MCM 2021

16.8 – 20.8.2021

PROGRAM AND GENERAL INFORMATION

Contents

Welcome to the MCM 2021

Welcome to the *13th International Conference on Monte Carlo Methods and Applications, MCM 2021*, the first virtual one of this series due to the world wide effects of the COVID-19 Pandemic. We hope that you will have an interesting and productive week!

We wish to thank all the contributors and are thankful for their participation in making this conference a Monte Carlo highlight for 2021 – despite the extraordinary circumstances. We hope you will have the opportunity to both learn new ideas and start to build new collaborations and friendships.

Our scientific program features nine one hour (including questions) invited plenary talks. We also have more than 150 invited and contributed talks of 30 minutes each (again including questions). The different time zones of the participants are a challenge and we are trying to take that into account with an adapted schedule.

We again hope that you will enjoy your time in this virtual conference 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.

Andreas Neuenkirch and the Local Organizers

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.

The locations of the 14 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	**virtual conference** *hosted in Mannheim, Germany*
2023	???

2.1 Committees

Local Organizing Committee

Karin Bühl
Eike Fokken
Jan Friedrich
Xaver Kohlbrenner
Annalena Mickel
Andreas Neuenkirch (chair)
Peter Parczewski
Thomas Schillinger

Steering Committee

Zdravko Botev	(UNSW Sydney, Australia)
Ronald Cools	(KU Leuven, Belgium)
Emmanuel Gobet	(École Polytechnique, Palaiseau, France)
Pierre L'Ecuyer (chair)	(Université de Montréal, Canada)
Thomas Müller-Gronbach	(Universität Passau, Germany)
Karl Sabelfeld	(Russian Academy of Science, Russia)
Wolfgang Schmid	(Universität Salzburg, Austria)
Bruno Tuffin	(INRIA Rennes Bretagne-Atlantique, France)

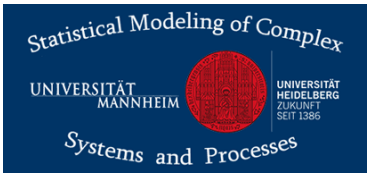
Program Committee

Zdravko Botev	(UNSW Sydney, Australia)
Nicolas Chopin	(ENSAE, France)
Ronald Cools	(KU Leuven, Belgium)
Ton Dieker	(Columbia University, USA)
Steffen Dereich	(WWU Münster, Germany)
Arnaud Doucet	(University of Oxford, United Kingdom)
Victor Elvira	(University of Edinburgh, United Kingdom)
Mark Girolami	(Imperial College, United Kingdom)
Emmanuel Gobet	(École Polytechnique, France)
Stefan Heinrich	(TU Kaiserslautern, Germany)
Aicke Hinrichs	(JKU Linz, Austria)
Fred Hickernell	(Illinois Institute of Technology, USA)
Alexander Keller	(NVIDIA, Germany)
Peter Kritzer	(Austrian Academy of Sciences, Austria)
Frances Kuo	(UNSW Sydney, Australia)
Pierre L'Ecuyer	(University of Montreal, Canada)
Christiane Lemieux	(University of Waterloo, Canada)
Makoto Matsumoto	(University of Hiroshima, Japan)
Thomas Müller-Gronbach	(Universität Passau, Germany)
Dirk Nuyens	(KU Leuven, Belgium)
Chris Oates	(Newcastle University, United Kingdom)
Art Owen	(Stanford University, USA)
Raghu Pasupathy	(Purdue University, USA)
Friedrich Pillichshammer	(JKU Linz, Austria)
Christian Robert	(Université Paris-Dauphine, France)
Karl Sabelfeld	(Russian Academy of Science, Russia)
Wolfgang Schmid	(Universität Salzburg, Austria)
Michaela Szölgyenyi	(Universität Klagenfurt, Austria)
Bruno Tuffin	(INRIA Rennes, France)
Liangliang Wang	(Simon Fraser University, Canada)

Sponsors

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

- Universität Mannheim
- DFG-Research Training Group 1953 "Statistical Modeling of Complex Systems"



Practical information

MCM 2021 will be a virtual conference. We will use zoom for the talks and gather.town for the coffee breaks and informal discussions. Links and access codes will be provided to registered participants via e-mail **until 13.8.2021**.

