV models, the model considered here has infinite activity jumps. We calculate, in this specific model, the strong rate of convergence and give some numerical illustrations. Jump extended models of this type were initially studied in the context of branching processes and was recently introduced to the financial mathematics literature to model sovereign interest rates, power and energy markets.

Monday 15:00–17:00, **Room: CB08.03.005**

Efficient computations for Bayesian inference using Monte Carlo methods

Chair: Matias Quiroz

Efficiently Combining Pseudo Marginal and Particle Gibbs Sampling

David Gunawan
University of New South Wales,

Christopher K. Carter
University of New South Wales,

Robert Kohn
University of New South Wales, r.kohn@unsw.edu.au

Particle Markov Chain Monte Carlo (PMCMC) is a general approach to carry out Bayesian inference in non-linear and non-Gaussian state space models. We show how to scale up PMCMC in terms of the number of parameters and number of time points by generating parameters that are highly correlated with the states with the states integrated out using a pseudo marginal step while the rest of the parameters are generated conditional on the states using particle Gibbs. We make the PMCMC scalable in the number of observations by using the same random numbers in the Metropolis-Hastings ratio of the pseudo marginal step. We do so by expressing the target density of the PMCMC in terms of the basic uniform or standard normal random numbers rather than in terms of the particles, as has been done till now, and develop a constrained version of conditional sequential Monte Carlo algorithm. We illustrate the methods using a high dimensional factor stochastic volatility having both a large number of parameters and a large number of latent states and show that our proposed method makes the computation much more efficient.

Robust Approximate Bayesian Inference with Synthetic Likelihood

Christopher Drovandi
Queensland University of Technology, Australia, c.drovandi@qut.edu.au

David Frazier
Monash University, Australia, david.frazier@monash.edu

Bayesian synthetic likelihood (BSL) [1] is now a well-established method for conducting approximate Bayesian inference in complex models where exact Bayesian approaches are either infeasible, or computationally demanding, due to the intractability of likelihood function. Similar to other likelihood-free methods, such as approximate Bayesian computation, implicit in the application of BSL is the maintained assumption that the data generating process can generate simulated summary statistics that mimic the behaviour of the observed summary statistics. This notion of model compatibility with the observed summaries is critical for the performance of BSL and its variants. We demonstrate theoretically, and through several examples, that if the assumed data generating process (DGP) differs from the true DGP, model compatibility may no longer be satisfied and BSL can give unreliable inferences. To circumvent the issue of incompatibility between the observed and simulated summary statistics, we propose two robust versions of BSL that can deliver reliable performance regardless of whether or not the observed and simulated summaries are compatible. Simulation results and two empirical examples demonstrate the good performance of this robust approach to BSL, and its superiority over standard BSL when model compatibility is not in evidence.

[1] Price, L. F., Drovandi, C. C., Lee, A., and Nott, D. J. *Bayesian synthetic likelihood.* Journal of Computational and Graphical Statistics, 27:1–11, 2018.

Focused Bayesian Prediction

David T. Frazier
Monash University,

Ruben Loaiza-May
Monash University,

Gael Martin
Monash University, gael.martin@monash.edu.au

Bayesian predictive distributions quantify uncertainty about out-of-sample values of a random process conditional only on observed data, where uncertainty about the model-specific parameters is integrated out via the usual probability calculus. Uncertainty about the assumed model itself can, in turn, be accommodated via model-averaging, with the implicit assumption being that the true data generating process (DGP) is contained in the set over which the averaging occurs. We move away from this so-called M-closed world, in which the true DGP is assumed to be either known with certainty, or known to lie in a finite set of models. Herein, we propose a novel method for constructing Bayesian predictive distributions that explicitly acknowledges that practitioners operate in an M -open world. This new approach is not based upon the construction of a conventional Bayesian predictive for a given model, or set of models, but is instead driven by a user-supplied concept of predictive performance loss. To develop such machinery in the Bayesian paradigm, we rely on the principles of approximate Bayesian computation (ABC). A posterior distribution defined over predictive functions is constructed by selecting draws from a prior class of predictives that minimize the loss over a pre-specified evaluation period. We illustrate the potential of our method in a simulation exercise based on misspecified models for a stochastic volatility process, where we find substantial improvements over exact Bayesian methods. Improved forecast accuracy is also a feature of a series of empirical analyses, auguring well for the potential of the new paradigm to reap benefits more broadly.

Pictures and Fear: Some thoughts on convergence

Daniel Simpson
University of Toronto, simpson@utstat.toronto.edu

Do we ever really compute the thing we think we do? And can we ever be sure our code worked? In this talk I'm going to outline some ways we can assess the output of algorithms for computing posterior distributions and suggest that we only have two things that we can trust: pictures and fear.

Monday 15:00–17:00, **Room: CB08.03.002**

Importance Sampling for Robust Model Simulation

Chair: Ad Ridder

Safety evaluation of black-box prediction models via rare-event simulation

Henry Lam
Columbia University, New York, NY, USA, khl2114@columbia.edu

We study the design of good importance samplers to simulate the probability that a black-box predictor, built for instance from machine learning tools such as neural networks and random forests, outputs an exceedingly high prediction value. This problem is motivated from robustness and safety estimations that arise in recent applications of autonomous vehicles and other physical systems. Our approach utilizes mixed integer programming and a "cutting plane" approach to sequentially locate dominant points that guide the tilting of distributions in the importance samplers.

Nonlinear space-filling design and non-parametric importance sampling

Chang-Han Rhee
Northwestern University, Evanston, IL, USA, chang-han.rhee@northwestern.edu

Enlu Zhou
Georgia Institute of Technology, Atlanta, GA, USA,

Peng Qiu
Georgia Institute of Technology, Atlanta, GA, USA,

Traditional space-filling designs for computer experiments aim to fill the parameter space with design points that are as "uniform" as possible. However, the resulting design points may be non-uniform in the model output space failing to provide a reliable representation of the output manifold, and becoming highly inefficient or even misleading in case the computer experiments are non-linear. In this talk, we propose an iterative algorithm [1] that fills in the model output manifold uniformly—rather than the parameter space uniformly—so that one could obtain a reliable understanding of the model behaviors with the minimal number of design points. This problem and the proposed solution has close connection to the rare-event simulation problems and non-parametric importance sampling algorithms. This connection leads to a novel convergence analysis of non-parametric importance sampling algorithms in Wasserstein metric.

[1] C.-H Rhee, E. Zhou, and P. Qiu. Space filling design for non-linear models. arXiv:1710.11616, 2019.

Variance reduction of quantile sensitivities

Ad Ridder
Vrije University, Amsterdam, Netherlands, ad.ridder@vu.nl

Warren Volk-Makarewicz
RWTH Aachen, Germany,

Yijie Peng
Peking University, Beijing, China,

Bernd Heidergott
Vrije University, Amsterdam, Netherlands,

This paper is a study of applying variance reduction techniques in the estimation of quantile sensitivities in simulation models. Sensitivity analysis allows us to determine the main contributing parameter to a simulation model and potential areas of model invalidation. Stochastic gradient estimation is a central approach in sensitivity analysis as the means to compute unbiased gradients in simulation models. Our running example model is a Jackson queueing network with feedback. We consider the computation of gradients of the waiting time quantiles with respect to various design parameters. The variance reduction methods we consider are antithetic variates, control variates, importance sampling, and Latin hypercube sampling. In the study, we will compare three approaches: a quantile sensitivity estimator via the finite difference methods based of the construction in [1], the kernel estimation method in [2], and via the generalized likelihood ratio method of [3].

[1] F. Chu and M. K. Nakayama. Confidence intervals for quantiles when applying variance reduction techniques. *ACM Transactions on Modeling and Computer Simulation 22* (2), p. 10, 2012.

[2] L. J. Hong. Estimating quantile sensitivities. *Operations Research 57* (1), 118Ð-130, 2009.

[3] Y. Peng, M. C. Fu, J.-Q. Hu, and B. Heidergott. A new unbiased stochastic derivative estimator for discontinuous sample performances with structural parameters. *Operations Research 66* (2), 487-Ð499, 2018.

Semi-parametric Dynamic Asymmetric Laplace Models for Tail Risk Forecasting, Incorporating Realized Measures

Richard Gerlach
The University of Sydney, Australia, `richard.gerlach@sydney.edu.au`

The joint Value-at-Risk (VaR) and expected shortfall (ES) quantile regression model of Taylor (2017) is extended, via incorporating a realized measure to drive the tail risk dynamics, as a potentially more efficient driver than daily returns. Further, a new model for the dynamics of the ES component is proposed and tested. Both a maximum likelihood and an adaptive Bayesian Markov Chain Monte Carlo method are employed for estimation, whose properties are compared in a simulation study; results favour the Bayesian approach, subsequently employed in a forecasting study of seven financial market indices. The proposed models are compared to a range of parametric, non-parametric and semi-parametric competitors, including GARCH, Realized GARCH, Extreme Value Theory method and the joint VaR and ES models in Taylor (2017), in terms of accuracy of one-day-ahead VaR and ES forecasts, over a long forecast sample period that includes the global financial crisis in 2007-2008. The results are favorable for the proposed models incorporating a realized measure, especially when employing the sub-sampled Realized Variance and the sub-sampled Realized Range.

Keywords: Realized Variance, Realized Range, Semi-parametric, Markov Chain Monte Carlo, Value-at-Risk, Expected Shortfall.

Tuesday 10:30–12:30, **Room: CB08.03.005**

Stochastic Computation and Complexity II

Chair: Sotirios Sabanis

On the performance of the Euler-Maruyama scheme for SDEs with discontinuous drift coefficient

Thomas Müller-Gronbach
U Passau, Germany, thomas.mueller-gronbach@uni-passau.de

Larisa Yaroslavtseva
U Passau, Germany, Larisa.Yaroslavtseva@uni-passau.de

Recently a lot of effort has been invested to analyze the L_p-error rate of the Euler-Maruyama scheme in the case of stochastic differential equations (SDEs) with a drift coefficient that may have discontinuities in space. For scalar SDEs with a piecewise Lipschitz drift coefficient and a Lipschitz diffusion coefficient that is non-zero at the discontinuity points of the drift coefficient so far only an L_p-error rate of at least $1/(2p)-$ has been proven. In this talk we show that under the latter conditions on the coefficients of the SDE the Euler-Maruyama scheme in fact achieves an L_p-error rate of at least $1/2$ for all $p \in [1, \infty)$ as in the case of SDEs with Lipschitz coefficients. The proof of this result is based on a detailed analysis of appropriate occupation times for the Euler-Maruyama scheme.

A strong order $3/4$ method for SDEs with discontinuous drift coefficient

Thomas Müller-Gronbach
U Passau, Germany, thomas.mueller-gronbach@uni-passau.de

Larisa Yaroslavtseva
U Passau, Germany, Larisa.Yaroslavtseva@uni-passau.de

We study strong approximation of scalar SDEs at the final time in the case when the drift coefficient may have discontinuities in space. We present a method based on finitely many evaluations of the driving Brownian motion that achieves an L_p-error rate of at least $3/4$ for all $p \in [1, \infty)$ under piecewise smoothness assumptions on the coefficients. To obtain this result we show in particular that a quasi-Milstein scheme achieves an L_p-error rate of at least $3/4$ in the case of coefficients that are both Lipschitz continuous and piecewise differentiable with Lipschitz continuous derivatives, which is of interest in itself. The latter error rates are obtained via a detailed analysis of the average size of increments of the time-continuous quasi-Milstein scheme over time-intervals in which the scheme crosses a point of non-differentiability of the coefficients.

Convergence order of the Euler-Maruyama scheme in dependence of the Sobolev regularity of the drift

Andreas Neuenkirch
University of Mannheim, Germany, neuenkirch@kiwi.math.uni-mannheim.de

Michaela Szölgyenyi
University of Klagenfurt, Austria, michaela.szoelgyenyi@aau.at

We study the strong convergence rate of the Euler-Maruyama scheme for scalar SDEs with additive noise and irregular drift. We provide a framework for the error analysis by reducing it to a weighted quadrature problem for irregular functions of Brownian motion. By analysing the quadrature problem we obtain for abitrarily small $\epsilon > 0$ a strong convergence order of $(1 + \kappa)/2 - \epsilon$ for a non-equidistant Euler-Maruyama scheme, if the drift has Sobolev-Slobodeckij-type regularity of order $\kappa \in (0, 1)$.

Derivative-free Mistein scheme for strong approximation of solutions of SDEs in analytic noise model

Paweł Morkisz
AGH University of Science and Technology, Krakow, Poland, pprzybyl@agh.edu.pl

We consider approximate solving of the following scalar SDE

$$\begin{cases} dX(t) = a(t, X(t))dt + b(t, X(t))dW(t), \quad t \in [0, T], \\ X(0) = \eta, \end{cases} \tag{7.2}$$

driven by a standard one-dimensional Wiener process $W = (W(t))_{t\in[0,T]}$.

Inspired by increasing popularity of computations with low precision (used on Graphics Processing Units -GPUs and standard Computer Processing Units -CPUs), we introduce a suitable analytic noise model of standard noisy information about a and b. In this model we show that, under standard assumptions on a and b, the upper bounds on the error of the derivative-free Milstein scheme is proportional to $n^{-1} + \delta_1 + \delta_2$, where n is a number of noisy evaluations of a and b and $\delta_1, \delta_2 \geq 0$ are precision parameters for values of a and b, respectively. We report numerical experiments performed on both CPU and GPU that confirm our theoretical findings. We also present some computational performance comparison between those two architectures.

This is a joint work with Paweł Przybyłowicz (AGH UST, Krakow, pprzybyl@agh.edu.pl)

[1] R. Kruse, and Y. Wu *A randomized Milstein method for stochastic differential equations with non-differentiable drift coefficients* Discrete and Continuous Dynamical Systems Series B, Volume 22, 2017, doi: 10.3934/dcdsb.2018253

[2] P. M. Morkisz, P. Przybyłowicz, *Optimal pointwise approximation of SDE's from inexact information*, Journal of Computational and Applied Mathematics, Volume 324, 2017, 85–100.

[3] A. Kałuża, P. M. Morkisz, and P. Przybyłowicz, *Optimal approximation of stochastic integrals in analytic noise model.* to appear in Applied Mathematics and Computation, https://doi.org/10.1016/j.amc.2019.03.022.

Tuesday 10:30–12:30, **Room: CB08.02.002**

Monte Carlo Methods for Spatial Problems

Chair: Shev MacNamara

Spatial Cauchy Processes with Local Tail Dependence

Pavel Krupskiy
University of Melbourne, Australia, pavel.krupskiy@unimelb.edu.au

Raphaël Huser
KAUST, Saudi Arabia, raphael.huser@kaust.edu.sa

We study a class of models for spatial data obtained using Cauchy convolution processes with random indicator kernel functions. We show that the resulting spatial processes have some appealing dependence properties including tail dependence at smaller distances and asymptotic independence at larger distances. We derive extreme-value limits of these processes and consider some interesting special cases. We show that estimation is feasible in high dimensions and the proposed class of models allows for a wide range of dependence structures.

Computing the Australian Cancer Atlas: Getting it 'just right'

Susanna Cramb
Cancer Council Queensland, Australia, susannacramb@cancerqld.org.au

Earl Duncan
ACEMS Queensland University of Technology (QUT), Australia, earl.duncan@qut.edu.au

Peter Baade
Cancer Council Queensland, Australia, peterbaade@cancerqld.org.au

Kerrie Mengersen
ACEMS QUT, Australia, k.mengersen@qut.edu.au

The online Australian Cancer Atlas (atlas.cancer.org.au) revealed the patterns of small-area estimates of incidence and survival for many types of cancer. The atlas was focused on providing robust, reliable estimates and showing the uncertainty around these estimates. To achieve this, Bayesian spatial hierarchical model estimates were computed using MCMC: the Metropolis adjusted Langevin algorithm for incidence (via the R package CARBayes), and a slice-sampling algorithm for survival (via WinBUGS). The challenges involved determining the preferred Bayesian spatial model, considering if the spatial smoothing was appropriate, and deciding which computational approach was preferred. Despite examining 2148 areas often with sparse data over 50 combinations of cancer types and sexes, the computational time required for MCMC was feasible. Other advantages of MCMC included the simplicity of calculating posterior probabilities and densities, which were used in reporting quantitative estimates of uncertainty. MCMC remains a viable and useful approach for Bayesian small-area cancer atlases.

Multi-Scale Process Modelling of Massive Spatial Datasets

Andrew Zammit-Mangion
University of Wollongong, Australia, azm@uow.edu.au

Jonathan Rougier
University of Bristol, UK, j.c.rougier@bristol.ac.uk

Recent years have seen a huge development in spatial modelling and prediction methodology, driven by the increased availability of remote-sensing data and the reduced cost of distributed-processing technology. It is well known that modelling and prediction using infinite-dimensional process models is not possible with large data sets, and that both approximate models and, often, approximate-inference methods, are needed. The problem of fitting simple global spatial models to large datasets has been solved through the likes of multi-resolution approximations and nearest-neighbour techniques. Here we tackle the next challenge, that of fitting complex, nonstationary, multiscale, models to large datasets. We propose doing this through the use of superpositions of spatial processes with decreasing spatial scale and increasing degrees of nonstationarity. Computation is facilitated through the use of Gaussian Markov random fields and parallel Markov chain Monte Carlo based on graph colouring. The resulting model allows for both distributed computing and distributed data. Importantly, it provides opportunities for valid model and data scalability and yet is still able to borrow strength across large spatial scales. We compare our approach to state-of-the-art spatial modelling and prediction methods and we illustrate a two-scale version on a dataset of sea-surface temperature containing on the order of one million observations.

Bayesian Indirect Likelihood for the Potts Model

Matt Moores
NIASRA, University of Wollongong, Australia, mmoores@uow.edu.au

Geoff Nicholls
University of Oxford, UK, nicholls@stats.ox.ac.uk

Anthony Pettitt
Queensland University of Technology, Australia, a.pettitt@qut.edu.au

Kerrie Mengersen
Queensland University of Technology, Australia, k.mengersen@qut.edu.au

The Potts model is commonly used for classification, where the labels are spatially-correlated. The strength of spatial association is governed by a smoothing parameter β, known as the inverse temperature. A difficulty arises from the dependence of an intractable normalising constant on the value of β, thus there is no closed-form solution for sampling from the posterior distribution directly. There are a variety of Markov chain Monte Carlo methods for sampling from the posterior without evaluating the normalising constant, including the exchange algorithm and approximate Bayesian computation (ABC). A serious drawback of these algorithms is that they do not scale well for models with a large state space, such as images with a million or more pixels. In this talk, I will introduce the parametric functional approximate Bayesian (PFAB) algorithm [1], which uses an integral curve to approximate the score function of the Potts model. PFAB incorporates known properties of the likelihood, such as heteroskedasticity and critical temperature. I will demonstrate this method using synthetic data as well as remotely-sensed imagery from the Landsat-8 satellite. The proposed algorithm achieves up to a hundredfold improvement in the elapsed runtime, compared to the exchange algorithm or ABC. An open source implementation of PFAB is available in the R package '**bayesImageS**' [2].

[1] M. T. Moores, G. K. Nicholls, A. N. Pettitt, and K. Mengersen. Scalable Bayesian Inference for the Inverse Temperature of a Hidden Potts Model. *Bayesian Analysis*, in press, 2018.

[2] M. T. Moores, D. Feng, and K. Mengersen. bayesImageS: Bayesian Methods for Image Segmentation using a Potts Model. R package version 0.5-3, 2018.

[3] M. T. Moores, C. C. Drovandi, K. Mengersen, and C. P. Robert. Pre-Processing for Approximate Bayesian Computation in Image Analysis. *Statistics & Computing* **25**(1): 23–33, 2015.