During the conference there will be seven zoom channels:

- CHANNEL 1 for the plenary talks,
- CHANNEL 2-6 for the special sessions and contributed talks,
- CHANNEL 7 as a Help Desk.

4.1 Tools and Communication

zoom

We assume that everybody is familiar with zoom now – after more than one year of digital talks and teaching. If not, please have a look on the webpage https://zoom.us/. We recommend to install zoom on your computer for a stable connection.

gather.town

As a substitute for the coffee breaks and informal discussions we will use gather.town. Here you will have your own avatar (in retro graphics of the 1990s) and you can walk around to see who is present. If you approach other people (or groups of people) you will be able to talk to the group. Moreover, there will be several virtual whiteboards for discussions. For further information please have a look on the webpage https://gather.town/.

Question hour

For remaining technical questions, we will offer two question hours:

- Thursday, 12.8, 16:00 – 17:00 (UTC+2)
- Friday, 13.8, 10:00 – 11:00 (UTC+2)

If you want to participate in one of the question hours, please send an e-mail to *mcm2021@mail.uni-mannheim.de* until **until 10.8.2021**

Help Desk

There will be a Help Desk during the conference. It can reached by writing an e-mail to `mcm2021@mail.uni-mannheim.de` or the designated zoom channel.

4.2 Presentations

Instructions for all participants

Please mute your microphone during the talks (unless you are a chair or a speaker). If you have a question, please use the "CHAT" function of zoom to announce it to the chair and wait until the chair addresses you.

Please be "zoom-disciplined". We are expecting some of the sessions and talks to have many participants.

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: use the "SHARE SCREEN" option of zoom for your presentation and turn on your video. If you are not familiar with zoom and this way of a giving a scientific talk, please familiarize yourself with this before the conference.

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. For handling the announcement of questions by participants please use the "CHAT" option of zoom. Please turn on your microphone and video as a chair of the session.

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.

In case a speaker does not show up, please leave that slot empty and wait for the announced time in the program to continue.

Program Overview

5.1 Plenaries

See also https://www.uni-mannheim.de/mcm-2021/program/plenary-talks/. All plenaries are on Zoom Channel 1.

Zdravko Botev. *Ridge Regression and variants: 50 years later*

Lester Mackey. *Probabilistic Inference and Learning with Stein's Method*

Sandeep Juneja. *Shift, Scale and Reset Smaller Models to Estimate Larger Ones: Agent-based Simulators in Epidemiology*

Larisa Yaroslavtseva. *On strong approximation of stochastic differential equations with a discontinuous drift coefficient*

Mireille Bossy. *Some numerical aspects of SDEs modelling particle dynamics in turbulent flow*

Elena Akhmatskaya. *Hamiltonian Monte Carlo: Standard Practices Revisited*

Heping Wang. *Tractability in the randomized setting*

Rob Scheichl. *Multilevel Quasi-Monte Carlo Methods for PDE Eigenvalue Problems*

Daniel Rudolf. *Slice Sampling*

5.2 Special Sessions

See also https://www.uni-mannheim.de/mcm-2021/program/special-sessions/.

Stein's Method: Zoom Channel 2
Organized by: Chris J. Oates (Newcastle University) and Francois-Xavier Briol (University College London)

New directions in non-reversible Markov chain / process Monte Carlo: Zoom Channel 2
Organized by: Kengo Kamatani (ISM, Japan) and Gareth Roberts (University of Warwick)

Probabilistic Numerical Methods: Zoom Channel 2
Organized by: Toni Karvonen and Jonathan Cockayne (The Alan Turing Institute)

Quasi-Monte Carlo Software: Zoom Channel 6
Organized by: Fred Hickernell (Illinois Institute of Technology)

Variance Reduction Techniques: Zoom Channel 3
Organized by: Nadhir Ben Rached (RWTH Aachen),Chiheb Ben Hammouda (RWTH Aachen) and Raul Tempone (RWTH Aachen, KAUST)

Monte Carlo methods for discontinuous functions: Zoom Channel 3
Organized by: Sebastian Krumscheid (RWTH Aachen) and Abdul-Lateef Haji-Ali (Heriot-Watt University)

Theory and Applications of Particle Systems: Zoom Channel 3
Organized by: Abdul-Lateef Haji-Ali (Heriot-Watt University) and Raul Tempone (RWTH Aachen, KAUST)

Numerical Methods for SDEs with Boundary Issues: Zoom Channel 3
Organized by: Madalina Deaconu (INRIA, France) and Francisco Bernal (Carlos III of Madrid University)

Stochastic Computation and Complexity: Zoom Channel 4
Organized by: Stefan Heinrich (TU Kaiserslautern), Thomas Müller-Gronbach (Universität Passau) and Larisa Yaroslavtseva (Universität Passau)

Stochastic partial differential equations – Simulation and Modelling: Zoom Channel 4
Organized by: Andreas Roessler (Universität Lübeck) and Claudine von Hallern (Universität Hamburg)

Recent advances on the Laplace approximation and related sampling methods: Zoom Channel 6
Organized by: Daniel Rudolf (University of Göttingen), Claudia Schillings (University of Mannheim), Björn Sprungk (TU Bergakademie Freiberg) and Philipp Wacker (FAU Erlangen-Nürnberg)

Uncertainty quantification for hyperbolic partial differential equations: Zoom Channel 5
Organized by: Simone Göttlich and Thomas Schillinger (University of Mannheim)

Quantitative Aspects of SDEs and Applications: Zoom Channel 4
Organized by: Stefan Geiss (University of Jyväskylä)

Wasserstein gradient flows and their applications: Zoom Channel 5
Organized by: Daniel Adams (University of Edinburgh), Goncalo dos Reis (University of Edinburgh), Hong Duong (University of Birmingham)

5.3 Contributed Talks

The contributed talks will be grouped into the following Sessions:

Markov chain Monte Carlo and related topics: Zoom Channel 2

Numerical Methods for SPDEs: Zoom Channel 6

MCQMC in Statistics: Zoom Channel 6

Monte Carlo Methods for PDMPs and ODEs: Zoom Channel 5

Applications in Physics and Engineering: Zoom Channel 5

Regression and Learning: Zoom Channel 3

Optimization: Zoom Channel 5

Discrepancy and Dispersion: Zoom Channel 6

Sampling Methods: Zoom Channel 5

MCQMC Integration: Zoom Channel 6

Schedule and Abstracts

The conference will run on seven zoom channels. Channel 1 is reserved for plenary talks, while Channel 2 to 6 cover the Special Sessions and Contributed Talks. Channel 7 is the help desk.

6.1 Zoom Channel 1 – Plenary Talks

back to table of contents

Clicking on the name of the speaker will lead you to the abstract. `NOTE THAT ALL TIMES ARE UTC+2.`

	Monday	**Tuesday**	**Wednesday**	**Thursday**	**Friday**
	11:15 conference opening				
11:30 – 12:30	Zdravko Botev	Sandeep Juneja	Mireille Bossy	Heping Wang	Daniel Rudolf
					12:30 conference closing
17:30 – 18:30	Lester Mackey	Larisa Yaroslavtseva	Elena Akhmatskaya	Rob Scheichl	
	18:30 announcement MCM 2023				

Back to overview

zoom 1, Monday, 11:30

Ridge regression and variants: 50 years later

Zdravko Botev
University of New South Wales Sydney, botev@unsw.edu.au

In the age of multi-layer neural networks, we ask the question: Do we really understand the simple linear ridge regression model?