Tuesday 10:30–12:30, **Room: CB08.03.004**

Simulation Optimization

Chair: Raghu Pasupathy

Confidence Region for Model Parameters in Stochastic Gradient Descent via Batch Means

Yi Zhu
Northwestern University, USA, yichen2016@u.northwestern.edu

Jing Dong
Columbia University, USA, jing.dong@gsb.columbia.edu

Stochastic Gradient Descent (SGD) and variants of it have been widely used in model parameters estimation for either online learning or when data sizes are very large. As the estimators we constructed via SGD is not exact, it is desirable to be able to quantify the estimation error incurred. While most there is a rich literature studying convergence rate of the objective function or the parameter estimation error based on SGD, much less is known about the statistical inference for true model parameters. Recently, [1] and [2] propose novel techniques to quantify uncertainty for model parameters in SGD. Following this line of work, in this paper, we propose a simple procedure to construct asymptotically valid confidence regions for model parameters based on a cancellation method known as the batch means. The confidence region we constructed can accommodate multi-dimensional joint inference, which takes the covariance structure of the estimators into account.

[1] X. Chen, J. D. Lee, X. T. Tong and Y. Zhang *Statistical Inference for Model Parameters in Stochastic Gradient Descent.* Annals of Statistics, to appear, 2019.

[2] W. Su and Y. Zhu *Uncertainty Quantification for Online Learning and Stochastic Approximation via Hierarchical Incremental Gradient Descent.* working paper, 2018.

Efficient Iterative Schemes for Optimal Transport Distributionally Robust Optimization

José Blanchet
Stanford University, USA, jose.blanchet@gmail.com

Karthyek Murthy
SUTD, SG, karthyek@gmail.com

Fan Zhang
Stanford University, USA, fzh@stanford.edu

We consider optimal transport based distributionally robust optimization (DRO) problems with locally strongly convex transport cost functions and affine decision rules. Under conventional convexity assumptions on the underlying loss function, we obtain structural results about the value function, the optimal policy, and the worst-case optimal transport adversarial model. These results expose a rich structure embedded in the DRO problem (e.g., strong convexity even if the non-DRO problem was not strongly convex, a suitable scaling of the Lagrangian for the DRO constraint, etc. which are crucial for the design of efficient algorithms). As a consequence of these results, one can develop optimization procedures which have the same sample and iteration complexity as a natural non-DRO benchmark algorithm such as stochastic gradient descent; and sometimes even better complexity. Our analysis provides insights into the fine structure and convexity properties of the DRO value function, the optimal policy, and the worst-case optimal transport adversarial model.

Risk-sensitive Inference

Harsha Honnappa
Purdue University, School of Industrial Engineering, USA, honnappa@purdue.edu

We consider a single-stage stochastic optimization problem, where the stochastic variable is only known through a parametrized predictive model. Operating in a Bayesian setting, we study the interplay between posterior estimation and decision-making through a "risk-sensitive" lens by leveraging dual representations of convex risk measures. We prove finite sample probabilistic bounds on an oracle regret function, and illustrate these results through simulations on an example problem. Time permitting, we will discuss extensions of the methodology to non-Bayesian and multi-stage problems.

Retrospective approximation for Stochastic Optimization

Raghu Pasupathy
Purdue University, Department of Statistics, USA, pasupath@purdue.edu

David Newton
Purdue University, Department of Statistics, USA, newton34@purdue.edu

Aaron Nung Kwan Yip
Purdue University, Department of Mathematics, USA, yipn@purdue.edu

The Stochastic Approximation (SA) recursion $X_{k+1} = X_k - \gamma_k \nabla F(X_k), k = 1, 2, \ldots$, also known as Stochastic Gradient Descent (SGD), is the "workhorse" stochastic optimization recursion for machine learning. For smooth, strongly convex objectives, it is well-known that the SA iteration achieves the information theoretic Cramér-Rao lower bound (not to be confused with the more common but crude optimal rate $\mathcal{O}(k^{-1/2})$) when the step size is "optimally" chosen. However, such optimal choice depends on unknown curvature constants of the objective function; moreover, incorrect choice can lead to a severe deterioration in convergence rates. There is strong indication that this situation is not merely "theoretical" but manifests frequently in practice because the eigen values of the Hessian at the optimum are rarely known in advance. As a remedy to this impasse, we present Retrospective Approximation (RA), where individual sample-paths generated (implicitly) with a chosen sample size are solved to a chosen error tolerance. We demonstrate that by keeping the sample size and the error tolerance in a specified relation to each other, convergence rates nearly identical to those attained by "optimal" SGD can be achieved. Most importantly, RA's structure lends itself to the use of the best deterministic solvers, e.g., L-BFGS with line search and the use of conjugated gradients and Hessian vector products, thereby seamlessly combining practical performance with near optimal asymptotic convergence rates.

Tuesday 10:30–12:30, **Room: CB08.03.002**

QMC Theory and Applications I

Chair: Vesa Kaarnioja

A Quasi-Monte Carlo Method for an Optimal Control Problem

Philipp Guth
University of Mannheim, Germany, pguth@mail.uni-mannheim.de

Vesa Kaarnioja
UNSW Sydney, Australia, v.kaarnioja@unsw.edu.au

Frances Kuo
UNSW Sydney, Australia, f.kuo@unsw.edu.au

Claudia Schillings
University of Mannheim, Germany, c.schillings@uni-mannheim.de

Ian Sloan
UNSW Sydney, Australia, i.sloan@unsw.edu.au

In this work we apply a quasi-Monte Carlo (QMC) method to an optimal control problem constrained by an elliptic partial differential equation (PDE) equipped with an uncertain diffusion coefficient. In particular the optimization problem is to minimize the expected value of a tracking type cost functional with an additional penalty on the control. The uncertain coefficient in the PDE is parametrized by a countably infinite number of terms via a Karhunen–Loève expansion (KLE) and the expected value is considered as an infinite-dimensional integral in the corresponding parameter space.

We discretize the optimization problem by truncating the KLE after s terms, approximating the solution of the PDE using finite elements (FE) and approximating the expected value by an n-point QMC rule in s dimensions. It is shown that the discretization error of the solution to the optimization problem is bounded by the discretization error of the adjoint state. For the convergence analysis the latter is decomposed into truncation error, FE error and QMC error, which are then analysed separately. Numerical experiments confirming our theoretical convergence results will be presented.

An analogue of QMC integration over a finite group: difference sets and its generalization

Hiroki Kajiura
Graduate School of Science, Hiroshima University, Japan, hikajiura@hiroshima-u.ac.jp

Let G be a finite group, f a complex function on G, and X a subset of G. We study an analogue of quasi-Monte Carlo (QMC) integration over G. Let us put $v := \#G$, $k := \#X$. We define the average $I(f) := (1/v)\sum_{g\in G} f(g)$ of f over G as an analogue of the integration, and then the average $I_X(f) := (1/k)\sum_{x\in X} f(x)$ over X as an analogue of QMC approximation. Our aim is to minimize the integration error

$$\mathrm{Err}(f;X) := |I(f) - I_X(f)|$$

for a fixed k and G by choosing a "good" subset X.

When G has a good structure (e.g. a d-dimensional torus), there are large amount of researches. Here, we only consider an abstract finite group G. We lack an appropriate replacement of "frequencies."

Instead, we use representation theory of finite groups. Let $\widehat{G}$ be the set of isomorphism classes of irreducible representation of G. Then, the space of complex functions on G decomposes into a direct sum of H_ρ for $\rho \in \widehat{G}$, and hence any function f decomposes to $\sum_{\rho\in\widehat{G}} f^\rho$, where f^ρ's are mutually orthogonal with respect to the standard inner product. Let I_X be the normalized characteristic function defined by $I_X(x) = 1/k$ if $x \in X$, and $I_X(x) = 0$ otherwise. We define $V(f)$ (an analogue of variance) by $\sum_{\rho\neq 1_G}(\dim\rho||f^\rho||)$, and $D(X)$ (an analogue of discrepancy) by $\max_{\rho\neq 1_G}(||I_X^\rho||/\dim\rho)$. Then we have a Koksma-Hlawka type inequality:

$$\mathrm{Err}(f;X) \le V(f)D(X).$$

The strange weight using $\dim\rho$ is chosen so that the following properties hold. We are not sure whether the bound is practical or not. We show a lower bound $D(X) \ge \sqrt{\frac{1/k-1/v}{v-1}}$ depending only on k and v. We show that the equality holds if and only if X is a *pre-difference set*, which we define here. For any $a \in G$, denote its conjugacy class by $[a] \subset G$. We define $\lambda_{[a]}$ by $\#\{(x,y) \in X^2 \mid x^{-1}y \in [a]\}/\#[a]$. X is a pre-difference set if the value $\lambda_{[a]}$ is independent of the choice of a, except $a = e$. This notion generalizes difference set[1] in combinatorics, and they coincide if G is commutative. (We could not find the notion of pre-difference set in the literatures.) The dihedral group of order 16 has no "non-trivial" difference set[2], but we find non-trivial pre-difference set in this group, using a computer.

[1] R. H. Bruck. *Difference sets in a finite group.* Amer. Math. Soc. 78, 1955, pp. 464-481.

[2] Robert E. Kibler. *A summary of noncyclic difference sets, $k < 20$.* J. Combinatorial Theory Ser. A 25.1, 1978, pp. 62-67.

On the Use of Quasi-Monte Carlo in Quantile Estimation

Zhijian He
South China University of Technology, China, `hezhijian@scut.edu.cn`

Quantiles are usually used to measure risk of stochastic systems. We consider quantile estimation based on quasi-Monte Carlo (QMC) methods, which have the potential to improve the rate of Monte Carlo. We first prove the convergence of QMC-based estimates under very mild conditions. We then establish a deterministic error bound of $O(N^{-1/d})$ for the QMC estimates, where d is the dimension of the QMC point sets used in the simulation and N is the sample size.

Simulating empirical copulas: Empirical comparison MC and QMC

Hozumi Morohosi
National Graduate Insititute for Policy Studies, Tokyo, Japan, `morohosi@grips.ac.jp`

This work proposes a simulation method to generate a point set which follows an empirical copula induced by a given high-dimensional sample. Copula has been a generic modeling tool for the dependence between variables, which is getting increasingly popular among simulation people [1]. While most of research on copula seem to be made around parametric models, nonparametric models would be of their own use for several applications. We consider a simulation method based on empirical beta copula and empirical Bernstein copula [2]. For those models, although conditional distribution inversion method is available due to their simple structure, we usually have a great difficulty in obtaining the inversion of conditional distribution component by component successively in a high dimension. Assuming exchangeability between variables leads us to a method for dealing with this problem. Some computational experiments show our method works fairly well. We also report a comparison of MC and QMC for our method.

[1] M. Cambou, M. Hofert, and C. Lemieux. Quasi-random numbers for copula models. *Statistics and Computing*, 27(2017), 1307–1329.

[2] J. Segers, M. Sibuya, and H. Tsukahara. The empirical beta copula. *Journal of Multivariate Analysis*, 155(2017), 35–51.

Tuesday 15:00–17:00, **Room: CB08.02.002**

Monte Carlo Algorithms in Bayesian Statistics

Chair: Ziwen An

Efficient Bayesian Estimation for GARCH-type Models via Sequential Monte Carlo

Dan Li
QUT Brisbane, Australia, d31.li@hdr.qut.edu.au

Christopher Drovandi
QUT Brisbane, Australia, c.drovandi@qut.edu.au

Adam Clements
QUT Brisbane, Australia, a.clements@qut.edu.au

The popularity of Bayesian statistics has increased in recent years because it provides principled uncertainty quantification of model parameter estimates and model predictions. Sequential Monte Carlo (SMC) is an efficient method for sampling from the posterior distribution in Bayesian statistics. SMC may be preferred over the popular Markov chain Monte Carlo (MCMC) for certain classes of problems, since SMC is easily parallelisable, adaptable and more effective in representing complex posterior distributions. However, the useful properties of Bayesian statistics and SMC are yet to be fully realised in certain areas of Econometrics. We exploit the advantages of SMC to develop efficient parameter estimation, model prediction and selection methods for GARCH (Generalized AutoRegressive and Conditional Heteroskedasticity) models and their extensions. This approach provides an alternative method for quantify estimation uncertainty relative to standard asymptotic assumptions typically made with classical statistical inferences. We demonstrate that even with long time series, the posterior distribution for our models of interest can be far from normally distributed, highlighting the need for a Bayesian approach and an efficient posterior sampling method. We also make contributions to the SMC literature by exploring the most efficient way of constructing the sequence of distributions in SMC for long time series data. This is joint work with Associate Professor Christopher Drovandi and Professor Adam Clements.

Barkers algorithm for Bayesian inference in intractable likelihood models

Krys Latuszynski
University of Warwick, UK, K.G.Latuszynski@warwick.ac.uk

Flavio Gonçalves
Universidade Federal de Minas Gerais, Brasil, fbgoncalves@est.ufmg.br

Gareth Roberts
University of Warwick, UK, G.O.Roberts@warwick.ac.uk

Dootika Vats
University of Warwick, UK, D.Vats@warwick.ac.uk

Constructing MCMC algorithms for Bayesian inference in intractable likelihood models is problematic since difficulties in evaluating the likelihood make the application of the standard Metropolis-Hastings acceptance formula impossible. This enforces either (1) approximate methods that introduce bias of unknown magnitude, or (2) the pseudomarginal approach that is exact, but slows down MCMC convergence, sometimes dramatically. In this talk I will present a new approach (3) based on unbiased estimators of the likelihood and the Barkers acceptance ratio. The approach is exact and retains the per iteration convergence rate comparable to that of the standard Metropolis-Hastings. I will illustrate the approach with examples of exact inference for stochastic differential equations.

Subsampling Sequential Monte Carlo for Static Bayesian Models

David Gunawan
UNSW Sydney, ACEMS, david.gunawan40@gmail.com

Robert Kohn
UNSW Sydney, ACEMS, r.kohn@unsw.edu.au

Matias Quiroz
UNSW Sydney, ACEMS, quiroz.matias@gmail.com

Khue-Dung Dang
UNSW Sydney, ACEMS, doankhuedung.dang@unsw.edu.au

Minh-Ngoc Tran
University of Sydney, ACEMS, minh-ngoc.tran@sydney.edu.au

Our article shows how to carry out Bayesian inference by combining data subsampling with Sequential Monte Carlo (SMC). This takes advantage of the attractive properties of SMC for Bayesian computations with the ability of subsampling to tackle big data problems. SMC sequentially updates a cloud of particles through a sequence of densities, beginning with a density that is easy to sample from such as the prior and ending with the posterior density. Each update of the particle cloud consists of three steps: reweighting, resampling, and moving. In the move step, each particle is moved using a Markov kernel and this is typically the most computationally expensive part, especially when the dataset is large. It is crucial to have an efficient move step to ensure particle diversity. Our article makes two important contributions. First, in order to speed up the computation, we use an approximately unbiased and efficient annealed likelihood estimator based on data subsampling. The subsampling approach is more memory efficient than the corresponding full data SMC, which is a great advantage for parallel computation. Second, we use a Metropolis within Gibbs kernel with two conditional updates. First, a Hamiltonian Monte Carlo update makes distant moves for the model parameters. Second, a block pseudo-marginal proposal is used for the particles corresponding to the auxiliary variables for the data subsampling. We demonstrate the usefulness of the methodology through a series of examples.

Robust Bayesian Synthetic Likelihood via a Semi-Parametric Approach

Ziwen An
Queensland University of Technology, ziwen.an@hdr.qut.edu.au

David J. Nott
National University of Singapore, standj@nus.edu.sg

Christopher Drovandi
Queensland University of Technology, c.drovandi@qut.edu.au

Bayesian synthetic likelihood (BSL) [1] is now a well established method for performing approximate Bayesian parameter estimation for simulation-based models that do not possess a tractable likelihood function. BSL approximates an intractable likelihood function of a carefully chosen summary statistic at a parameter value with a multivariate normal distribution. The mean and covariance matrix of this normal distribution are estimated from independent simulations of the model. Due to the parametric assumption implicit in BSL, it can be preferred to its non-parametric competitor, approximate Bayesian computation, in certain applications where a high-dimensional summary statistic is of interest. However, despite several successful applications of BSL, its widespread use in scientific fields may be hindered by the strong normality assumption. Here we develop a semi-parametric approach to relax this assumption to an extent and maintain the computational advantages of BSL without any additional tuning. We test our new

method, semiBSL, on several challenging examples involving simulated and real data and demonstrate that semiBSL can be significantly more robust than BSL and another approach in the literature. The new method is implemented with the R package BSL, which is available on Comprehensive R Archive Network (CRAN).

[1] Price, L. F., Drovandi, C. C., Lee, A., and Nott, D. J. *Bayesian synthetic likelihood*. Journal of Computational and Graphical Statistics, 27:1–11, 2018.

Tuesday 15:00–17:00, **Room: CB08.03.002**

QMC Theory and Applications II

Chair: Ian Sloan

A Numerical Study of the $R_d(\phi_d)$ Sequence

Jordan Van
UNSW Sydney, Australia, jordan.van@student.unsw.edu.au

Frances Kuo
UNSW Sydney, Australia, f.kuo@unsw.edu.au

Vesa Kaarnioja
UNSW Sydney, Australia, v.kaarnioja@unsw.edu.au

Recently, a new low discrepancy sequence was introduced in [1]. It is argued in [1] that the new $R_d(\phi_d)$ sequence has superior approximation properties across a variety of problems such as integration and distributing points evenly on a sphere. In this talk, we investigate the performance of $R_d(\phi_d)$ for a variety of high dimensional integrals, such as those that arise in finance, in comparison to more well-known point sets. In addition, we investigate the merits of the minimum packing distance used in [1] as a measure of discrepancy.

[1] Martin Roberts. The unreasonable effectiveness of quasirandom sequences. http://extremelearning.com.au/unreasonable-effectiveness-of-quasirandom-sequences/, 2018.

Uncertainty Quantification using Periodic Random Variables and Lattice Quasi-Monte Carlo Rules

Vesa Kaarnioja
UNSW Sydney, Australia, v.kaarnioja@unsw.edu.au

Frances Kuo
UNSW Sydney, Australia, f.kuo@unsw.edu.au

Ian Sloan
UNSW Sydney, Australia, i.sloan@unsw.edu.au

Many studies in uncertainty quantification have been carried out under the assumption of an input random field in which a countable number of independent random variables are each uniformly distributed on an interval, with these random variables entering linearly in the input random field – the so-called "affine" model. In this talk we propose an alternative model of the random field, in which the random variables have the same uniform distribution on an interval, but the random variables enter the input field as periodic functions. The field is constructed in such a way as to have the same mean and covariance function as the affine random field. Higher moments differ from the affine case, but in general the periodic model seems no less desirable. The periodicity is beneficial when used in association with "lattice" quasi-Monte Carlo rules: we naturally obtain higher order convergence rates!

Uncertainty Quantification for Capacitors with Uncertain Boundaries Using QMC

Harri Hakula
Aalto University, Finland, harri.hakula@aalto.fi

Vesa Kaarnioja
UNSW Sydney, Australia, v.kaarnioja@unsw.edu.au

Frances Kuo
UNSW Sydney, Australia, f.kuo@unsw.edu.au

In engineering applications, it is important to account for the nonlinear effect that manufacturing imperfections have on the material properties of electronic components. One way to carry out an analysis of this kind is to utilize techniques used in the field of uncertainty quantification and model these imperfections as domain uncertainties parametrized by using a Karhunen–Loève expansion.