In this back-to-basics talk, I argue that after more than 50 years since its widespread adoption, the simplest Tikhonov regularization method is still a treasure trove of insights waiting to be explored. Along the way, we explain why current popular methods for selecting the regularization parameters do not perform well and how to resolve this issue. In addition, we explain how collinearity amongst the features affects both prediction and inference in precise quantifiable ways. It also turns out that a more sophisticated version of the ridge regression model can be used for model selection, in some cases performing better than much of the current 21 century technology. In view of some of the findings, the talk thus calls for a reassessment of some model selection methods currently in use.

Back to overview

zoom 1, Monday, 17:30

Probabilistic Inference and Learning with Stein's Method

Lester Mackey
Microsoft Research New England and Stanford University, lmackey@stanford.edu

Stein's method is a powerful tool from probability theory for bounding the distance between probability distributions. In this talk, I'll describe how this tool designed to prove central limit theorems can be adapted to assess and improve the quality of practical inference procedures. Along the way, I'll highlight applications to Markov chain Monte Carlo sampler selection, goodness-of-fit testing, variational inference, de novo sampling, post-selection inference, and nonconvex optimization and close with several opportunities for future work.

Back to overview

zoom 1, Tuesday, 11:30

Shift, Scale and Reset Smaller Models to Estimate Larger Ones: Agent-based Simulators in Epidemiology

Sandeep Juneja
Tata Institute of Fundamental Research, Mumbai, juneja@tifr.res.in

Agent-based simulators are a popular epidemiological modelling tool to study the impact of various non-pharmaceutical interventions in managing an evolving pandemic. They provide the flexibility to accurately model a heterogeneous population with time and location varying, person specific interactions. To accurately model detailed behaviour, typically each person is separately modelled. This however, may make computational time prohibitive when the region population is large and when time horizons involved are large. We observe that simply considering a smaller aggregate model and scaling up the output leads to inaccuracies. In this talk we primarily focus on the COVID-19 pandemic and dig deeper into the

underlying probabilistic structure of an associated agent based simulator (ABS) to arrive at modifications that allow smaller models to give accurate statistics for larger models. We exploit the observations that in the initial disease spread phase, the starting infections behave like a branching process. Further, later once enough people have been infected, the infected population closely follow its mean field approximation. We build upon these insights to develop a shifted, scaled and reset version of the simulator that accurately evaluates the ABS's performance using a much smaller model while essentially eliminating the bias that otherwise arises from smaller models.

Back to overview

zoom 1, Tuesday, 17:30

On strong approximation of stochastic differential equations with a discontinuous drift coefficient

Larisa Yaroslavtseva
Universität Passau, larisa.yaroslavtseva@uni-passau.de

The classical assumption in the literature on numerical approximation of stochastic differential equations (SDEs) is global Lipschitz continuity of the coefficients of the equation. However, many SDEs arising in applications fail to have globally Lipschitz continuous coefficients.

In the last decade an intensive study of numerical approximation of SDEs with non-globally Lipschitz continuous coefficients has begun. In particular, strong approximation of SDEs with a drift coefficient that is discontinuous in space has recently gained a lot of interest. Such SDEs arise e.g. in mathematical finance, insurance, neuroscience and stochastic control problems. Classical techniques of error analysis are not applicable to such SDEs and well known convergence results for standard methods do not carry over in general. In this talk I will give an overview of recent results in this area.

Back to overview

zoom 1, Wednesday, 11:30

Some numerical aspects of SDEs modelling particle dynamics in turbulent flow

Mireille Bossy
Inria Sophia Antipolis – Méditerranée, Mireille.Bossy@inria.fr

Particle-laden flows occur in many industrial and environmental systems, such as pollutant dispersion, agglomeration, or deposition phenomena. When dealing with turbulent flow, at play in numerous situations, complex SDE descriptions are coupled with numerical models for the flow, leading to systems of SDEs representing for instance near wall particles dynamics, or colliding systems of Langevin or Brownian type. This talk will present some typical SDE models of this flied and the associated numerical difficulties. These difficulties or numerical constraints are increasing with the improvement of the represented physics, which requires to adapt the time integration schemes and their convergence analysis.

Back to overview

zoom 1, Wednesday, 17:30

Hamiltonian Monte Carlo: Standard Practices Revisited

Elena Akhmatskaya
Basque Center for Applied Mathematics (BCAM) and IKERBASQUE, Basque Foundation for Science,
eakhmatskaya@bcamath.org

With the recently increased interest in probabilistic models, such as Bayesian epidemic models or probabilistic deep learning, the efficiency of an underlying sampler becomes a crucial consideration. A Hamiltonian Monte Carlo (HMC) sampler is one popular choice for models of this kind. We revisit the standard practices of the HMC, and propose several promising alternatives. The topics of our discussion include a formulation of the HMC, numerical integration methods for Hamiltonian dynamics, and the choice of simulation parameters and settings.

This talk is based on joint work with L. Nagar, M. Fernández-Pendás, M. Parga Pazos, F. Puchhammer, T. Radivojević and J. M. Sanz-Serna.

Back to overview

zoom 1, Thursday, 11:30

Tractability in the randomized setting

Heping Wang
Capital Normal University, Beijing, wanghp@cnu.edu.cn

In this talk we discuss two multivariate approximation problems. First we study approximation of multivariate functions from a separable Hilbert space in the randomized setting with the error measured in the weighted L_2 norm. We consider the power of standard information $\Lambda^{\rm std}$ for tractability under the normalized or absolute error criterion and show that it is the same as that of general linear information $\Lambda^{\rm all}$ for all notions of tractability. Specifically, we solve Open Problems 98, 101, 102 and almost solve Open Problem 100 as posed by Novak and Woźniakowski in the book: Tractability of Multivariate Problems, Volume III: Standard Information for Operators, EMS Tracts in Mathematics, Zürich, 2012.

Second we investigate uniform approximation of functions in the tensor product reproducing kernel Hilbert spaces using $\Lambda^{\rm all}$ in the randomized setting and show that it is polynomially tractable under the normalized error criterion. This is contrary to the deterministic case setting, in which uniform approximation problems of tensor product spaces suffer from the curse of dimensionality.

Back to overview

zoom 1, Thursday, 17:30

Multilevel Quasi-Monte Carlo Methods for PDE Eigenvalue Problems

Rob Scheichl
Ruprecht-Karls-Universität Heidelberg, r.scheichl@uni-heidelberg.de

Eigenvalue problems involving partial differential operators appear naturally when modelling physical phenomena, such as the buckling of mechanical structures, the criticality of nuclear reactors or the optical properties of photonic crystals. The model parameters are often only partially known or uncertain. To quantify these uncertainties, stochastic approaches are common. The stochasticity in the coefficients causes the eigenvalues and eigenfunctions to also be stochastic, and so our goal will be to compute statistics, such as expectation or variance, of these eigenvalues and eigenfunctions. Spatially distributed

uncertainty, e.g. in the PDE coefficients or in the geometry of the domain, leads to problems with infinitely many stochastic parameters.

In this talk, we will present a multilevel quasi-Monte Carlo method for approximating the expectation of the minimal eigenvalue of a second-order elliptic eigenvalue problem and provide a rigorous error analysis guaranteeing dimension-independent convergence at a rate of (almost) 1/N in terms of the number of samples N (under mild assumptions on the regularity of the random coefficient and without any cost increase due to finite element mesh refinement). In practice, to approximate this expectation one must: 1) truncate the stochastic dimension; 2) discretise the eigenvalue problem in space (e.g., by finite elements); and 3) apply a quadrature rule to estimate the expected value. The multilevel quadrature method is based on a hierarchy of finite element meshes and truncation dimensions. Furthermore, to make each eigensolve on a given level more efficient, we utilise the two-grid scheme from [Xu & Zhou 1999] to obtain the eigenvalue on the fine mesh from the coarse eigenvalue (and eigenfunction) with a single linear solve. All the approximations are rigorously analysed. Numerical experiments confirm our theoretical results and show excellent scalability.

This is joint work with Alexander Gilbert.