In this talk, we consider uncertainty quantification for capacitors equipped with an uncertain boundary. The mathematical model for the capacitor can be derived from Maxwell's equations and, by assuming a homogeneous medium within the capacitor, the mathematical model can be further reduced into a mixed Dirichlet–Neumann problem for the Laplacian. Letting u denote the harmonic electric potential within a planar capacitor D, the capacitance of D is proportional to the modulus

$$\mathrm{M}[D] = \int_D |\nabla u|^2 \,\mathrm{d}x.$$

We take the modulus as our quantity of interest and consider its expectation computed over the parameter space describing the boundary variations. We provide an error estimator for the deterministic PDE problem and carry out a regularity analysis for the moduli computed over the parameter space, which justifies the use of quasi-Monte Carlo methods for computing the expectation. The theoretical results are assessed in an ensemble of numerical experiments.

Kernel-Based Lattice Point Interpolation for Uncertainty Quantification using Periodic Random Variables

Vesa Kaarnioja
UNSW Sydney, Australia, v.kaarnioja@unsw.edu.au

Yoshihito Kazashi
EPFL, Lausanne, Switzerland,

Frances Kuo
UNSW Sydney, Australia, f.kuo@unsw.edu.au

Fabio Nobile
EPFL, Lausanne, Switzerland,

Ian Sloan
UNSW Sydney, Australia, i.sloan@unsw.edu.au

In this talk (related to one by Frances Kuo) we use kernel approximation based on periodic random variables and lattice rules to approximate the solution of an elliptic PDE with random field as input. The method is shown to give good (but not optimal) rates of convergence independently of truncation dimension, while being simple and cheap to implement.

Tuesday 15:00–17:00, **Room: CB08.03.004**

Applied Probability and Monte Carlo Methods I

Chair: Chang-Han Rhee

Wasserstein-based Distributionally Robust Performance Analysis with Wasserstein Martingale Constraints

Jose Blanchet
Stanford University, USA, jblanche@stanford.edu

Peter Glynn
Stanford University, USA, glynn@stanford.edu

Zhengqing Zhou
Stanford University, USA, zqzhou@stanford.edu

We study computational algorithms for distributionally robust expectations among all martingales which lie within a prescribed tolerance (in Wasserstein distance) of a baseline martingale model. We obtain rates of convergence showing that the solutions to these types of problems can be computed with a canonical sample complexity rate (square-root of the sample size). This is joint work with Peter Glynn and Zhengqing Zhou.

Validating Optimization with Probabilistic Constraints

Henry Lam
Columbia University, USA, henry.lam@columbia.edu

Huajie Qian
Columbia University, USA, h.qian@columbia.edu

We consider optimization with stochastic or probabilistic constraints under the availability of limited data or Monte Carlo samples. In this situation, the obtained solutions are subject to statistical noises that affect both the feasibility and the objective performance. Common approaches in data-driven optimization impose constraint reformulations that are "safe" enough to ensure solution feasibility with high confidence. Often times, selecting this safety margin relies on loose statistical estimates, in turn leading to overly conservative and suboptimal solutions. We investigate a validation-based framework to balance the feasibility-optimality tradeoff more efficiently, by leveraging the typical low-dimensional structure of solution paths in these data-driven reformulations. We demonstrate how our approach can lead to feasible solution with nearly dimension-free safety adjustment and confidence guarantees.

On Counting Distributions for Mortality Projections

Ad Ridder
Vrije Universiteit Amsterdam , Netherlands, ad.ridder@vu.nl

Shaul K. Bar-Lev
University of Haifa, Israel, barlev@stat.haifa.ac.il

In this paper we consider counting distributions $\{p(n), n = 0, 1, \ldots\}$ that belong to the class of natural exponential families (NEF's), meaning that they are parameterized by

$$p(n) = \mu(n)e^{\theta n - \kappa(\theta)}, \quad n = 0, 1, \ldots.$$

Here, $\{\mu(n), n = 0, 1, \ldots\}$ are nonnegative numbers called the kernel of the NEF. Denote m for the mean, and V for the variance of the NEF distribution. Then we consider distributions having variances as functions of the mean of the form

$$V(m) = m\left(1 + \frac{m}{p}\right)^r, \quad r = 0, 1, \ldots.$$

here, p is the dispersion parameter. The class of these distributions is called the ABM class after [1]. The special cases are $r = 0$ which gives the Poisson, $r = 1$ gives negative binomial, and $r = 2$ gives generalized Poisson (also known as Abel). In all these special cases, the kernels and the corresponding probabilities are exactly computable. However, this does not hold for $k \geq 3$. In [1] it is shown that the members of the ABM class predict better than the Poisson the mortality rates of elderly age people. In this paper we utilize the machinery of [2] to construct a numerical method for computing and sampling from the ABM class, also for $k \geq 3$.

[1] Awad, Y., Bar-Lev, S.K. and Makov, U. "A new class counting distributions embedded in the Lee-Carter model for mortality projections: A Bayesian approach" *Technical Report.* Actuarial Research Center, University of Haifa, Israel. 2016.

[2] G. Letac, and M. Mora. "Natural real exponential families with cubic variance functions." *The Annals of Statistics 18*(1), 1–37. 1990.

Central Limit Theorems for Coupled Particle Filters

Ajay Jasra
National University of Singapore, SG, staja@nus.edu.sg

In this talk we give new central limit theorems (CLT) for several coupled particle filters (CPFs). CPFs are used for the sequential estimation of the difference of expectations w.r.t. filters which are in some sense close. Examples include the estimation of the filtering distribution associated to different parameters (finite difference estimation) and filters associated to partially observed discretized diffusion processes (PODDP) and the implementation of the multilevel Monte Carlo (MLMC) identity. We develop new theory for CPFs and based upon several results, we propose a new CPF which approximates the maximal coupling (MCPF) of a pair of predictor distributions. In the context of ML estimation associated to PODDP with discretization Δ_l we show that the MCPF and the approach in [1] have, under assumptions, an asymptotic variance that is upper-bounded by an expression that is (almost) $\mathcal{O}(\Delta_l)$, uniformly in time. The $\mathcal{O}(\Delta_l)$ rate preserves the so-called forward rate of the diffusion in some scenarios which is not the case for the CPF in [2].

[1] Jasra, A., Ballesio, M., Von Schwerin, E., & Tempone, R. (2018). A coupled particle filter for multilevel estimation. Technical Report.

[2] Jasra, A., Kamatani, K., Law K. J. H. & Zhou, Y. (2017). Multilevel particle filters. *SIAM J. Numer. Anal.*, **55**, 3068-3096.

Tuesday 15:00–17:00, **Room: CB08.03.005**

Stochastic Computation and Complexity III

Chair: Thomas Müller-Gronbach

Integration by parts formulae for killed processes and their unbiased Monte Carlo simulation.

Noufel Frikha
Paris Diderot University, France, frikha@math.univ-paris-diderot.fr

Arturo Kohatsu-Higa
Ritsumeikan University, Japan, khts00@fc.ritsumei.ac.jp

Libo Li
UNSW Sydney, Australia, libo.li@unsw.edu.au

In this presentation, we provide new probabilistic representations for two integration by parts formulas, one being of Bismut-Elworthy-Li's type, for the marginal law of a one-dimensional uniformly elliptic diffusion process killed at a given level. These formulas are established by combining a Markovian perturbation argument with a tailor-made Malliavin calculus for the underlying Markov chain structure involved in the probabilistic representation of the original marginal law. Among other applications, an unbiased Monte Carlo path simulation method for both integration by parts formula stems from the previous probabilistic representations.

Complexity of stochastic integration in fractional Sobolev classes

Stefan Heinrich
University of Kaiserslautern, Germany, heinrich@informatik.uni-kl.de

Inspired by a recent paper of Eisenmann and Kruse [1] we study the complexity of stochastic integration $\int_0^1 f(t)dW_t$ of stochastic functions f with fractional Sobolev regularity $W_p^r([0,1])$ in t. We consider all $0 < r < \infty$ and $1 \le p \le \infty$. Information about f consists of function values while that about W_t may be function values or scalar products with polynomials of a given degree. Both deterministic and randomized algorithms are considered. We determine the order of the complexity, which includes finding and analyzing algorithms of optimal order and proving matching lower bounds. This extends results from [1], who considered the case $0 < r < 2$ and did not establish lower bounds, and from [2], where only deterministic integrands were considered.

[1] Monika Eisenmann, Raphael Kruse. Two quadrature rules for stochastic Itô-integrals with fractional Sobolev regularity. arXiv:1712.08152

[2] S. Heinrich, Complexity of stochastic integration in Sobolev classes, J. Math. Anal. Appl. (2019), https://doi.org/10.1016/j.jmaa.2018.12.077.

The role of the Eikonal equation in numerics for bounded SDEs

Francisco Bernal
Carlos III University of Madrid, Spain, francisco.bernal@uc3m.es

The accurate numerical integration of spatially bounded SDEs is challenging owing to the interaction of discretised chains with the boundary. With the Brownian motion, the efficiency of the Walk on Spheres method is affected by trajectories approaching too slowly to absorbing surfaces, while lingering off reflected ones: exactly the opposite to what would be wished. With general SDEs, the presence of boundaries spoils weak rates of convergence to $\mathcal{O}(h^{1/2})$ and worse. In particular, boundary corners can typically only be handled with ad hoc procedures, if at all.

In this talk, we will demonstrate how many of these difficulties can be overcome by the simple expedient of extracting the full information hidden in the distance map which describes the domain of integration. This idea is underpinning two novel schemes which will be discussed: a first-ever implementation of Milstein's method for bounded diffusions [1, chapter 5], recently published [2]; and a new version of the Walk on Spheres for domains with partially planar surfaces, currently in progress.

[1] G. N. Milstein and M. V. Tretyakov. *Stochastic Numerics for Mathematical Physics.* Springer Berlin, 2004.

[2] F. Bernal *An implementation of Milstein's method for general bounded diffusions.* Journal of Scientific Computing, 1-24 2018. (https://doi.org/10.1007/s10915-018-0884-6)

How good is random information?

Aicke Hinrichs
JKU Linz, Austria, `aicke.hinrichs@jku.at`

David Krieg
JKU Linz, Austria, `david.krieg@jku.at`

Erich Novak
FSU Jena, Germany, `erich.novak@uni-jena.de`

Joscha Prochno
KFU Graz, Austria, `joscha.prochno@uni-graz.at`

Mario Ullrich
JKU Linz, Austria, `mario.ullrich@jku.at`

The solution $S(f)$ of a numerical problem often depends on a function f of many variables. For example, $S(f)$ could be the integral or the maximum of f. It could also be the function itself. We want to compute $S(f)$, but we only have incomplete information about f: a certain a priori knowledge and the outcome of finitely many measurements.

The measurements might be point evaluations or other linear functionals. Usually we assume that we can choose the measurements at will. We try to choose the measurements in a way that allows us to minimize the error. In this case, we talk about optimal information.

In this talk we assume that we do not get to choose the measurements. Instead, we imagine that the information comes random and ask: How much do we loose in comparison to optimal information? We study this question for several examples, where the answers range from *almost nothing* to *almost everything*.

[1] A. Hinrichs, D. Krieg, E. Novak, J. Prochno, and M. Ullrich. *On the power of random information.* arXiv:1903.00681, 2019.

[2] A. Hinrichs, D. Krieg, E. Novak, J. Prochno, and M. Ullrich. *Random sections of ellipsoids and the power of random information.* arXiv:1901.06639, 2019.

Wednesday 10:30–12:30, **Room: CB08.03.002**

Stochastic Computation and Complexity IV

Chair: Larisa Yariskavtseva

CLTs for stochastic approximation schemes under nonstandard assumptions

Steffen Dereich
Universität Münster, Germany, steffen.dereich@wwu.de

Sebastian Kassing
Universität Münster, Germany, sebastian.kassing@wwu.de

Stochastic approximation algorithms are a standard tool for the numerical computation of minimizers. Starting from the works of Robbins and Monro [3] the field attracted significant attention and the rich scientific output is part of several monographs. In the case where the dynamical scheme is attracted by isolated points stochastic approximation schemes are well understood in the classical case treated in the before-mentioned article as well as for the Polyak-Ruppert average treated for instance in [2]. In machine learning applications though minimizers typically are not isolated points so that the classical theory is not applicable. Under assumptions similar to the ones imposed on the objective function in [1] we provide new central limit theorems for the perturbation around the attracting manifold. Moreover, we derive a weak limit theorem for the difference between the values of the objective function at the estimated and optimal parameter.

[1] B. Fehrman, B. Gess, and A. Jentzen. *Convergence rates for the stochastic gradient descent method for non-convex objective functions.* arXiv:1904.01517 [math.NA], 2019.

[2] B. T. Polyak and A. B. Juditsky. *Acceleration of stochastic approximation by averaging.* SIAM J. Control Optim., 30(4):838–855, 1992.

[3] H. Robbins and S. Monro. *A stochastic approximation method.* Ann. Math. Statistics, 22:400–407, 1951.

Central limit theorem for discretization errors based on stopping time sampling

Emmanuel Gobet
Ecole Polytechnique, France, emmanuel.gobet@polytechnique.edu

Nicolas Landon
Ecole Polytechnique, France, nico.landon@gmail.com

Uladzislau Stazhynski
Ecole Polytechnique, France, uladzislau.stazhynski@polytechnique.edu

We study the convergence in distribution of the renormalized error arising from the discretization of a Brownian semimartingale sampled at stopping times. Our mild assumptions on the form of stopping times allow the time grid to be a combination of hitting times of stochastic domains and of Poisson-like random times. Remarkably, a Functional Central Limit Theorem holds under great generality on the semimartingale and on the form of stopping times. Furthermore, the asymptotic characteristics are quite explicit. Along the derivation of such results, we also establish some key estimates related to approximations and sensitivities of hitting time/position with respect to model and domain perturbations.

SGLD with dependent data streams: Convex and Non-Convex case

Mathias Barkhagen
The University of Edinburgh, UK,

Ngoc Huy Chau
Alfréd Rényi Institute of Mathematics, Hungary,

Éric Moulines
CMAP, Ecole Polytechnique, France,

Miklos Rásonyi
Alfréd Rényi Institute of Mathematics, Hungary,

Sotirios Sabanis
The University of Edinburgh & The Alan Turing Institute, UK., S.Sabanis@ed.ac.uk,

Ying Zhang
The University of Edinburgh, UK,

Stochastic Gradient Langevin Dynamics (SGLD) is a combination of a Robbins-Monro type algorithm with Langevin dynamics. It has been extremely popular lately due to its links with ML algorithms, see [1]. In this talk, the SGLD method with fixed step size λ is considered in order to sample from a (not necessarily logconcave) target distribution π, known up to a normalisation factor. We assume, as in [2], that unbiased estimates of the gradient from possibly dependent observations are available and estimate the distance of the SGLD law from π in a suitable Wasserstein-type metric.

[1] M. Raginsky, A. Rakhlin, and M. Telgarsky. Non-convex learning via Stochastic Gradient Langevin Dynamics: a nonasymptotic analysis. *Proceedings of Machine Learning Research*, (65)1674–1703, 2017.

[2] M. Barkhagen, N. H. Chau, É. Moulines, M. Rásonyi, S. Sabanis and Y. Zhang. On stochastic gradient Langevin dynamics with stationary data streams in the logconcave case. *Preprint*, 2018. arXiv:1812.02709

Deep Artificial Neural Networks Overcome the Curse of Dimensionality in PDE Approximation

Arnulf Jentzen
ETH Zürich, Switzerland, arnulf.jentzen@sam.math.ethz.ch

Diyora Salimova
ETH Zürich, Switzerland, diyora.salimova@sam.math.ethz.ch

Timo Welti
ETH Zürich, Switzerland, timo.welti@sam.math.ethz.ch

In recent years deep artificial neural networks (DNNs) have very successfully been employed in numerical simulations for a multitude of computational problems. Such simulations indicate that DNNs seem to be able to overcome the curse of dimensionality in the sense that the number of real parameters used to describe the DNN grows at most polynomially in both the reciprocal of the prescribed approximation accuracy and the dimension of the function which the DNN aims to approximate in these computational problems. However, while there is a large number of rigorous mathematical approximation results for

artificial neural networks in the scientific literature, there are only a few special situations where results in the literature can rigorously explain the success of DNNs when approximating high-dimensional functions.

In this talk it is revealed that DNNs do indeed overcome the curse of dimensionality in the numerical approximation of Kolmogorov PDEs with constant diffusion and nonlinear drift coefficients. A crucial ingredient in our proof of this result is the fact that the artificial neural network used to approximate the solution of the PDE really is a *deep* artificial neural network with a large number of hidden layers. This talk is based on a joint work with Arnulf Jentzen and Diyora Salimova, see [https://arxiv.org/abs/1809.07321].

Wednesday 10:30–12:30, **Room: CB08.03.005**

Recent Advances in Monte Carlo Methods for Robustness, Long Memory Processes and Rough Differential Equations

Chair: José Blanchet

Perfect Sampling for Queues with Autoregressive Arrivals

Xinyun Chen
The Chinese University of Hong Kong (Shenzhen), China, chenxinyun@cuhk.edu.cn

Motivated by the applications in social media, healthcare and financial systems, queueing models with autoregressive arrivals, such as the Hawkes or shot-noise-driven arrivals, have been studied in several recent works. So far, there is no analytic approach to compute the stationary distribution of single- or multiple - server queues. In this talk, we propose a perfect sampling algorithm that generates i.i.d. samples exactly following the stationary distribution of single- and multiple-server queues with autoregressive arrivals.

Enhanced Balancing of Bias-Variance Tradeoff in Stochastic Simulation

Henry Lam
Columbia University, USA, henry.lam@columbia.edu

Xinyu Zhang
Columbia University, USA, zhang.xinyu@columbia.edu

Xuhui Zhang
University of Science and Technology of China, zxh1998@mail.ustc.edu.cn

Biased stochastic estimators, such as finite-differences for noisy gradient estimation, often contain tuning parameters that balance impacts from the bias and the variance. While the optimal order of these parameters in terms of the simulation budget can be readily established, the precise best values depend on model characteristics that are typically unknown in advance. We investigate a framework to construct new estimators by combining simulation runs on sequences of tuning parameter values, such that the estimators consistently outperform a given tuning parameter choice in the conventional approach, regardless of the unknown model characteristics. We argue this outperformance via a minimax risk ratio and demonstrate, under this framework, a consistently outperforming estimator constructed by a weighting scheme that is characterized by a sum of two components with distinct decay rates.

ϵ-Strong Simulation of Fractional Brownian Motion and Related Stochastic Differential Equations

Yi Chen
Northwestern University, USA, yichen2016@u.northwestern.edu

Jing Dong
Columbia University, USA, jing.dong@gsb.columbia.edu

Hao Ni
University College London, UK, h.ni@ucl.ac.uk

Consider the fractional Brownian Motion (fBM) $B^H = \{B^H(t) : t \in [0,1]\}$ with Hurst index $H \in (0,1)$. We construct a probability space supporting both B^H and a fully simulatable process $\hat{B}^H_\epsilon$ such that

$$\sup_{t\in[0,1]} |B^H(t) - \hat{B}^H_\epsilon(t)| \leq \epsilon$$

with probability one for any user specified error parameter $\epsilon > 0$. When $H > 1/2$, we further enhance our error guarantee to the α-Hölder norm for any $\alpha \in (1/2, H)$. This enables us to extend our algorithm to the simulation of fBM driven stochastic differential equations $Y = \{Y(t) : t \in [0,1]\}$. Under mild regularity conditions on the drift and diffusion coefficients of Y, we construct a probability space supporting both Y and a fully simulatable process $\hat{Y}_\epsilon$ such that

$$\sup_{t\in[0,1]} |Y(t) - \hat{Y}^H_\epsilon(t)| \leq \epsilon$$

with probability one. Our algorithms enjoy the tolerance-enforcement feature, i.e., the error bounds can be updated sequentially. Thus, the algorithms can be readily combined with other simulation techniques like multilevel Monte Carlo to estimate expectation of functionals of fBMs efficiently.