Back to overview

zoom 1, Friday, 11:30

Slice Sampling

Daniel Rudolf
Georg-August-Universität Göttingen, daniel.rudolf@uni-goettingen.de

For approximate sampling of a partially known distribution the slice sampling methodology provides a machinery for the design and simulation of a Markov chain with desirable properties. In the machine learning community it is a frequently used approach, which appears not only their as standard sampling tool. In particular, the elliptical slice sampler attracted in the last decade, as tuning-free and dimension robust algorithm, considerable attention. However, from a theoretical point of view it is not well understood. In general, the theoretical results, which testify qualitatively robust and "good" convergence properties of classical slice sampling methods, are mostly not applicable because of idealized implementation assumptions. Motivated by that the aim of the talk is

1. to provide an introduction into the slice sampling methodology;
2. to discuss different interpretations;
3. to talk about convergence results; as well as
4. to point to open questions.

Back to overview

6.2 Zoom Channel 2

back to table of contents

Clicking on the name of the speaker will lead you to the abstract. NOTE THAT ALL TIMES ARE UTC+2.

	Monday	Tuesday	Wednesday	Thursday	Friday
		Special Session **Non-reversible Markov Chain / Process MC** (part A)	Contributed talks on **MCMC** (part B)		
09:00		Kengo Kamatani	Khai Xiang Au		
09:30		Joris Bierkens	Sanket Agrawal		
10:00		Pierre Monmarché	Mary Llewellyn		
10:30		Chris Sherlock	Neil K. Chada		
	Special Session **Stein's method: Alternatives to MC**	Special Session **Non-reversible Markov Chain / Process MC** (part B)	Contributed talks on **MCMC** (part C)	Special Session **Probabilistic Numerics**	
13:30			J. P. Madrigal Cianci		
14:00	Qiang Liu	Alice Corbella	Luke Kelly	Jon Cockayne	
14:30	Andrew Duncan	Giorgos Vasdekis	Guo-Jhen Wu	Toni Karvonen	
15:30	Zhuo Sun	Murray Pollock	James Foster	Xing Liu	
16:00	Takuo Matsubara	Andi Wang	Katharina Schuh	Jonathan Wenger	
16:30			Jeremie Coullon		
	Special Session **Stein's method: Postprocessing of MC**	Contributed talks on **MCMC** (part A)			
19:00	Marina Riabiz	Saifuddin Syed			
19:30	Onur Teymur	Niloy Biswas			
20:00	Matthew Fisher	Guanyang Wang			
20:30	Jackson Gorham	Jessica Forsyth			

Special Session Stein's method: Alternatives to MC

Back to overview

zoom 2, Monday, 14:00

Uncertainty Quantification with Intractable Models – A Bless from Stein's Method

Qiang Liu
University of Texas at Austin, US, lqiang@cs.utexas.edu

Stein's method is a powerful technique for deriving fundamental theoretical results on approximating and bounding distances between probability measures, such as central limit theorem. Recently, it was found that the key ideas in Stein's method, despite being originally designed as a pure theoretical technique, can be repurposed to provide a basis for developing practical statistics and computational methods for learning and uncertainty quantification using large scale, intractable probabilistic models. This talk will discuss recent applications of Stein's method for uncertainty quantification with intractable probabilistic models.

Back to overview

zoom 2, Monday, 14:30

The Geometry of Stein Variational Gradient Descent and Its Applications

Andrew Duncan
Imperial College London, UK, a.duncan@imperial.ac.uk

Bayesian inference problems require sampling or approximating high-dimensional probability distributions. The focus of this talk is on the recently introduced Stein variational gradient descent methodology, a class of algorithms that rely on iterated steepest descent steps with respect to a reproducing kernel Hilbert space norm. This construction leads to interacting particle systems, the mean-field limit of which is a gradient flow on the space of probability distributions equipped with a certain geometrical structure. This viewpoint is leveraged to shed some light on the convergence properties of the algorithm, addressing the problem of choosing a suitable kernel function. I will also discuss some recent developments in exploiting the underlying geometry of the target probability density to mitigate the curse of dimensionality inherent in Kernel Stein Discrepancy and its SVGD counterpart.