Convergence of one-step survival Brownian bridge Monte Carlo for barrier options

Thomas Gerstner
Goethe Universitaet, Germany, gerstner@math.uni-frankfurt.de

Bastian Harrach
Goethe Universitaet, Germany, harrach@math.uni-frankfurt.de

Daniel Roth
Goethe Universitaet, Germany, roth@math.uni-frankfurt.de

We consider the pricing and the sensitivity calculation of continuously monitored barrier options. Standard Monte Carlo algorithms work well for pricing these options. However, since they do not behave stable with respect to numerical differentiation, one would generally resort to regularized differentiation schemes or derive an algorithm for the precise differentiation. For barrier options the Brownian bridge approach leads to a precise, but non Lipschitz continuous, first differentiation.

In this work we will show weak convergence of order one for the Brownian bridge approach. Then, we generalize the idea of one-step survival, first introduced by Glasserman and Staum, to general scalar stochastic differential equations and combine it with the Brownian bridge approach leading to a new one-step survival Brownian bridge approximation. We show that the new technique can be adapted in such a way that its results satisfies stable second order Greeks. Besides studying the stability we will prove unbiasedness, leading to an equal convergence property, and variance reduction.

Furthermore we present the partial differentiations which allow to adapt a pathwise sensitivity algorithm. Moreover, we develop a one-step survival Brownian bridge Multilevel Monte Carlo algorithm to greatly reduce the computational cost in practice.

Monte Carlo Estimation: Miscellaneous Applications

Chair: Spiridon Penev

Choosing the Tuning Parameter for Goodness-of-Fit Tests that Employ Bootstrapped Critical Values by Using Estimated Sizes

Willem D. Schutte
Centre for Business Mathematics and Informatics, North-West University, South Africa,
wd.schutte@nwu.ac.za

Leonard Santana
Unit for Business Mathematics and Informatics, North-West University, South Africa,
leonard.santana@nwu.ac.za

James S. Allison
Unit for Business Mathematics and Informatics, North-West University, South Africa,
james.allison@nwu.ac.za

Many goodness-of-fit tests in the literature contain some tuning parameters that should be appropriately selected in order to ensure reasonable power for these tests. In [1], a data-dependent choice of the tuning parameter that appears in these goodness-of-fit tests is discussed. However, these methods are only applicable to a class of tests for which the null distribution can be approximated through simple Monte Carlo methods, e.g., for a location-scale family. This data-dependent choice of the tuning parameter is obtained by maximising the bootstrap power of a test when calculated from the Monte Carlo approximation of the appropriate critical value. Even though the bootstrap power is maximised, the procedure is able to maintain the nominal significance level of the test as the size is reasonably accurately approximated by Monte Carlo methods. However, in the slightly broader class of tests where this is not possible, i.e., the null distribution is unknown and cannot be approximated via Monte Carlo methods, one cannot rely on this procedure since the process of choosing the parameter based on maximising bootstrap power will lead to tests that are oversized. Unlike the approach followed in [1] the approach followed for the tests in this class is to approximate the null distribution (and resulting critical values) using the bootstrap. Typical tests that fall in this class include, for example, testing for symmetry and testing goodness-of-fit for the gamma distribution with a shape parameter. The new method to obtain the data-dependent choice of the tuning parameter in these tests does not rely on maximising bootstrap power, but rather relies on attempting to find the value of the tuning parameter that allows the test to come as close as possible to the specified nominal significance level. However, a challenge that one faces in the setting where the estimated nominal sizes of the test differ for the various choices of the tuning parameter is that one cannot easily compare powers across tests employing these different choices. Therefore, size adjustments to reported powers are incorporated to facilitate correct power comparisons by making use of the procedures described in [3]. An iterative bootstrap algorithm which employs the warp speed method of [2] is provided and the relative performance of our new method for choosing the parameter is investigated using different tuning parameter-based test statistics in the two example scenarios mentioned above. The proposed method of tuning parameter selection works well for the gamma distribution, selecting the method that both attains the nominal significance level and produces high powers, but the method applied to the test for symmetry in regression models does not fare as well and seldom selects the parameters that produce high power.

[1] J. S. Allison, and L. Santana. *On a data-dependent choice of the tuning parameter appearing in certain goodness-of-fit tests.* Journal of Statistical Computation and Simulation, Volume 85(16):3276–3288, 2015.

[2] R. Giacomini, D. N. Politis, and H. White. *A warp-speed method for conducting Monte Carlo experiments involving bootstrap estimators.* Econometric Theory, Volume 29: 567–589, 2013.

[3] C. J. Lloyd. *Estimating test power adjusted for size.* Journal of Statistical Computation and Simulation, Volume 75(11): 921–933, 2005.

Assessing Weighted Inter-rater Agreement Under Missing Data

Anil Dolgun
RMIT University Melbourne, Australia, anil.dolgun@rmit.edu.au

Haydar Demirhan
RMIT University Melbourne, Australia, haydar.demirhan@rmit.edu.au

In inter-rater agreement studies, missing data commonly occur when at least one rating of a unit is missing. Such missingness can be a result of a dropout during a clinical trial, a missed appointment in an observational study or because at least one of the raters fails to rate the unit. Missing data can have a substantial effect on conclusions drawn from the data [1, 3], therefore the assessment of inter-rater agreement should be carefully done for the cases where missing ratings exist. In this study, we propose a new weighting strategy for assessing inter-rater agreement assuming the missing ratings are missing completely at random (MCAR). We give possible implications of this weighting strategy for different inter-rater agreement coefficients and under different missingness mechanisms. We compare the performance of the proposed method with Gwet's weighted kappa [2] and listwise deletion kappa using a Monte Carlo simulation study, and conclude by summarizing the usefulness of this approach in the analysis of this particular type of agreement data.

References

[1] Craig K Enders. The impact of missing data on sample reliability estimates: Implications for reliability reporting practices. *Educational and psychological measurement*, 64(3):419–436, 2004.

[2] Kilem L Gwet. *Handbook of inter-rater reliability: The definitive guide to measuring the extent of agreement among raters.* Advanced Analytics, LLC, 2014.

[3] Jan-Willem Strijbos and Gerry Stahl. Methodological issues in developing a multi-dimensional coding procedure for small-group chat communication. *Learning and Instruction*, 17(4):394–404, 2007.

Robust index tracking

Spiridon Penev
UNSW Sydney, Australia, s.penev@unsw.edu.au

Pavel Shevchenko
Macquarie University, Sydney, Australia, pavel.shevchenko@mq.edu.au

Wei Wu
UNSW Sydney, Australia, wei.wu1@unsw.edu.au

Index tracking is a popular form of asset management. Typically, a quadratic function is used to define the tracking error of a portfolio and the look back approach is applied to solve the index tracking problem. We argue that a forward looking approach is more suitable, whereby the tracking error is expressed as expectation of a function of the difference between the returns of the index and of the

portfolio. We also assume that there is an uncertainty in the distribution of the assets, hence a robust version of the optimization problem needs to be adopted. We use Bregman divergence in describing the deviation between the nominal and actual distribution of the components of the index. In this scenario, we derive the optimal robust index tracking strategy in a semi-analytical form as a solution of a system of nonlinear equations. Calculation of some expected values involved in the semi-analytical solutions is implemented via Monte Carlo simulation. We compare the performance of the robust strategy with the optimal non-robust strategy.

[1] S. Penev, and K. Naito. Locally robust methods and near-parametric asymptotics. *Journal of Multivariate Analysis*. 167, 395–417, 2018.

[2] P. Glasserman, and X. Xu. Robust risk measurement and model risk. *Quantitative Finance*. 14, 29–58, 2014

Projection Monte Carlo Methods For Estimation of the Bidirectional Scattering-Surface Reflectance Distribution Function

Natalya Tracheva
Novosibirsk State University, Novosibirsk, Russia and Institute of Computational Mathematics and Mathematical Geophysics of SB RAS, Novosibirsk, Russia, tnv@osmf.sscc.ru

Sergey Ukhinov
Novosibirsk State University, Novosibirsk, Russia and Institute of Computational Mathematics and Mathematical Geophysics of SB RAS, Novosibirsk, Russia, sau@osmf.sscc.ru

Accurate modeling of the scattered by medium radiation is essential for a wide range of applications in atmospheric optics and image rendering. The usual approach is focused on considering models for the bidirectional reflectance distribution function. In this talk, we discuss the problem of approximation of a bidirectional scattering-surface reflectance distribution function with the accounting of subsurface particle interactions and polarization.

We consider the vector radiation transport equation that enables the introduction of the polarization effect in the model of radiation transfer. In the process of subsurface interaction during the propagation of the radiation, there occur multiple changes in the polarization status of the light, which depend on the characteristics of the scattering and absorbing medium. A change in the polarization state can significantly influence on the angular probability densities of back-scattered or transmitted by the medium radiation.

To estimate the bidirectional scattering-surface reflectance distribution function, we consider an approach based on the projection expansion. We discuss particular cases of this approach. The first one is a classical two-dimensional histogram-based estimation and its generalization — a kernel density estimator. The second approach is an expansion of the angular bidirectional scattering-surface reflectance distribution function on the orthonormal polynomial basis with Lambertian weight. The possibility of applying the Monte Carlo method is determined by the fact that expansion coefficients are the mathematical expectations of weighted random values of the standard functions.

We discuss a specific for the posed problem technique of choosing optimal kernel bandwidth for the uniform kernel function and provide a comparative review of numerical results, obtained for the angular bidirectional scattering-surface reflectance distribution function of the intensity and the degree of polarization of radiation.

The reported study was partially funded by RFBR according to the research projects 17-01-00823, 18-01-00356, 18-31-00213.

Wednesday 10:30–12:30, **Room: CB08.02.002**

Scalable Bayesian Inference for Complex Statistical Models

Chair: Liangliang Wang

Effective Online Bayesian Phylogenetics Via Sequential Monte Carlo with Guided Proposals

Mathieu Fourment
University of Technology Sydney, mathieu.fourment@uts.edu.au

Brian Claywell
Fred Hutchinson Cancer Research Center, bclaywel@fredhutch.org

Vu Dinh
University of Delaware, vucdinh@udel.edu

Connor McCoy
Google, connorm@gmail.com

Erick Matsen
Fred Hutchinson Cancer Research Center, matsen@fredhutch.org

Aaron Darling
University of Technology Sydney, aaron.darling@uts.edu.au

Modern infectious disease outbreak surveillance produces continuous streams of sequence data which require phylogenetic analysis as data arrives. Current software packages for Bayesian phylogenetic inference are unable to quickly incorporate new sequences as they become available, making them less useful for dynamically unfolding evolutionary stories. This limitation can be addressed by applying a class of Bayesian statistical inference algorithms called sequential Monte Carlo to conduct *online inference*, wherein new data can be continuously incorporated to update the estimate of the posterior probability distribution. In this paper we describe and evaluate several different online phylogenetic sequential Monte Carlo (OPSMC) algorithms. We show that proposing new phylogenies with a density similar to the Bayesian prior suffers from poor performance, and we develop 'guided' proposals that better match the proposal density to the posterior. Furthermore, we show that the simplest guided proposals can exhibit pathological behavior in some situations, leading to poor results, and that the situation can be resolved by heating the proposal density. The results demonstrate that relative to the widely-used MCMC-based algorithm implemented in MrBayes, the total time required to compute a series of phylogenetic posteriors as sequences arrive can be significantly reduced by the use of OPSMC, without incurring a significant loss in accuracy.

Parallel tempering: Scaling limits, and optimality

Saifuddin Syed
University of British Columbia, Canada, saif.syed@stat.ubc.ca

Alexandre Bouchard-Côté
University of British Columbia, Canada, bouchard@stat.ubc.ca

George Deligiannidis
University of Oxford, UK, deligian@stats.ox.ac.uk

Arnaud Doucet
University of Oxford, UK, doucet@stats.ox.ac.uk

MCMC methods are widely used to approximate intractable expectations with respect to high-dimensional un-normalized probability distributions. We construct a Markov chain with our desired stationary distribution to explore our sample space. In theory, the chain should accurately explore the state space asymptotically, but in practice, it can get trapped exploring local regions of high probability and suffer from poor mixing in a finite time.

Parallel tempering (PT) algorithms were introduced to try and tackle this issue. We delegate the task of exploration to additional heated chains running in parallel with better mixing properties. They then communicate with the target chain of interest and help it discover new unexplored regions of the sample space. Since their introduction in the 90's, PT algorithms are still extensively used to improve mixing in hard sampling problems arising in statistics, physics, computational chemistry, phylogenetics, and machine learning. In this talk, we will give an introduction to PT algorithms, determine their scaling behaviour, efficiency, and limitations. Consequentially, we will establish the theoretically optimal temperature spacing, communication scheme, and a number of chains for a general class of sampling problems, as well as guidelines for efficient implementation.

Variance Reduction Properties of the Reparameterization Trick

Ming Xu
Queensland University of Technology, mingda.xu@hdr.qut.edu.au

Matias Quiroz
University of Technology Sydney,

Robert Kohn
University of New South Wales,

Scott A. Sisson
University of New South Wales,

The reparameterization trick is widely used in variational inference as it yields more accurate estimates of the gradient of the variational objective than alternative approaches such as the score function method. Although there is overwhelming empirical evidence in the literature showing its success, there is relatively little research exploring why the reparameterization trick is so effective. We explore this under the idealized assumptions that the variational approximation is a mean-field Gaussian density and that the log of the joint density of the model parameters and the data is a quadratic function that depends on the variational mean. From this, we show that the marginal variances of the reparameterization gradient estimator are smaller than those of the score function gradient estimator. We apply the result of our idealized analysis to real-world examples.

An Annealed Sequential Monte Carlo Method for Bayesian Phylogenetics

Liangliang Wang
Simon Fraser University, Canada, lwa68@sfu.ca

Shijia Wang
Simon Fraser University, Canada, shijiaw@sfu.ca

Alexandre Bouchard-Côté
University of British Columbia, Canada, bouchard@stat.ubc.ca

In [1], we described an "embarrassingly parallel" method for Bayesian phylogenetic inference, annealed Sequential Monte Carlo, based on recent advances in the Sequential Monte Carlo literature such as adaptive determination of annealing parameters. The algorithm provides an approximate posterior distribution over trees and evolutionary parameters as well as an unbiased estimator for the marginal likelihood. This unbiasedness property can be used for the purpose of testing the correctness of posterior simulation software. We evaluate the performance of phylogenetic annealed Sequential Monte Carlo by reviewing and comparing with other computational Bayesian phylogenetic methods, in particular, different marginal likelihood estimation methods. Unlike previous Sequential Monte Carlo methods in phylogenetics, our annealed method can utilize standard Markov chain Monte Carlo tree moves and hence benefit from the large inventory of such moves available in the literature. Consequently, the annealed Sequential Monte Carlo method should be relatively easy to incorporate into existing phylogenetic software packages based on Markov chain Monte Carlo algorithms. We illustrate our method using simulation studies and real data analysis.

[1] L. Wang, S. Wang, and A. Bouchard-Côté. An Annealed Sequential Monte Carlo Method for Bayesian Phylogenetics. *Systematic Biology*, (In Press), 2019.

Wednesday 13:30–15:00, **Room: CB08.02.002**

Probabilistic Numerical Methods

Chair: Toni Karvonen

Convergence rates of adaptive Bayesian quadrature

Motonobu Kanagawa
University of Tübingen, motonobu.kanagawa@uni-tuebingen.de

Bayesian quadrature is a probabilistic numerical method for the task of numerical integration. It uses a stochastic process (typically a Gaussian process) as a prior distribution of the integrand, and derives a posterior distribution of the value of the integral, treating the function values at design points as "observed data." Thus, the key question is how to set the set of design points. Recently there are adaptive approaches that select design points in a sequential way depending on the previously observed function values. These methods have been shown to be empirically superior to non-adaptive methods for some cases where the integrand is known to possess certain properties, such as it being non-negative, but no convergence guarantee has been established. In this talk, I'll explain our recent results on convergence rates of adaptive Bayesian quadrature.

Numerical Integration Method for Training Neural Networks

Sho Sonoda
RIKEN AIP, Tokyo, Japan, sho.sonoda@riken.jp

We propose a novel numerical integration method for training a shallow neural network by using the ridgelet transform with a uniform convergence guarantee. Given a training dataset, the ridgelet transform can provide the parameters of the neural network that attains the global optimum of the training problem. In other words, we can obtain the global minimizer of the training problem by numerically computing the ridgelet transform, instead of by numerically optimizing the so-called backpropagation training problem. We employ the kernel quadrature for the basis of the numerical integration. Originally, the kernel quadrature has been developed for the purpose of computing posterior means, where the measure is assumed to be a probability measure, and the final product is a single number. On the other hand, our problem is the computation of an integral transform, where the measure is generally a signed measure, and the final product is a function. In this paper, we develop a generalized kernel quadrature method for signed measures with a uniform convergence guarantee, and propose a natural choice of kernels.

Methods for Large-Scale and High-Dimensional Probabilistic Integration

Toni Karvonen
Aalto University, Finland, toni.karvonen@aalto.fi

For N data points, the naive implementation of Bayesian cubature and other kernel-based numerical integration methods is based on solving a linear system of N equations. The resulting cubic computational and quadratic memory cost in N restrict the applicability of these methods. We show how relatively flexible *fully symmetric sets*, obtained from given vectors via coordinate permutations and sign-changes, can be exploited for efficient computation of the weights of kernel-based cubature rules for up to tens of millions of points. If the point set is a union of J fully symmetric sets, computational complexity is reduced from $\mathcal{O}(N^3)$ to $\mathcal{O}(J^3 + JN)$ and memory complexity from $\mathcal{O}(N^2)$ to $\mathcal{O}(J^2)$. In its current state the algorithm does not adapt to efficient optimization of kernel hyperparameters, which is a crucial step

in uncertainty calibration of a Bayesian cubature method. The talk is mainly based on the articles [1, 2], but we also briefly discuss some other recent approaches based on sparse grids [3] and a combination of low discrepancy points and shift-invariant kernels [4].

[1] T. Karvonen and S. Särkkä. Fully symmetric kernel quadrature. *SIAM Journal on Scientific Computing*, 40(2):A697–A720, 2018.

[2] T. Karvonen, S. Särkkä, and C. J. Oates. Symmetry exploits for Bayesian cubature methods. To appear in *Statistics and Computing*, 2019.

[3] J. Oettershagen. *Construction of Optimal Cubature Algorithms with Applications to Econometrics and Uncertainty Quantification*. PhD thesis, Institut für Numerische Simulation, Universität Bonn, 2017.

[4] J. Rathinavel and F. Hickernell. Fast automatic Bayesian cubature using lattice sampling. Preprint, *arXiv:1809.09803v1*, 2018.

Wednesday 13:30–15:00, **Room: CB08.03.002**

Approximate Bayesian Computation

Chair: Grégoire Claré

An attempt to make ABC cheaper

Julien Stoehr
U Paris-Dauphine, stoehr@ceremade.dauphine.fr

Approximate Bayesian computation (ABC) are now a well established simulation method to deal with parametric statistical models whose likelihood is intractable. A burden of such a method stems from the computational cost when simulating data sets from a model is too time-consuming so that the posterior distribution cannot be approximated or sampled from in a reasonable amount of time. The Bayesian inference is then not straightforward. In this talk, we present an attempt to reduce the cost of the algorithm by plugging a non-parametric estimator based on an ABC reference table within a MCMC scheme.

Approximate Bayesian computation to model passenger flow within airport terminals

Anthony Ebert
QUT, Australia, ac.ebert@qut.edu.au

Rapidly increasing numbers of air passengers have led to congestion within airport terminals. To help manage this congestion, we develop statistical tools for decision support. Existing decision support tools for passenger flows in airports do not allow parameters to be inferred systematically and do not account for parameter uncertainty. This is because the complexity of processes occurring within the terminal make the likelihood intractable, meaning that approximate Bayesian computation (ABC), which is based on a distance between data and model realisations, should be appropriate. However, standard queueing simulation techniques are computationally expensive which make ABC infeasible for such problems. A new queueing simulation algorithm called queue departure computation has made ABC feasible for such models. We repurpose maximum mean discrepancy, a distance on probability measures as a distance on functional data between observed passenger flow counts and model realisations. We demonstrate how it is used to infer parameters and provide tools for decision support with a case study of a delayed flight. Finally, we adapt curve registration techniques to the likelihood-free domain to further enhance our ABC distance.

ABC within Gibbs sampling

Grégoire Claré
U Paris-Dauphine, clarte@ceremade.dauphine.fr

ABC methods become exponentially less efficient as the parameter space grows in dimension. Our idea to reduce this difficulty is to explore the parameter space according to the Gibbs sampler, by plugging ABC approximations of the conditionals into the scheme. Several numerical simulations show the efficiency of this approach, although it shares the limitations of both Gibbs sampler and ABC methods: compatibility of the conditionals and choice of summary statistics. We use coupling techniques to show the convergence of the algorithm under some conditions and compute a rough upper bound on the total variation distance between the resulting law and the law of the exact Gibbs sampler.

Wednesday 13:30–15:00, **Room: CB08.03.004**

Advances in Variational Inference

Chair: Ming Xu

Streamlined Variational Inference for Inverse Problems Models

Luca Maestrini
University of Technology Sydney, Australia, luca.maestrini@uts.edu.au

Matt P. Wand
University of Technology Sydney, Australia, matt.wand@uts.edu.au

Robert G. Aykroyd
University of Leeds, United Kingdom, r.g.aykroyd@leeds.ac.uk

Inverse problems are essentially statistical regression problems where a response depending on a number of causal factors or parameters is measured and the goal is to estimate the parameter values. However, they may be highly multivariate and have predictors which are highly correlated. Hence even linear inverse problems cannot be solved by classical regression methods, nor can they be solved using standard dimension reduction or regularised regression techniques. A remedy is to use Markov random field models, which can be slow to fit via Markov chain Monte Carlo methods. Variational message passing updating algorithms for factor graph fragments arising in inverse problem Bayesian models are identified, catalogued and derived. The resultant factor graph fragments facilitate streamlined implementation of fast approximate algorithms for inverse problems and form the basis for software development. Contemporary inverse problems models give rise to new factor graph fragment types for different penalization strategies. Nevertheless, the variational message passing approach on factor graph fragments is such that algorithm updates and streamlining steps only need to be derived once for a particular fragment, which can be integrated in an arbitrarily complex model. The first applications are one- and two-dimensional deconvolution problems motivated by archaeology data.

Stochastic Variational Bayes with Particle Filter

Nghia Nguyen
University of Sydney, Australia, nghia.nguyen@sydney.edu.au

Minh-Ngoc Tran
University of Sydney, Australia, minh-ngoc.tran@sydney.edu.au

We propose an efficient Bayesian inference framework for a class of non-linear and non-Gaussian state space models using Gaussian variational approximation, with a parsimonious but fexible factor parametrization of the covariance matrix, and particle filter. We implement natural gradient methods for the optimization, exploiting the factor structure of the variational covariance matrix in computation of the natural gradient. The intractable likelihood is unbiasedly estimated from a particle filter. We illustrate the method using a high dimensional factor stochastic volatility having both a large number of parameters and a large number of latent states and show that our proposed method makes the computation much more efficient than Particle Markov Chain Monte Carlo framework.

Variational forecasts of observation-driven models

Alex Cooper
Monash University, alexander.cooper@monash.edu

Bonsoo Koo
Monash University, bonsoo.koo@monash.edu

David Frazierr
Monash University, david.frazier@monash.edu

Gael Martin
Monash University, gael.martin@monash.edu

We demonstrate that approximate Bayesian forecasts constructed using variational methods are both accurate and extremely fast to compute. We focus on the class of generalized linear autoregressions (GLAR), a simple but large class of observation-driven forecasting models with wide practical application. We present both theoretical and numerical evidence that variational forecasts can have the same predictive power as 'exact' (ie unapproximated) Bayesian forecasts constructed using sampling methods.

Our theoretical results provide conditions for variational GLAR forecasts to be asymptotically equivalent to exact forecasts, in two respects. First, we show that variational approximate probabilistic forecast distributions 'agree' with their exact counterparts in the sense of statistical merging. Second, we show that in the large-data limit, variational forecasting performance is no worse than for exact forecasts, as measured by a proper scoring rule.

We also present numerical experiments on variational forecasts for different data types, including continuous, binary, and count data. In each case, the experiments confirm our theoretical results and demonstrate that variational forecasts can be several orders of magnitude faster than forecasts based on sampling methods.

Wednesday 13:30–15:00, **Room: CB08.03.005**

Applied Probability and Monte Carlo methods II

Chair: Chang-Han Rhee

Limiting distributional fixed points in large banking networks

Anand Deo
Tata Institute of Fundamental Research, India, anand.deo@tifr.res.in

We analyse the equilibrium behaviour of a large network of banks in presence of incomplete information, where inter-bank borrowing and lending is allowed, and banks suffer shocks to assets. In a two time period graphical model, we show that the equilibrium wealth distribution is the fixed point of a complex, high dimensional distribution-valued map.

Fortunately, there is a dimension collapse in the limit as the network size increases, where the equilibriated system converges to the unique fixed point involving a simple, one dimensional distribution-valued operator, which, we show, is amenable to simulation. Specifically, we develop a Monte-Carlo algorithm that computes the fixed point of a general distribution-valued map and derive sample complexity guarantees for it. We numerically show that this limiting one-dimensional regime can be used to obtain useful structural insights and approximations for networks with as low as a few hundred banks.

Distributed Stochastic Optimization

Raghu Pasupathy
Purdue University, Department of Statistics, USA, pasupath@purdue.edu

The distributed stochastic optimization problem is that of co-operatively minimizing the sum of n real-valued functions $f_i, i = 1, 2, \ldots, n$ by n "agents." Agent i is assumed to have access to "noisy" estimates of the gradient of function f_i alone. Furthermore, the agents are assumed to be located on a connected graph $\mathcal{G}$ so that information can be passed by each agent to its neighbors, but usually at high cost. We review the problem setup for distributed stochastic optimization and its wide applicability, and discuss the key stochastic approximation results to solve this problem. We will specifically compare existing rate results against the "centralized" context.

The Skipping Random Walk Metropolis Sampler

Alessandro Zocca
California Institute of Technology, USA, azocca@caltech.edu

John Moriarty
Queen Mary University of London , UK, j.moriarty@qmul.ac.uk

Jure Vogrinc
Warwick Univeristy, UK, Jure.Vogrinc@warwick.ac.uk

In this talk I will introduce a novel algorithm, named skipping Random Walk Metropolis (skipping RWM), to sample from conditional distributions with densities of the form

$$\pi = \frac{\rho \mathbf{1}_C}{\rho(C)},$$

where ρ is a density on $\mathbf{R}^d$ and C is a set of 'extreme' values of ρ. This sampling algorithm is designed to addresses in particular the situation where the event C is rare or multimodal or has a disconnected support. Such issues can arise, for example, in the study of risk and reliability, or while using the slice sampling algorithm, where the density ρ is uniform and the set C may be disconnected. While the density π must be known explicitly (up to a multiplicative constant), the other assumptions are otherwise minimal, e.g., no knowledge is required about the derivatives of π and indeed it need not be smooth. I will prove a strong law of large numbers for skipping RWM, show comparison results with the classical Metropolis-Hastings algorithm and present an application to power grid reliability, where the goal is to understand how random disturbances and fluctuations (due for instance to renewable energy sources) can cause frequency violations.

This talk is based on joint work with John Moriarty (QMUL) and Jure Vogrinc (University of Warwick).

Thursday 10:30–12:30, **Room: CB08.02.002**

Multilevel Monte Carlo methods

Chair: Tomohiko Hironaka

Unbiased Estimation of the Reciprocal Mean for non-negative Random Variables

Sarat Moka
The University of Queensland, Brisbane, Australia, s.babumoka@uq.edu.au

Dirk Kroese
The University of Queensland, Brisbane, Australia, kroese@maths.uq.edu.au

Sandeep Juneja
TIFR, Mumabi, India, juneja@tifr.res.in

Many simulation problems require the estimation of a ratio of two expectations. In recent years Monte Carlo estimators have been proposed that can estimate such ratios without bias. We investigate the theoretical properties of such estimators for the estimation of $\beta = 1/\mathbb{E}[Z]$, where $Z \geq 0$. The estimator, $\widehat{\beta}(w)$, is of the form $w/f_w(N)\prod_{i=1}^{N}(1-wZ_i)$, where $w < \beta$ and N is any random variable with probability mass function f_w on the positive integers. For a fixed w, the optimal choice for f_w is well understood, but less so the choice of w. We study the properties of $\widehat{\beta}(w)$ as a function of w and show that its expected time variance product decreases as w decreases, even though the cost of constructing the estimator increases with w. We also show that the estimator is asymptotically equivalent to the maximum likelihood (biased) ratio estimator and establish practical confidence intervals.

Multilevel Monte Carlo for simulation and inference of biochemical reaction networks

David Warne
School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, david.warne@qut.edu.au

Ruth Baker
Mathematical Institute, University of Oxford, Oxford, UK, ruth.baker@maths.ox.ac.uk

Matthew Simpson
School of Mathematical Sciences, Queensland University of Technology, Brisbane, Australia, matthew.simpson@qut.edu.au

Stochastic models of biochemical reaction networks are often more realistic descriptions of cellular processes over deterministic counterparts when small populations of certain chemical species are considered. Realistic applications of these models are computationally intensive since analytical solutions to the chemical master equation are rarely available. Multilevel Monte Carlo (MLMC) methods have been demonstrated to be very effective for the acceleration of expectation calculations with respect to the forwards problem. The statistical inference of reaction rate parameters, however, is an even more computationally intensive task that often relies upon likelihood-free methods, such as approximate Bayesian computation (ABC). We investigate a modified ABC approach that is based on and application of MLMC applied directly to rejection sampling. Our method constructs an approximation of the posterior distribution function through a telescoping summation of biased estimators using a decreasing sequence of acceptance thresholds. We briefly review the concept of MLMC for the forwards problem, discuss challenges specific to the inference problem, then demonstrate the effectiveness of our MLMC method for ABC using several stochastic models of biochemical reaction networks.

Multilevel quasi Monte Carlo for optimization of PDEs

Andreas Van Barel
KU Leuven, Belgium, andreas.vanbarel@cs.kuleuven.be

Stefan Vandewalle
KU Leuven, Belgium, stefan.vandewalle@cs.kuleuven.be

The multilevel Monte Carlo method can be used to efficiently generate a gradient (or Hessian vector product) for robust PDE constrained optimization problems [3]. Considering multiple levels reduces the computational time by orders of magnitude. However, the slow convergence rates of the Monte Carlo method apply. The computational effort scales at best as the square of the required accuracy on the gradient. It has been shown in, e.g., [1] that the quasi Monte Carlo method can significantly reduce the number of samples required. Multilevel quasi Monte Carlo methods have also been investigated for elliptic PDEs [2]. In ideal circumstances, the computational effort may scale linearly with the required accuracy. In this talk, we apply the multilevel quasi Monte Carlo method to obtain gradients (and Hessian vector products) to solve robust optimization problems. We compare the performance of this approach for a tracking type robust optimization problem constrained by a model elliptic diffusion PDE with lognormal diffusion coefficient.

[1] I.G. Graham, F.Y. Kuo, D. Nuyens, R. Scheichl, I.H. Sloan, *Quasi-Monte Carlo methods for elliptic PDEs with random coefficients and applications*. Journal of Computational Physics, 230(10), pp. 3668–3694, 2011.

[2] F.Y. Kuo, R. Scheichl, C. Schwab, I.H. Sloan, E. Ullmann, *Multilevel quasi-Monte Carlo methods for lognormal diffusion problems*. Mathematics of Computation, 86(308), pp. 2827–2860, 2017.

[3] A. Van Barel, S. Vandewalle, *Robust Optimization of PDEs with Random Coefficients Using a Multilevel Monte Carlo method*, SIAM/ASA Journal on Uncertainty Quantification, 7(1), pp. 174–202, 2019.

Multilevel Monte Carlo Methods for Estimating the Expected Value of Sample Information

Tomohiko Hironaka
School of Engineering, The University of Tokyo, Japan, hironaka-tomohiko@g.ecc.u-tokyo.ac.jp

Michael B. Giles
Mathematical Institute, University of Oxford, Oxford, United Kingdom, mike.giles@maths.ox.ac.uk

Takashi Goda
School of Engineering, The University of Tokyo, Japan, goda@frcer.t.u-tokyo.ac.jp

The Expected Value of Sample Information (EVSI) is widely used for decision making under uncertainty particularly in the medical field. EVSI is defined as a nested expectation in which an outer expectation is taken with respect to one random variable Y and an inner expectation to the other random variable θ. Although the nested (Markov Chain) Monte Carlo estimator has been often used, a root-mean-square-accuracy of ϵ is achieved notoriously at a cost of $O(\epsilon^{-2-1/\alpha})$, where α is typically between 1/2 and 1.

We study a novel efficient Monte Carlo estimator of EVSI by applying a multilevel Monte Carlo(MLMC) method. Instead of fixing the number of inner samples to θ, we use geometric progression on the number of inner samples. Our MLMC estimator is constructed by telescoping sum, a sum of the Monte Carlo estimators of the differences between successive approximation levels on the inner conditional expectation.

Under a set of assumptions on decision and information models, we show that the differences of successive levels are small, and therefore our MLMC estimator improves the necessary computational cost to optimal $O(\epsilon^{-2})$. Numerical experiments confirm the considerable computational savings as compared to the nested Monte Carlo estimator.

Thursday 10:30–12:30, **Room: CB08.03.002**

Monte Carlo Methods for Large Dependent Data

Chair: Feng Li

Approximate Bayesian Conditional Copula

Clara Grazian
UNSW Sydney, Australia, clara.grazian@unich.it

Luciana Della Valle
University of Plymouth, United Kingdom, luciana.dallavalle@plymouth.ac.uk

Julian Stander
University of Plymouth, United Kingdom, J.Stander@plymouth.ac.uk

Brunero Liseo
Sapienza Università di Roma,Italy, brunero.liseo@uniroma1.it

Many proposals are now available to model complex data, in particular thanks to the recent advances in computational methodologies and algorithms which allow to work with complicated likelihood function in a reasonable amount of time. However, it is, in general, difficult to analyse data characterized by complicated forms of dependence. Copula models have been introduced as probabilistic tools to describe a multivariate random vector via the marginal distributions and a copula function which captures the dependence structure among the vector components, thanks to the Sklar's theorem [1], which states that any d-dimensional absolutely continuous density can be uniquely represented as the product of the marginal distributions and the copula function.

While it is often straightforward to produce reliable estimates for the marginal distributions, making inference on the dependence structure is more difficult. Major areas of application include econometrics, hydrological engineering, biomedical science, signal processing and finance.

Bayesian methods to analyse copula models [2, 3] tend to be computational intensive or to rely on the choice of a particular copula function, even if methods of model selection are not yet fully developed in this setting. We will present a general method to estimate some specific quantities of interest of a generic copula (such as, for example, tail dependence indices, Spearman's ρ or the Kendall's τ) by adopting an approximate Bayesian approach along the lines of [4]. In particular, we discuss the use of an approximate Bayesian computation algorithm, based on the empirical likelihood approximation of the integrated likelihood of the quantity of interest [5].

Our approach is general, in the sense that it could be adapted both to parametric and nonparametric modelling of the marginal distributions and can be generalised in presence of covariates. It also allow to avoid the definition of the copula function in a setting where it is in general difficult to apply model selection procedures.

The class of algorithms proposed allows the researcher to model the joint distribution of a random vector in two separate steps: first the marginal distributions and, then, a copula function which captures the dependence structure among the vector components.

In particular, the extension which allow to consider covariates is useful, since many available approaches are known to be based on not-consistent estimate of the copula function.

[1] Sklar M. Fonctions de répartition à n dimensions et leurs marges. *Publ Inst Statist Univ Paris* **8**, 229–231, (1959)

[2] Craiu V.R., Sabeti A. In mixed company: Bayesian inference for bivariate conditional copula models with discrete and continuous outcomes. *J Multivar Anal* **110**, 106–120, 2012.

[3] Min A., Czado C. Bayesian inference for multivariate copulas using pair-copula constructions. *J Financ Econometrics* **8**(4),511–546, 2010.

[4] Mengersen K., Pudlo P., Robert C.P. Bayesian computation via empirical likelihood. *Proc Natl Acad Sci* **110**(4),1321–1326, 2013.

[5] Grazian, C., Liseo, B. Approximate Bayesian inference in semiparametric copula models. *Proc Natl Acad Sci Bayesian Analysis*, **12**(4), 991–1016. 2017.

Variational Bayes Estimation of Discrete-Margined Copula Models with Application to Time Series

Rubén Loaiza-Maya
Monash University, Australia, ruben.loaizamaya@monash.edu

Michael S Smith
University of Melbourne, Australia, m.smith@mbs.edu

We propose a new variational Bayes estimator for high-dimensional copulas with discrete, or a combination of discrete and continuous, margins. The method is based on a variational approximation to a tractable augmented posterior, and is faster than previous likelihood-based approaches. We use it to estimate drawable vine copulas for univariate and multivariate Markov ordinal and mixed time series. These have dimension rT, where T is the number of observations and r is the number of series, and are difficult to estimate using previous methods. The vine pair-copulas are carefully selected to allow for heteroskedasticity, which is a feature of most ordinal time series data. When combined with flexible margins, the resulting time series models also allow for other common features of ordinal data, such as zero inflation, multiple modes and under- or over-dispersion. Using six example series, we illustrate both the flexibility of the time series copula models, and the efficacy of the variational Bayes estimator for copulas of up to 792 dimensions and 60 parameters. This far exceeds the size and complexity of copula models for discrete data that can be estimated using previous methods.

Predicting forecasting algorithm performance with diverse time series features

Yanfei Kang
Beihang University, China, yanfeikang@buaa.edu.cn

Feature-based time series representation has attracted substantial attention in a wide range of time series analysis methods. Recently, the use of time series features for forecast model selection and model averaging has been an emerging research focus in the forecasting community. That calls for a more diverse time series benchmarks as the training data to model time series features and forecasting algorithm performances. We propose GeneRAting TIme Series with diverse and controllable characteristics, named GRATIS, with the use of mixture autoregressive (MAR) models. We generate sets of time series using MAR models and investigate the diversity and coverage of the generated time series in a time series feature space. Efficient Bayesian surface regression is used on the generated data to examine how time series features influence forecasting method performances, which enables us to predict the performances of the forecasting methods on test data. Our experiments on the largest forecasting competition data (M4) show that the proposed methodology yields comparable performances with the top best methods in M4 competition. (Joint work with Feng Li, Rob J. Hyndman, Thiyanga Talagala and George Athanasopoulos).

Bayesian high-dimensional covariate-dependent copula modeling with application to stocks and text sentiments

Feng Li
Central University of Finance and Economics, China, feng.li@cufe.edu.cn

Tail-dependence modeling with flexible marginal and copula distributions is widely used in financial time series. We extend the covariate dependent copula model of Li and Kang (2018) to vine copulas and we allow all pair copula parameters including those that describe conditional dependence to be functions of linear predictors. Although vine decompositions are not unique, fast heuristic algorithms exist for selecting vine structures. The implementation of these algorithms with covariate dependent copula parameters is not straightforward. Instead for our empirical studies we exploit the structure of the data themselves. We update the copula component together with the marginal components jointly. The joint posterior is not tractable and we use a stochastic gradient MCMC within Gibbs sampler, i.e. a Gibbs sampler is used for updating the joint parameter components, with each conditional parameter block updated by stochastic gradient MCMC algorithm. Our model is applied to financial stock returns with text sentiments data.

Thursday 10:30–12:30, **Room: CB08.03.004**

Diffusion Processes: Theory and Applications

Chair: Alexander Shkolnik

Efficient simulation of a two-factor stochastic volatility jump-diffusion model with non-zero correlation between variance factors

Sumei Zhang
School of Science, Xi'an University of Posts and Telecommunications, China, zhanggsumei@163.com

Nikolai Dokuchaev
Department of Mathematics and Statistics, Curtin University, Australia, N.Dokuchaev@curtin.edu.au

We extend the double Heston model by assuming non-zero correlation between variance factors while allowing double exponential jumps in asset price. Based on the transformed volatility scheme, we provide a simulation scheme for the extended model. We apply the scheme to arithmetic Asian options. Numerical results show that the scheme is efficient and outperforms the Euler scheme. Variance reduction technique based on antithetic variate and control variate further improves the scheme. By investigating the impacts of the introduced correlation and jumps on Asian option price and implied volatility, we find the impact of the correlation on arithmetic Asian option price is more significant than jumps and the correlation can help the double Heston model to capture comprehensive feature of implied volatility.

Exact Simulation of Coupled Wright-Fisher Diffusions

Celia García-Pareja
Karolinska Institutet, Stockholm, Sweden, celia.garcia.pareja@ki.se

Henrik Hult
KTH Royal Institute of Technology, Stockholm, Sweden, hult@kth.se

Timo Koski
KTH Royal Institute of Technology, Stockholm, Sweden, tjtkoski@kth.se

In this talk we introduce an exact rejection algorithm for simulating paths of a family of multivariate Wright-Fisher diffusions, the coupled Wright-Fisher diffusion, which models the co-evolution of multiple genetic traits at different locations on the genoma. Our algorithm uses independent one-dimensional neutral Wright-Fisher diffusions as candidate proposals. The candidates can be sampled exactly by means of existing algorithms and are only needed at a finite number of points. Once a candidate is accepted, the remaining of the path can be recovered by sampling from a neutral multivariate Wright-Fisher bridge, for which we also provide an exact sampling strategy. The technique relies on a modification of the alternating series method and extends existing algorithms that are currently only available for the one-dimensional case. Finally, our algorithm's complexity is derived and its performance evaluated in a simulation study which shows promising results.

Simulation of the time needed by a diffusion process in order to exit from a given interval (WOMS algorithm).

Nicolas Massin
Institut de Mathématiques de Bourgogne, Dijon, France, nicolas.massin@u-bourgogne.fr

For many applications, it is important to describe or simulate the exit time from an interval for a stochastic process. In the particular context of diffusion processes which are solutions of stochastic differential equations, time discretization schemes like Euler scheme are usually used. They permit the simulation of a paths skeleton and lead to an approximation of the exit time as a by-product. The aim of the talk is to present a completely different approach based on the study of the Brownian paths. It is possible to find some domains, called for instance spheroids, such that both the distribution of Brownian exit time from this spheroid and its exit position are well-known.

From that point, we define an iterative procedure : a walk on spheres for the Brownian motion and so called WOMS algorithm. We can show that it is hard to find an algorithm as precise and efficient than this algorithm.

Starting from this particular case, we generalize the iterative procedure in order to deal with a family of diffusion written as functions of a changed time Brownian motion. This strong relation permits to define a new WOMS algorithm by deduction.Theoretical results and numerical illustrations point out the efficiency of such an algorithm. The particular Ornstein-Uhlenbeck case will be presented in details.

[1] Muller, Mervin E., *Some continuous Monte Carlo methods for the Dirichlet problem*, Annals of Mathematical Statistics vol.27, 1956

Unbiased Simulation of Multivariate Jump-Diffusions

Guanting Chen
Stanford University, guanting@stanford.edu

Kay Giesecke
Stanford University, giseceke@staford.edu

Alexander Shkolnik
University of California, Santa Barbara, shkolnik@ucsb.edu

Techniques for the simulation of stochastic differential equations (SDEs) have attracted a significant amount of interest in the Monte Carlo methods and applied probability communities. Recent breakthroughs on generic algorithms for multivariate diffusions include exact schemes (e.g., [1]) and unbiased methods (e.g., [2] & [3]). However, the addition of state-dependent jumps to the SDE presents significant challenges for these approaches.

We develop the first (to our knowledge) method for simulation of multivariate jump-diffusions with general drift, volatility and jump-intensity coefficients. The approach allows one to embed essentially any method for simulating the pure diffusion in between the jump times. While exact sampling of state-dependent jumps is in general not feasible, we propose a change of probability measure to circumvent this obstacle. The change of measure is induced by certain point process martingales and is of independent interest. Numerical results for applications in finance demonstrate the advantages of the algorithm.

[1] J. Blanchet and F. Zhang. *Exact Simulation for Multivariate Ito Diffusions.* preprint arXiv:1706.05124, 2017.

[2] V. Bally, Vlad and A. Kohatsu-Higa, Arturo. *A probabilistic interpretation of the parametrix method* The Annals of Applied Probability, Volume 25, (6), 2015.

[3] P. Andersson and A. Kohatsu-Higa. *Unbiased simulation of stochastic differential equations using parametrix expansions*, Bernoulli, Volume 23, (3) 2018.

Thursday 10:30–12:30, **Room: CB08.03.005**

Monte Carlo Methods in Environmental Applications

Chair: Andrew Zammit-Mangion

Climate inference on Australian daily rainfall using distributed MCMC

Michael Bertolacci
UWA, Perth, Australia, michael.bertolacci@research.uwa.edu.au

Edward Cripps
UWA, Perth, Australia, edward.cripps@uwa.edu.au

Sally Cripps
USyd, Perth, Australia, sally.cripps@sydney.edu.au

John Lau
UWA, Perth, Australia, john.lau@uwa.edu.au

Ori Rosen
UTEP, El Paso, Texas, USA, orosen@utep.edu

We analyse daily rainfall records from across the Australian continent to understand their variability over time and space, as well as to quantify how global climate drivers influence rainfall. The data comprise 294 million daily rainfall measurements made between 1876 and 2015, spanning 17,606 locations. Daily rainfall data is challenging because it has many zeros as well as long-tailed positive values, necessitating models that accommodate mixed discrete-continuous observations. Furthermore, Australia is a large continent with varied rainfall patterns over both time and space. We present a Bayesian hierarchical model that handles these features via a mixture structure incorporating site-specific temporal and climatic influences on both whether and how much rainfall occurs. These influences are summarised in an interpretable manner at another level of the hierarchy. The model has many parameters, and the data set is large, so for estimation we develop a parallel and distributed Markov chain Monte Carlo algorithm. The uncertainty in the resulting inference is fully quantified, providing insight into both the extent and limitations of our understanding of daily rainfall processes.

[1] M. Bertolacci, E. Cripps, S. Cripps, J. Lau, O. Rosen. *Climate inference on daily rainfall across the Australian continent, 1876–2015.* Annals of Applied Statistics, 2018. forthcoming.

Vision and Decision: creating actionable insights from Bayesian Hierarchical Models for the Great Barrier Reef

Petra Kuhnert
CSIRO Data61, Canberra Australia, Petra.Kuhnert@data61.csiro.au

Decision-making happens in everyday life, and for government bodies, it is an important task that is often informed by scientific research. For many problems, particularly those involving the environment, complex models are used to seek understanding of the real world. Accompanying these complex models are uncertainties related to the chosen modelling framework, inputs, and parameters used. To a decision-maker, the predictions arising from these models and their uncertainties are difficult to interpret. Furthermore, the scientist's approach to communicating this information is challenging.

Using a case study from the Great Barrier Reef (GBR), Australia, I will showcase a real world problem that uses Bayesian Hierarhichcal Modelling (BHM) to quantify pollutant load estimates and

uncertainties and present them in a way that is pallatable to decision-makers. The approach is a spatio-temporal statistical model that is mechanistically motivated by a catchment model that captures the underlying physical processes of flow and sediment. The model adopts a dimension reduction approach to accommodate seasonal and spatial characteristics of the system. The method was developed on a spatial river network in the Upper Burdekin subcatchment that consists of 411 nodes over a 20 year period, with limited monitoring conducted throughout the catchment.

Spatio-temporal predictions and posterior distributions from the model will be showcased through a series of four new visualisations that can assist catchment managers with regards to decisions in the catchment. These new visualisations are available as part of the *Vizumap* R package and include the bivariate choropleth map, map pixelation, glyph rotation and exceedance probability maps. Bivariate choropleth maps explore the "blending" of two colour schemes, one representing the estimate and a second representing the margin of error. The second approach uses map pixelation to convey uncertainty. The third approach uses a glyph to represent uncertainty and is what we refer to as glyph rotation. The final map based exploration of uncertainty is through exceedance probabilities.

This research is joint with Dan Gladish (CSIRO Data61), Dan Pagendam (CSIRO Data61), Chris Wikle (University of Missouri) and Lydia Lucchesi (University of Washington).

False Discovery Rates to Detect Signals From Incomplete Spatially Aggregated Data

Noel Cressie
Distinguished Professor, NIASRA, University of Wollongong, Australia , ncressie@uow.edu.au

This paper considers testing for the absence/presence of a spatial signal, where the data are defined on (irregular) small areas resulting from aggregation of fine-resolution pixels. When all pixels are observed, there are a number of ways to test for the presence of a spatial signal in the image; the one we shall focus on is a powerful nonparametric hypothesis-testing approach called the enhanced false discovery rate (EFDR) procedure. However, the methodology relies on data defined on a rectangular spatial domain with regular pixels. This paper generalises the existing methodology so that all irregularities in the small areas can be handled; the idea is to augment EFDR with a Monte Carlo method in spatial statistics known as conditional simulation. EFDR is applied to M conditional simulations resulting in M statistically dependent p-values for testing the null hypothesis of no signal. The final p-value for the hypothesis test of no signal based on the original data is derived. In the case of rejection of the null hypothesis , we give an estimate of the signal at the finest resolution. An environmental application to temperature change in the Asia-Pacific region is given. This is joint work with Hsin-Cheng Huang (Academia Sinica, Taiwan) and Andrew Zammit-Mangion (University of Wollongong, Australia).

Bayesian Atmospheric Tomography for Detection and Estimation of Methane Emissions

Laura Cartwright
University of Wollongong, Australia, lcartwri@uow.edu.au

The detection and quantification of greenhouse-gas fugitive emissions is of both national and global importance. Despite several decades of active research, it remains predominantly an open problem, largely due to model errors and misspecifications that appear at each stage of the flux-inversion processing chain. In 2015, a controlled-methane-release experiment headed by Geoscience Australia was carried out at the CO2CRC Ginninderra site, and a variety of instruments and methods were employed for quantifying the release rate. In this talk I will present a fully Bayesian approach to atmospheric tomography for inferring the flux. The Bayesian framework is designed to account for uncertainty in the measurements, the meteorological data, the temporally varying background concentration, and the atmospheric model itself,

when doing inversion using Markov chain Monte Carlo. I will show that through Bayesian atmospheric tomography we are able to learn about the dispersion model, and propagate uncertainty from unknown parameters within the model to that on the flux. We show the utility of our approach in detecting and quantifying methane emissions from both point and path instruments, using data collected during the Ginninderra experiment. This work is joint with Andrew Zammit-Mangion, Andrew Feitz, and Nicholas Deutscher.

Thursday 15:00-17:00, **Room: CB08.03.002**

Recent advances on Bayesian computation

Chair: David Gunawan

High-dimensional Copula Variational Approximation through Transformation

Michael Stanley Smith
Melbourne Business School, University of Melbourne, Australia, mike.smith@mbs.edu

Rubén Albeiro Loaiza-Maya
Econometrics and Business Statistics, Monash University, Australia, rloaizma@gmail.com

David John Nott
Department of Statistics, National University of Singapore, Singapore, standj@nus.edu.sg

Variational methods are attractive for computing Bayesian inference for highly parametrized models and large datasets where exact inference is impractical. They approximate a target distribution—either the posterior or an augmented posterior—using a simpler distribution that is selected to balance accuracy with computational feasibility. Here we approximate an element-wise parametric transformation of the target distribution as multivariate Gaussian or skew-normal. Approximations of this kind are implicit copula models for the original parameters, with a Gaussian or skew-normal copula function and flexible parametric margins. A key observation is that their adoption can improve the accuracy of variational inference in high dimensions at limited or no additional computational cost. We consider the Yeo-Johnson and G&H transformations, along with sparse factor structures for the scale matrix of the Gaussian or skew-normal. We also show how to implement efficient reparametrization gradient methods for these copula-based approximations. The efficacy of the approach is illustrated by computing posterior inference for three different models using six real datasets. In each case, we show that our proposed copula model distributions are more accurate variational approximations than Gaussian or skew-normal distributions, but at only a minor or no increase in computational cost.

Approximation Properties of Variational Bayes for Vector Autoregressions

Reza Hajargasht
Swinburne University of Technology, rhajargasht@swin.edu.au

Variational Bayes (VB) is a recent approximate method for Bayesian inference. It has the merit of being a fast and scalable alternative to Markov Chain Monte Carlo (MCMC) but its approximation error is often unknown. In this paper, we derive the approximation error of VB in terms of mean, mode, variance, predictive density and KL divergence for the linear Gaussian multi-equation regression. Our results indicate that VB approximates the posterior mean perfectly. Factors affecting the magnitude of underestimation in posterior variance and mode are revealed. Importantly, We demonstrate that VB estimates predictive densities accurately.

Implicit Langevin Algorithms for Sampling From Log-concave Densities

Liam Hodgkinson
University of Queensland,

Robert Salomone
University of New South Wales, r.salomone@unsw.edu.au

Fred Roosta
University of Queensland,

For sampling from a log-concave density, we study implicit integrators resulting from theta-method discretization of the overdamped Langevin diffusion stochastic differential equation. Theoretical and algorithmic properties of the resulting sampling methods for different values of theta and a range of step sizes are established. Our results generalize and extend prior works in several directions. In particular, for theta greater than or equal to 1/2, we prove geometric ergodicity and stability of the resulting methods for all step sizes. We show that obtaining subsequent samples amounts to solving a strongly-convex optimization problem, which is readily achievable using one of numerous existing methods. Numerical examples supporting our theoretical analysis are also presented.

Flexible Density Tempering Approaches for State Space Models with an Application to Factor Stochastic Volatility Models

David Gunawan
UNSW Sydney, Australia, david.gunawan@unsw.edu.au

Robert Kohn
UNSW Sydney, Australia, r.kohn@unsw.edu.au

Chris Carter
UNSW Sydney, Australia, chris.carter@unsw.edu.au

Minh Ngoc Tran
University of Sydney, Australia, minh-ngoc.tran@sydney.edu.au

Density tempering or annealing is a sequential Monte Carlo approach that moves a collection of parameters and latent state variables through a number of levels, with each level having its own target density, in such a way that it is easy to generate both the parameters and latent state variables at the initial level while the target density at the final level is the posterior density of interest. A critical component of the annealing or density tempering method is the Markov move component that is implemented at every stage of the annealing process. The Markov move component effectively runs a small number of Markov chain Monte Carlo iterations for each combination of parameters and latent variables so that they are better approximations to that level of the tempered target density. Duan and Fulop (2015) used a pseudo marginal Metropolis-Hastings (PMMH) approach with the likelihood estimated unbiasedly and a random walk MCMC kernel for the Markov move step of the model parameters, which is inefficient in high dimensions. We propose using instead more flexible Markov move steps that are based on particle Gibbs and Hamiltonian Monte Carlo and demonstrate the proposed methods using a high dimensional factor stochastic volatility model having both a large number of parameters and a large number of latent state variables. An estimate of the marginal likelihood is obtained as a byproduct of the estimation procedure.

Thursday 15:00-17:00, **Room: CB08.03.004**

Random Number Generation

Chair: Pierre L'Ecuyer

Study on upper bounds of sample sizes for two-level tests in NIST SP800-22

Hiroshi Haramoto
Ehime University, Japan, haramoto@ehime-u.ac.jp

NIST SP800-22 is one of the most widely used statistical testing tools for pseudorandom number generators (PRNGs) [1]. It consists of 15 tests (one-level tests) and two additional tests (two-level tests). One of the two-level tests is to measure the uniformity of p-values provided by a certain one-level test via a chi-squared goodness-of-fit test. This two-level test is, in some sense, more powerful than the one-level test. However, determining an appropriate sample size at the second level is not easy.

In this talk, we give an upper bound of the sample size at the second level for each of the following six one-level tests: the Test for the Longest-Run-of-Ones in a Block, the Overlapping Template Matching test, the Linear Complexity test, the Frequency Test within a Block, the Random Excursions test, and the Discrete Fourier Transform test. These bounds are derived by the chi-squared discrepancy [2, 3] which measures discrepancy between the actual distribution of p-values (which may not be uniform even if a true random source is used because of approximation errors in the test) and the uniform distribution $U(0, 1)$.

We also show some experimental results of this two-level test on some PRNGs, which support a justification of the criterion. In addition, we propose a method for increasing the sample size at the second level, which makes the two-level test more sensitive than NIST's recommended usage.

[1] Lawrence E. Bassham, III et al. SP 800-22 Rev. 1a. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. Technical report, National Institute of Standards & Technology, 2010.

[2] Hiroshi Haramoto and Makoto Matsumoto. A Method to Compute an Appropriate Sample Size of a Two-Level Test for the NIST Test Suite. *Monte Carlo and Quasi-Monte Carlo Methods 2016*, pages 283–294, 2018.

[3] Makoto Matsumoto and Takuji Nishimura. A Nonempirical Test on the Weight of Pseudorandom Number Generators. *Monte Carlo and Quasi-Monte Carlo methods 2000*, pages 381–395, 2002.

Tiny Mersenne Twister with Parameterization

Mutsuo Saito
Hiroshima University, Japan, sai10@hiroshima-u.ac.jp

Makoto Matsumoto
Hiroshima University, Japan, m-mat@math.sci.hiroshima-u.ac.jp

Recently, modified xorshift generators, such as Vigna's xorshift128+ generators, are getting popular. This is because (1) fast, (2) compact size (state space is 128 bit) (3) passing Big Crush statistical test in TestU01. Tiny Mersenne Twisters[1] are similar to them, namely, having compact size, fast, so that they match parallel use in GPUs. An advantage of Tiny MTs over these generators is that Tiny MTs are parameterized. Kenji Rikitake offered a table of 2^{20} different parameters, which means that 2^{20} distinct pseudo random number generators are available. Several Tiny MTs are tested by TestU01, and

they all passed. We also show some common flaws in the randomness of the sequence generated by the xorshift128+ family.

[1] M. Saito and M. Matsumoto. *Tiny Mersenne Twister (TinyMT): A small-sized variant of Mersenne Twister*. http://www.math.sci.hiroshima-u.ac.jp/~m-mat/MT/TINYMT/index.html, 2011.

[2] K. Rikitake. *TinyMT pre-calculated parameter list*. https://github.com/jj1bdx/tinymtdc-longbatch, 2013.

A Software Tool to Analyze the Lattice Structure of RNGs

Marc-Antoine Savard
DIRO, Université de Montréal, Canada, ma2112savard@gmail.com

Pierre L'Ecuyer
DIRO, Université de Montréal, Canada, lecuyer@iro.umontreal.ca

Erwan Bourceret
École Polytechnique, Palaiseau, France, ebourceret@hotmail.fr

We present the latest version of LatMRG, a software tool to analyze the lattice structure of multiple recursive and matrix linear congruential generators, simple or combined, and to search for good parameter values when building new generators under various types of constraints on the parameters. We also present examples of what the software can do.

LatMRG contains both executable programs and a C++ library of classes and methods to test for maximal period, construct a basis for the lattice generated by the vectors of successive or non-successive output values from the generator, compute various figures of merit for such lattices or their dual, and much more. For lattice basis reduction and computing shortest nonzero lattice vectors, it uses a sister library named *Lattice Tester*, which in turns relies on the *NTL* library [5] for certain operations on large integers and on lattices. The present software is a complete redesign and rewrite of the old Modula-2 version discussed in [2] and used in [1], for example. The new version was used recently for [3, 4]. It will be available freely on GitHub in the Summer of 2019.

[1] P. L'Ecuyer. Good parameters and implementations for combined multiple recursive random number generators. *Operations Research*, 47(1):159–164, 1999.

[2] P. L'Ecuyer and R. Couture. An implementation of the lattice and spectral tests for multiple recursive linear random number generators. *INFORMS Journal on Computing*, 9(2):206–217, 1997.

[3] P. L'Ecuyer and R. Simard. On the lattice structure of a special class of multiple recursive random number generators. *INFORMS Journal on Computing*, 26(2):449–460, 2014.

[4] P. L'Ecuyer, P. Wambergue, and E. Bourceret. Spectral analysis of the MIXMAX random number generators. *INFORMS Journal on Computing*, 2019. To appear.

[5] V. Shoup. *NTL: A Library for doing Number Theory*. Courant Institute, New York University, New York, NY, July 2018. Available at https://shoup.net/ntl/.

Generating High-Quality Pseudorandom Sequences Using Chaotic True Orbits on Algebraic Integers

Asaki Saito
Future University Hakodate, Japan, saito@fun.ac.jp

Akihiro Yamaguchi
Fukuoka Institute of Technology, Japan, aki@fit.ac.jp

Pseudorandom sequences with high (dimensional) uniformity and quasirandom sequences are known to be very useful for Monte Carlo computation. It is, however, still an interesting question how we can generate a *pseudorandom sequence in the original sense*, i.e., a computer-generated sequence of numbers that appear similar to a typical sample of independently identically distributed random variables. According to ergodic theory, the two simplest chaotic maps on the unit interval, namely the Bernoulli map $x \mapsto 2x \bmod 1$ and the tent map $x \mapsto 1 - |2x - 1|$, can generate ideal random binary sequences. However, it is also known that one cannot simulate these maps with conventional simulation methods such as those using double-precision binary floating-point numbers or arbitrary-precision rational numbers. In this work, by using algebraic integers, we realize generators of high-quality pseudorandom binary sequences based on these chaotic maps. The distinguishing characteristic of our generators is that they generate chaotic true orbits of these maps by exact computation. In particular, we clarify a way to properly prepare a set of seeds (i.e., initial points), which is needed when generating multiple pseudorandom sequences. With this seed selection method, we can choose seeds without bias and can avoid overlaps in latter parts of the pseudorandom sequences derived from them. We also demonstrate through statistical testing that the generated sequences have good statistical properties.

Thursday 15:00-17:00, **Room: CB08.02.002**

Advances in Scalable Monte Carlo

Chair: Leah South

Distributed Bayesian Inference for Varying Coefficient Spatiotemporal Models

Rajarshi Guhaniyogi
University of California Santa Cruz, U.S.A., guhaniyogi@ucsc.edu

Cheng Li
National University of Singapore, Singapore, stalic@nus.edu.sg

Terrance D. Savitsky
U.S. Bureau of Labor Statistics, U.S.A., tds151@gmail.com

Sanvesh Srivastava
The University of Iowa, U.S.A., sanvesh-srivastava@uiowa.edu

Bayesian varying coefficient models based on Gaussian processes are popular in many disciplines because they balance flexibility and interpretability. Markov chain Monte Carlo (MCMC) methods are available to fit these models, but they are inefficient even for moderately large data. Motivated by the task of modeling massive spatiotemporal data, we develop a divide-and-conquer Bayesian method for fitting spatiotemporal varying coefficient models based on multiple output Gaussian processes. Our method partitions the space-time tuples into a large number of overlapping subsets, obtains MCMC samples of parameters and predictions in parallel across the subsets, and combines the subset MCMC samples into an approximate full data posterior. By tuning the stochastic approximation in subset posteriors, we show theoretically that the combined posterior distribution can converge at an optimal rate towards the true underlying surface, and we provide guidance for choosing the number of subsets depending on the analytic properties of Gaussian processes. To improve the efficiency of MCMC sampling, we further develop a new data augmentation scheme based on parameter expansion. We demonstrate the excellent empirical performance of our method across diverse simulations and a real data application to the temperature and precipitation data over 100 years in the U.S.A.

Dimensional Scaling of Piecewise Deterministic Monte Carlo

Joris Bierkens
Delft Institute of Applied Mathematics, joris.bierkens@tudelft.nl

Kengo Kamatani
Osaka University, kamatani@sigmath.es.osaka-u.ac.jp

Gareth Roberts
University of Warwick, Gareth.O.Roberts@warwick.ac.uk

In recent years piecewise deterministic Markov processes (PDMPs) have emerged as a promising alternative to classical MCMC algorithms. In particular these PDMP based algorithms have good convergence properties due to their non-reversible nature and furthermore these algorithms allow for asymptotically exact subsampling. Although many different PDMP based algorithms can be designed, two algorithms play fundamental roles: the Bouncy Particle sampler and the Zig-Zag sampler. In this talk both algorithms will be briefly introduced and a comparison of properties of these algorithms will be presented, with particular attention to recent results concerning scaling with respect to dimension [1].

[1] Joris Bierkens, Kengo Kamatani, and Gareth O. Roberts. High-dimensional scaling limits of piecewise deterministic sampling algorithms. jul 2018.

Quasi-stationary Monte Carlo Methods

Andi Wang
University of Oxford, UK, a.wang@stats.ox.ac.uk

Divakar Kumar
University of Warwick, UK, d.kumar@warwick.ac.uk

Murray Pollock
University of Warwick, UK, m.pollock@warwick.ac.uk

Gareth Roberts
University of Warwick, UK, gareth.o.roberts@warwick.ac.uk

David Steinsaltz
University of Oxford, UK, steinsal@stats.ox.ac.uk

In this talk I will discuss the recent class of quasi-stationary Monte Carlo (QSMC) methods [6] for sampling from density functions. Such methods utilise the probabilistic notion of *quasi-stationarity*, which concerns the asymptotic distributions of killed Markov processes conditional on nonabsorption.

The first such method was the Scalable Langevin Exact (ScaLE) algorithm of [4]. The authors considered killed Brownian motions, whose quasi-stationary distributions coincided with the target distribution. This framework allowed the authors to utilise principled subsampling, which lead to an exact algorithm with a sublinear cost in n, the number of observations in the Bayesian experiment.

In this talk I will focus on a novel QSMC method known as ReScale (Regenerating ScaLE), which takes a stochastic approximation approach to simulating from quasi-stationary distributions, in the flavour of [2]. In ReScaLE, a diffusive particle is augmented with a Poisson clock, not unlike recent PDMP methods [3, 1]. When the clock rings, particle jumps to a new location drawn independently from its empirical occupation measure. The resulting ReScaLE process is non-Markovian and in fact adapts over time to the target. This approach is significantly simpler than the original sequential Monte Carlo approach of ScaLE, while maintaining good scaling properties; in particular subsampling techniques are still available. We established convergence of ReScaLE (in compact settings) in [7]. I will describe applications of ReScaLE to 'tall data' problems.

If time permits I will discuss exciting recent developments: a simple tweak to ReScaLE – fixing the rebirth distribution – enables the construction of a regenerative process which can produce exact draws from the target via coupling from the past [5].

[1] Joris Bierkens, Paul Fearnhead, and Gareth Roberts. The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data. *The Annals of Statistics*, 47(3):1288–1320, 2019.

[2] Jose Blanchet, Peter Glynn, and Shuheng Zheng. Analysis of a stochastic approximation algorithm for computing quasi-stationary distributions. *Advances in Applied Probability*, 48(10):792–811, 2016.

[3] Alexandre Bouchard-Côté, Sebastian J. Vollmer, and Arnaud Doucet. The Bouncy Particle Sampler: A Nonreversible Rejection-Free Markov Chain Monte Carlo Method. *Journal of the American Statistical Association*, 113(522):855–867, 2018.

[4] Murray Pollock, Paul Fearnhead, Adam M. Johansen, and Gareth O. Roberts. The Scalable Langevin Exact Algorithm: Bayesian inference for big data. https://arxiv.org/abs/1609.03436, 2016.

[5] J G Propp and D B Wilson. Example Sampling with Coupled Markov Chains and Applications to Statistical Mechanics. *Random Structures and Algorithms*, 9(2):223–252, 1996.

[6] Andi Q. Wang, Martin Kolb, Gareth O. Roberts, and David Steinsaltz. Theoretical properties of quasi-stationary Monte Carlo methods. *Annals of Applied Probability*, 29(1):434–457, 2019.

[7] Andi Q. Wang, Gareth O. Roberts, and David Steinsaltz. An Approximation Scheme for Quasi-stationary Distributions of Killed Diffusions. https://arxiv.org/abs/1808.07086, 2018.

Scalable Samplers for High-Dimensional Models with Stochastic Nets

Joshua J Bon
Queensland University of Technology, Brisbane, Australia, j.bon@qut.edu.au

The underlying geometry of continuous global-local shrinkage priors is utilised to construct scalable Gibbs samplers for high-dimensional models. The efficiency arises from targeting a truncated density rather than a density defined by a traditional hierarchy for some of the regression parameters. This is especially effective for sparsity-inducing priors, since a large proportion of the parameters are truncated very close to zero in each step of the algorithm. Exploiting this property can reduce expensive matrix operations, and increase the speed of the Gibbs sampler without approximation. We refer to the class of priors considered as stochastic nets, a name which arises due to their particular multivariate geometry. This class includes the horseshoe [1], Dirichlet-Laplace [2] or R2-D2 [3] priors for example. In simulations using multivariate normal models, the standard Gibbs sampler is numerically unstable for $p \in \{2500, 5000\}$ with $n = 100$, whereas the truncated Gibbs sampler is faster and remains stable.

[1] C. M. Carvalho, N. G. Polson, and J. G Scott (2010). The horseshoe estimator for sparse signals, *Biometrika* **97**(2): 465–480.

[2] A. Bhattacharya, D. Pati, N. S. Pillai, and D. B. Dunson (2015). Dirichlet-Laplace priors for optimal shrinkage, *Journal of the American Statistical Association* **110**(512): 1479–1490.

[3] Y. Zhang, B. J. Reich, and H. D. Bondell (2017). High dimensional linear regression via the R2-D2 shrinkage prior, *arXiv preprint arXiv:1609.00046v2* .

Thursday 15:00-17:00, **Room: CB08.03.005**

Applied Probability and Monte Carlo methods III

Chair: Thomas Taimre

Estimating Tails of IID Sums without the MGF

Alexander Shkolnik
University of California, Santa Barbara, shkolnik@ucsb.edu

Matthew Varble
University of California, Santa Barbara, mvarble@math.ucsb.edu

Estimating tails of i.i.d. sums of random variables in $\mathbf{R}^d$ is a classical problem with numerous applications in engineering and the sciences. The standard simulation method (originating from [1]) applies importance sampling to generate the increments of the sum under an exponentially tilted distribution. The sampling measure is obtained from the moment generating function (MGF) of the increment's distribution. Alternatively, when the dimension d is moderate, one may use the same MGF to invert a Fourier transform and recover the tail to arbitrary precision (e.g. see [2]). Both approaches require knowledge of the MGF, which may be intractable when the distrution of the increment is nonstandard.

We propose an importance sampling scheme that is useful in such cases. By defining a counting process associated with the increments of the random walk, we are able to leverage the machinery of stochastic exponentials (see [3] for a standard reference) to develop a change of measure which does not require the MGF. We prove the resulting scheme significantly outperforms crude Monte Carlo and illustrate this with numerics.

[1] D Siegmund. Importance Sampling in the Monte Carlo Study of Sequential Tests. *Ann. Statist.*, 4(4):673–684, 1976.

[2] Joseph Abate and Ward Whitt. The Fourier-series method for inverting transforms of probability distributions. *Queueing Systems*, 10(1):5–87, jan 1992.

[3] Philip E. Protter. *Stochastic Integration and Differential Equations*, volume 21 of *Stochastic Modelling and Applied Probability*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2 edition, 2005.

Efficient simulation of tail probabilities for subexponential sums with dependent random weights

Nicholas C. A. Johnson
The University of Queensland, Australia,

Thomas Taimre
The University of Queensland, Australia, t.taimre@uq.edu.au

This paper deals with the problem of efficiently estimating extreme tail probabilities for the sum of n dependent and non-identically-distributed random variables $H_1, \ldots, H_n$ where each summand $H_i = \Theta_i X_i$ is the product of a subexponential random variable X_i and a random weight Θ_i. Motivated by applications to insurance problems, the sequence of random weights $\Theta_1, \ldots, \Theta_n$ considered here has a recursive dependence structure driven by an underlying sequence of iid subexponential random variables $R_1, \ldots, R_n$. A novel Asmussen–Kroese-type estimator for the tail probabilities is presented, along with numerical results which suggest the estimator enjoys good efficiency properties.

Efficient Rare Event for Heavy-tailed Chance Constrained Optimization

José Blanchet
Stanford University, USA, jose.blanchet@gmail.com

Fan Zhang
Stanford University, USA, fzh@stanford.edu

Bert Zwart
CWI, NL, bert.zwart@cwi.nl

Chance constrained optimization (CC-opt) refers to a class of problems in which the optimizer wishes to minimize a convex function (typically linear) subject to a low tolerance probability in the violation of an outcome which can be expressed as a convex function being large. So, if the tolerance probability is zero, the problem reduces to a convex optimization problem. CC-opt arises naturally in a wide range of modeling settings. Unfortunately, however, CC-opt problems are typically NP-hard and therefore are challenging to solve. In the context of regularly varying distributions, we provide novel algorithms with probably efficient running times as the tolerance probability tends to zero (i.e. violation is a rare event).

Uncertainty Quantification for Stochastic Approximation Limits and applications to risk/performance metrics in finance

Linda Chamakh
BNP Paribas, France, linda.chamakh@hotmail.fr

Stéphane Crepey
Evry University, France, stephane.crepey@univ-evry.fr

Gersende Fort
CNRS Institut de Mathématiques de Toulouse, France, Gersende.Fort@math.univ-toulouse.fr

Emmanuel Gobet
Ecole polytechnique, France, emmanuel.gobet@polytechnique.edu

Uladzislau Stazhynski
Ecole polytechnique, France, uladzislau.stazhynski@polytechnique.edu

We analyze the uncertainty quantification for the limit of a Stochastic Approximation (SA) algorithm. In our setup, this limit ζ is defined as the zero of an intractable function and is modeled as uncertain through a parameter θ: we aim at deriving the probabilistic distribution of $\zeta(\theta)$, given a probability distribution for θ. We introduce the so-called Uncertainty Quantification for SA (UQSA) algorithm, an SA algorithm in increasing dimension for computing the basis coefficients of a chaos expansion of $\zeta(\theta)$ on an orthogonal basis of a suitable Hilbert space. Almost convergence, in the Hilbert space, of UQSA is established under mild, tractable conditions. Applications to uncertain Value-at-Risk and Expected Shortfall of some line of risk, or to uncertain Sharpe Ratio of Quantitative Investment Strategy will be discussed.

Friday 9:00-10:30, **Room: CB08.02.002**

Nets, Sequences, and Discrepancy

Chair: Wolfgang Schmid

Hyperuniform point sets on flat tori

Tetiana Stepaniuk
RICAM, Austrian Academy of Sciences, Austria, tania_stepaniuk@ukr.net

We study hyperuniformity on flat tori. The concept of hyperuniformity had been introduced by S. Torquato and F. Stillinger [3] to measure regularity of distributions of infinite particle systems in R^d. Hyperuniform point sets on the unit sphere have been studied by J. Brauchart, P. Grabner, W. Kusner and J. Ziefle in [1] and [2]. We show that point sets which are hyperuniform for large balls, small balls or balls of threshold order on the flat tori are uniformly distributed. Moreover, it is also shown that QMC–designs sequences for Sobolev classes, probabilistic point sets (with respect to jittered samplings) and some determinantal point process are hyperuniform.

[1] *J. S. Brauchart, P. J. Grabner, and W. Kusner.* Hyperuniform point sets on the sphere: Deterministic aspects, Constructive Approximation (2018), 1–17.

[2] *J. S. Brauchart, P. J. Grabner, W. B. Kusner, and J. Ziefle.* Hyperuniform point sets on the sphere: probabilistic aspects, arXiv:1809.02645.

[3] S. Torquato and F. H. Stillinger. *Local density fluctuations, hyperuniformity, and order metrics.* Phys. Rev. E 68 (2003), no. 4, 041–113.

Characterization of Digital $(0, 2)$-Sequences in Base 2 and Its Application

Kosuke Suzuki
Hiroshima University, Japan, kosuke-suzuki@hiroshima-u.ac.jp

The notions of a (t, m, s)-net and a (t, s)-sequence in base b represent how uniformly distributed the points from the net or the sequence are in the b-adic elementary intervals in $[0, 1]^s$. A powerful method to construct (t, m, s)-nets and (t, s)-sequences is the digital construction. Our purpose is to characterize digital nets and sequences with best possible quality parameter t. We give a characterization of all $m \times m$ generating matrices (C_1, C_2, C_3) which generate a digital $(0, m, 3)$-net in base 2 and all infinite matrices (C_1, C_2) which generate a digital $(0, 2)$-sequence in base 2. As an application, we show that digital nets consisting of 3-dimensional points which represent 3 consecutive numbers from any combined Tausworthe generators are not $(0, m, 3)$-nets.

This talk is based on a joint work with Roswitha Hofer and one with Makoto Matsumoto and Hiroki Kajiura.

On The Dependence Structure Of Scrambled (t, m, s)-Nets

Jaspar Wiart
RICAM Linz, Austria, jaspar.wiart@ricam.oeaw.ac.at

Christiane Lemieux
University of Waterloo, Ontario, clemieux@uwaterloo.ca

We study the dependence structure of scrambled (t, m, s)-nets and show that they have a negative lower/upper orthant dependence structure if and only if $t = 0$. This study allows us to gain a deeper understanding about the classes of functions for which the variance of estimators based on scrambled $(0, m, s)$-nets can be proved to be no larger than that of a Monte Carlo estimator.

Friday 9:00-10:30, **Room: CB08.03.004**

Network Reliability and the Splitting Method

Chair: Gerardo Rubino

Limit distributions of the upper order statistics for the Lévy-frailty Marshall-Olkin distribution

Javiera Barrera
Faculty of Engineering and Science, Universidad Adolfo Ibáñez, Santiago, Chile,
javiera.barrera@uai.cl

Guido Lagos
Industrial Engineering Department, University of Santiago, Santiago, Chile, guido.lagos@usach.cl

The Marshall-Olkin (MO) distribution has been considered a key model in reliability theory and in risk analysis, where it is used to model the lifetimes of dependent components or entities of a system and dependency is induced by "shocks" that hit one or more components at a time [1]. Of particular interest is the Lévy-frailty subfamily of the Marshall-Olkin (LFMO) distribution [2], since it has few parameters and because the nontrivial dependency structure is driven by an underlying Lévy subordinator process. The main contribution of our work is that we derive the precise asymptotic behavior of the upper order statistics of the LFMO distribution. More specifically, we consider a sequence of n univariate random variables jointly distributed as a multivariate LFMO distribution and analyze the order statistics of the sequence as n grows. Our main result states that if the underlying Lévy subordinator is in the normal domain of attraction of a stable distribution with index of stability α, then, after certain logarithmic centering and scaling, the upper order statistics converge in distribution to a stable distribution if $\alpha > 1$ or a simple transformation of it if $\alpha < 1$. Our result is especially useful in network reliability and systemic risk, when modeling the lifetimes of components in a system using the LFMO distribution, as it allows to understand the behavior of systems that rely on its last working components. Our result can also give easily computable confidence intervals for these components, provided that a proper convergence analysis is carried out first.

[1] Cherubini, U., Durante, F., Mulinacci, S.: Marshall-Olkin Distributions — Advances in Theory and Applications: Bologna, Italy, October 2013, vol. 141. Springer (2015)

[2] Mai, J.F., Scherer, M.: Reparameterizing Marshall-Olkin copulas with applications to sampling. Journal of Statistical Computation and Simulation **81**(1), 59–78 (2011)

Splitting–based method for network reliability estimation

Héctor Cancela
Facultad de Ingeniería, UDELAR, Montevideo, Uruguay, cancela@fing.edu.uy

Leslie Murray
Facultad de Ingeniería, UNR, Rosario, Argentina, lmurray@fceia.unr.edu.ar

Gerardo Rubino
INRIA Rennes–Bretagne Atlantique, France, Gerardo.Rubino@inria.fr

This work describes a variance reduction technique proposed to make efficient estimations of the unreliability of highly reliable stochastic flow networks (SFNs).

The Creation Process and the Destruction Process [1] are two algorithms that transform a static network model into a dynamic one. They are the basis of different variance reduction methods designed to make efficient reliability estimations on highly reliable networks for which links can only assume two possible values (operational or failed), and independently from each other (as an example, see our paper [2]). In a recent work [3], we have extended the Creation Process to network models for which links can assume more than two values—e.g. stochastic flow networks—and we designed a Splitting-based method that relies on these types of models for the highly reliable case.

In this work we are interested in the more general case of dependent components, where dependency is captured by the Marshall–Olkin copula. We moved to the Destruction Process here, because it fits the Marshall-Olkin model much better than the Creation Process since the network starts operational and fails due to the successive failures (shocks). However, applying Splitting to the Destruction Process, as we did on the Creation Process, is extremely inefficient. We then describe a new way to apply Splitting in this more general framework, that leads to significant improvements in variance reduction as in the case of independent components.

[1] T. Elperin, I. B. Gertsbakh, and M. Lomonosov. *Estimation of Network Reliability Using Graph Evolution Models.* IEEE Transactions on Reliability 40.5, pp. 572–581, 1991.

[2] L. Murray, H. Cancela, and G. Rubino. *A Splitting algorithm for network reliability estimation.* IIE Transactions 45.2, pp. 177–189, 2013.

[3] H. Cancela, L. Murray and G. Rubino. *Efficient Estimation of Stochastic Flow Network Reliability.* IEEE Transactions on Reliability, accepted January 22, 2019.

The Multilevel Splitting Method for Static Problems: Applications to Wireless Communications Systems Performances

Daniel Mackinlay
UNSW Sydney, Australia, d.mackinlay@unsw.edu.au

Nadhir Ben Rached
RWTH Aachen University, Germany, benrached@uq.rwth-aachen.de

Zdravko Botev
UNSW Sydney, Australia, botev@unsw.edu.au

Raúl Tempone
Alexander von Humboldt Professor, RWTH Aachen University, Germany
& KAUST, Saudi Arabia, tempone@uq.rwth-aachen.de

Mohamed-Slim Alouini
KAUST, Saudi Arabia, slim.alouini@kaust.edu.sa

We aim to use the multilevel splitting method for the performance analysis of wireless communication systems. In its original form, the multilevel splitting algorithm cannot be applied to static problems, since it requires an underlying Markov process that we can split. Our novel idea is to embed the static problem within a continuous time Markov process in such a way that, the Markov process has exactly the static distribution at a particular time instant. With this embedding, we will be able to employ a version of the dynamic splitting algorithm that suits our purposes.

Friday 9:00-10:30, **Room: CB08.03.002**

Methods for Acceleration and Variance Reduction

Chair: Florian Puchhammer

Constrained Hamiltonian Monte Carlo for PDE Inverse Problems

Khai Xiang Au
National University of Singapore, khai@u.nus.edu

Matthew Graham
National University of Singapore, m.m.graham@nus.edu.sg

Alexandre Thiery
National University of Singapore, a.h.thiery@nus.edu.sg

We consider the problem of inference in models of the form $y = \mathcal{G}(x)+\eta$ where y is a vector of observed variables, x is the quantity of interest to be inferred, $\mathcal{G}$ is the observation operator and $\eta \sim \mathcal{N}(0, \sigma^2 I)$ is a vector of additive noise in the observations. Specifically, we consider the case where $\mathcal{G}$ involves solving a system of partial differential equations (PDEs) in x. In such models, the posterior mass concentrates in the neighbourhood of a manifold as σ tends to zero. As a result, the efficiency of MCMC methods deteriorates due to the need to take increasingly small steps.

In this work, we present a constrained HMC algorithm that is robust to small σ values, i.e. low noise. Taking the observations generated by the model to be constraints on the prior, we define a manifold on which the constrained HMC algorithm generate samples. By exploiting the geometry of the manifold, our algorithm is able to take larger step sizes than more standard MCMC methods, resulting in a more efficient sampler. We explain the setup of manifold HMC and additional steps necessary to ensure the reversibility of the resulting Markov chains. Finally, we apply the algorithm to several PDE inverse problems to verify its efficiency in the low noise regime.

Fast Haar Convolution for Geometric Deep Learning

Ming Li
La Trobe University, Australia, ming.li@latrobe.edu.au

Zheng Ma
Princeton University, USA, zhengm@princeton.edu

Yu Guang Wang
The University of New South Wales, Australia, yuguang.wang@unsw.edu.au

Xiaosheng Zhuang
City University of Hong Kong, Hong Kong, xzhuang7@cityu.edu.hk

Graph Neural Networks (GNNs) have become a topic of intense research recently due to their powerful capability in high-dimensional classification and regression tasks for graph-structured data. However, as GNNs typically define the graph convolution by the orthonormal basis for the graph Laplacian, they suffer from high computational cost when the graph size is large. This paper introduces the Haar basis, a sparse and localized orthonormal system for graph, constructed from a coarse-grained chain on the graph. The graph convolution under Haar basis — the Haar convolution can be defined accordingly for GNNs. The sparsity and locality of the Haar basis allow Fast Fourier Transforms (FFTs) on graph, by which a fast evaluation of Haar convolution between the graph signals and the filters can be achieved. We conduct preliminary experiments on GNNs equipped with Haar convolution, which can obtain state-of-the-art results for a variety of geometric deep learning tasks.

[1] Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., and Vandergheynst, P. Geometric deep learning: going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

[2] Bruna, J., Zaremba, W., Szlam, A., and Lecun, Y. Spectral networks and locally connected networks on graphs. In *ICLR*, 2014.

[3] Li, M., Ma, Z., Wang, Y. G., and Zhuang, X. Fast Haar Convolution for Geometric Deep Learning. *submitted*.

[4] Wang, Y. G. and Zhuang, X. Tight framelets and fast framelet filter bank transforms on manifolds. *Applied and Computational Harmonic Analysis*, 2018.

Variance Reduction for Chemical Reaction Networks with Array-RQMC

Amal Ben Abdellah
DIRO, Université de Montré al, Canada, amal.ben.abdellah@umontreal.ca

Pierre L'Ecuyer
DIRO, Université de Montré al, Canada, lecuyer@iro.umontreal.ca

Florian Puchhammer
DIRO, Université de Montré al, Canada, florian.puchhammer@umontreal.ca

Mathematical processes in molecular biology frequently rely on path simulation algorithms for discrete time Markov chains. For instance, the fixed step τ-leap method by Gillepsie [2], a modification of kinetic Monte Carlo, can be used to study well-mixed chemically reacting systems. For the simulation of the sample paths, crude Monte Carlo (MC) is commonly seen as a viable approach, however, one might expect a smaller variance for more refined techniques. Very recently, Beentjes and Baker [1] used randomized quasi-Monte Carlo (RQMC) for simulating chemical reaction networks with τ-leaping but the gain in terms of variance was limited. We examine the application of a different path simulation algorithm, Array-RQMC, to this problem, which has already proven to significantly outperform MC in many other applications [3, 4, 5]. In the Array-RQMC algorithm, many chains are efficiently simulated in parallel, however, the states need to be sorted after each step. Even though standard sorting procedures exist, they sometimes lack efficiency when the states are high-dimensional.

In this talk we show emprically that combining Array-RQMC with τ-leaping for well-mixed chemical reaction networks can lead to a significantly faster convergence of the variance than MC. Moreover, we demonstrate that sorting with respect to a specific *importance function* of the states, which assigns to each state a real value, can clearly outperform standard sorting procedures in this setting in terms of both efficiency and variance.

[1] C. H. L. Beentjes and R. E. Baker. *Quasi-Monte Carlo Methods Applied to Tau-Leaping in Stochastic Biological Systems.* Bulletin of Mathematical Biology, Springer, 2018.

[2] D. T. Gillespie. *Approximate Accelerated Stochastic Simulation of Chemically Reacting Systems.* The Journal of Chemical Physics, Volume 115(4):1716-1733, 2001.

[3] P. L'Ecuyer, D. Munger, C. Lécot, and B. Tuffin. *Sorting Methods and Convergence Rates for Array-RQMC: Some Empirical Comparisons.* Mathematics and Computers in Simulation, Volume 143:191–201, 2018.

[4] P. L'Ecuyer, C. Lécot, and A. L'Archevê que-Gaudet. *On Array-RQMC for Markov Chains: Mapping Alternatives and Convergence Rates.* P. L'Ecuyer, A.B. Owen (Eds.), Monte Carlo and Quasi-Monte Carlo Methods 2008, Springer-Verlag, Berlin, pp. 485-500, 2009.

[5] P. L'Ecuyer, C. Lécot, and B. Tuffin. *A Randomized Quasi-Monte Carlo Simulation Method for Markov Chains.* Operations Research, Volume 56(4):958–975, 2008.

Friday 9:00-10:30, **Room: CB08.03.005**

Monte Carlo methods in change-point detection

Chair: Georgy Sofronov

Monte Carlo Methods for Spatial Clustering

Nishanthi Raveendran
Macquarie University, Sydney, Australia, nishanthi.raveendran@mq.edu.au

Georgy Sofronov
Macquarie University, Sydney, Australia, georgy.sofronov@mq.edu.au

Spatial clustering is an important component of spatial data analysis because of a possible heterogeneity of the data. Ignoring the spatial segmentation may lead to incorrect interpretation of parameters in the corresponding statistical model while spatial clustering allows us to develop an appropriate statistical model for each homogeneous domain. The problem of finding regional homogeneous domains is known as segmentation, partitioning or clustering which is commonly used in disease surveillance, spatial epidemiology, population genetics, landscape ecology, crime analysis and many other fields. For example, in epidemiology and public health, it is known that the disease risk varies across space and it is important to identify regions of safety and regions of risks. In this study, we focus on identifying homogeneous domains in binary data, which indicate the presence or absence of a certain plant species which are observed over a two-dimensional lattice. We consider this clustering problem within the change-point detection methodology. We developed a sequential importance sampling algorithm to estimate the average surface profile explaining the hetrogeneity of data [1]. We provide numerical experiments, which illustrate the effectiveness of this method and compare with the results obtained via MCMC algorithm within the generalized Gibbs sampler [2].

[1] N. Raveendran and G. Y. Sofronov. Identifying Clusters in Spatial Data via Sequential Importance Sampling. *Recent Advances in Computational Optimization, Studies in Computational Intelligence.* Springer,Cham, vol. 795, 175–189, 2019.

[2] J. M. Keith, D. P. Kroese, and D. Bryant A Generalized Markov Sampler *Methodology and Computing in Applied Probability.* 6(1): 29–53, 2004.

A Monte Carlo Method for Multiple Change-Point Detection in a Segmented ARMA Model

Lijing Ma
Macquarie University, Sydney, Australia, lijing.ma@mq.edu.au

Georgy Sofronov
Macquarie University, Sydney, Australia, georgy.sofronov@mq.edu.au

Justin Wishart
Macquarie University, Sydney, Australia, justin.wishart@mq.edu.au

In time series analysis, the model structure can change abruptly over time. In this talk, we propose a novel Monte Carlo method to detect these change-points that divide a time series into individual segments. We assume that within each segment the data follow a different autoregressive moving average (ARMA) model. The challenge here is the uncertainty of the location of change-points and the parameters of the model in each segment. This problem can be approached as a combinatorial optimization problem. The proposed algorithm is used to solve the optimization problem. The results of numerical experiments are provided to illustrate our method.

Change-Point Detection via the Cross-Entropy Method

Irene Hudson
Royal Melbourne Institute of Technology (RMIT), Melbourne, Australia, irene.hudson@rmit.edu.au

Madawa W. Jayawardana
Swinburne University of Technology, Melbourne, Australia, mjayawardana@swin.edu.au

Georgy Sofronov
Macquarie University, Sydney, Australia, georgy.sofronov@mq.edu.au

Change-point problems (or break point problems, disorder problems) can be considered one of the central issues of mathematical statistics, connecting asymptotic statistical theory and Monte Carlo methods, frequentist and Bayesian approaches, fixed and sequential procedures. In many real applications, observations are taken sequentially over time, or can be ordered with respect to some other criterion. The basic question, therefore, is whether the data obtained are generated by one or by many different probabilistic mechanisms. The change-point problem arises in a wide variety of fields, including bioinformatics, biomedical signal processing, speech and image processing, seismology, industry (e.g. fault detection) and financial mathematics. In this talk, we consider the Cross-Entropy method (for example, see [1]) to change-point detection to find estimates of change-points as well as parameters of the process on each segment.

[1] D. P. Kroese, T. Taimre, and Z. I. Botev. *Handbook of Monte Carlo methods*. John Wiley & Sons, 2011.

List of session chairs

An, Ziwen **Tuesday** 15:00

Blanchet, Jose **Wednesday** 10:30

Clare, Gregoire **Wednesday** 13:30

Dick, Josef **Wednesday** 9:00

Drovandi, Christopher **Monday** 10:30 **Monday** 13:30

Elvira, Victor **Monday** 10:30

Gobet, Emmanuel **Monday** 15:00 **Thursday** 9:00

Gunawan, David **Thursday** 15:00

Heinrich, Stefan **Monday** 15:00

Hickernell, Fred **Tuesday** 9:00

Hironaka, Tomohiko **Thursday** 10:30

Kaarnioja, Vesa **Tuesday** 10:30

Karvonen, Toni **Wednesday** 13:30

Kohn, Robert **Monday** 10:30

Kritzer, Peter **Monday** 10:30

Kuo, Frances **Friday** 11:00

L'Ecuyer, Pierre **Thursday** 13:30 **Thursday** 15:00

Li Feng, **Thursday** 10:30

MacNamara, Shev **Tuesday** 10:30

Müller-Gronbach, Thomas **Tuesday** 15:00

Pasupathy, Raghu **Tuesday** 10:30

Penev, Spiro **Wednesday** 10:30

Puchhammer, Florian **Friday** 9:00

Quiroz, Matias **Monday** 9:05 **Monday** 15:00

Rhee, Chang-Han **Tuesday** 15:00 **Wednesday** 13:30

Ridder, Ad **Monday** 15:00

Rubino, Gerardo **Friday** 9:00

Sabanis, Sotirios **Tuesday** 10:30

Schmid, Wolfgang **Friday** 9:00

Shkolnik, Alexander **Thursday** 10:30

Sloan, Ian **Tuesday** 13:30 **Tuesday** 15:00

Sofronov, Georgy **Friday** 9:00

South, Leah **Thursday** 15:00

Taimre, Thomas **Thursday** 15:00

Wang, Liangliang **Wednesday** 10:30

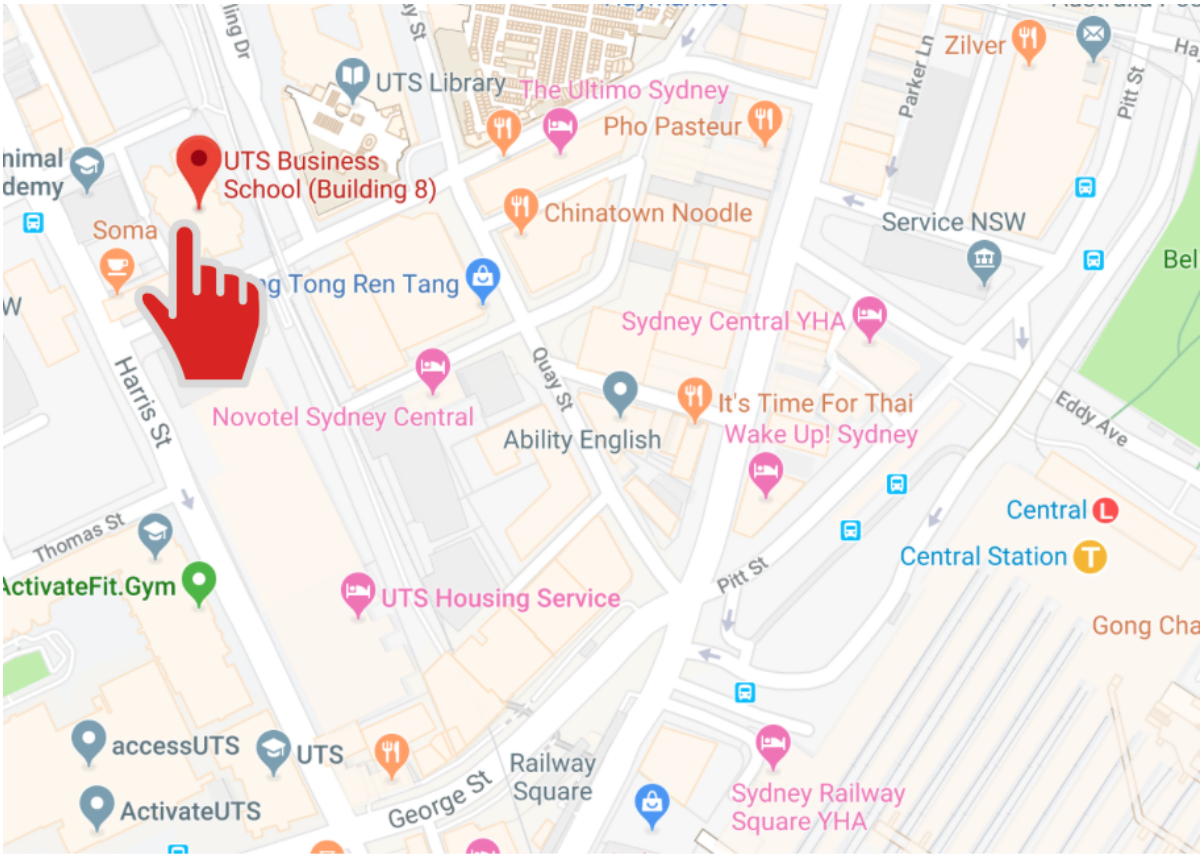
Xu, Ming **Wednesday** 13:30

Yaroslavtseva, Larisa **Wednesday** 10:30

Zammit-Mangion, Andrew **Thursday** 10:30

Index

Conference venue



Restaurant