Back to overview

zoom 2, Monday, 15:30

Vector-valued Control Variates

Zhuo Sun
University College London, UK, zhuo.sun.19@ucl.ac.uk

Control variates are post-processing tools for Monte Carlo estimators which can lead to significant variance reduction. This approach usually requires a large number of samples, which can be prohibitive for applications where the problem is computationally expensive. Furthermore, there are many scenarios where we need to compute multiple related integrals simultaneously or sequentially, which can further

exacerbate computational costs. In this paper, we propose vector-valued control variates, an extension of control variates which can be used to reduce the variance of multiple integrals jointly. This allows us to transfer information across integration tasks, and hence reduce the overall requirement for a large number of samples. Our construction is obtained through a generalised Stein identity and the development of novel matrix-valued Stein reproducing kernels.

Back to overview

zoom 2, Monday, 16:00

Robust Generalised Bayesian Inference for Intractable Likelihoods

Takuo Matsubara
Newcastle University, UK, tmatsubara@turing.ac.uk

Generalised Bayesian inference updates prior beliefs using a loss function, rather than a likelihood, and can therefore be used to confer robustness against possible misspecification of the likelihood. Here we consider generalised Bayesian inference with a Stein discrepancy as a loss function, motivated by applications in which the likelihood contains an intractable normalisation constant. In this context, the Stein discrepancy circumvents evaluation of the normalisation constant and produces generalised posteriors that are either closed form or accessible using standard Markov chain Monte Carlo. We provide numerical experiments on a range of intractable distributions, including applications to kernel-based exponential family models and non-Gaussian graphical models.

[1] T. Matsubara, J. Knoblauch, F-X. Briol, and C. J. Oates. *Robust generalised Bayesian inference for intractable likelihoods*. arXiv:2104.07359.

Special Session Stein's method: Postprocessing of MC

Back to overview

zoom 2, Monday, 19:00

Optimal Thinning of MCMC Output

Marina Riabiz
King's College London and Alan Turing Institute, UK, marina.riabiz@kcl.ac.uk

Wilson Ye Chen
University of Sydney, Australia, ye.chen@sydney.edu.au

Jon Cockayne
Alan Turing Institute, UK, jcockayne@turing.ac.uk

Pawel Swietach
Oxford University, UK, pawel.swietach@dpag.ox.ac.uk

Steven Niederer
King's College London, UK, steven.niederer@kcl.ac.uk

Lester Mackey
Microsoft Research, UK, lmackey@stanford.edu

Chris Oates
Newcastle University, UK, chris.oates@newcastle.ac.uk

The use of heuristics to assess the convergence and compress the output of Markov chain Monte Carlo can be sub-optimal in terms of the empirical approximations that are produced. Typically a number of the initial states are attributed to "burn in" and removed [2], whilst the remainder of the chain is "thinned" if compression is also required. In this work we consider the problem of retrospectively selecting a subset of states, of fixed cardinality, from the sample path such that the approximation provided by their empirical distribution is close to optimal. A novel method is proposed, based on greedy minimisation of a kernel Stein discrepancy [4, 3, 1], that is suitable when the gradient of the log-target can be evaluated and an approximation using a small number of states is required. Theoretical results guarantee consistency of the method and its effectiveness is demonstrated in the challenging context of parameter inference for ordinary differential equations. Software is available at http://stein-thinning.org/.

[1] K., Chwialkowski, H., Strathmann, and A., Gretton. *A kernel test of goodness of fit.* Proceedings of the 33rd International Conference on Machine Learning, 2016.

[2] A., Gelman, and D. B., Rubin. *Inference from iterative simulation using multiple sequences.* Statistical science 7, 457–472, 1992.

[3] J., Gorham, and L., Mackey. *Measuring sample quality with kernels.* Proceedings of the 34th International Conference on Machine Learning, 2017.

[4] Q., Liu, and J. D., Lee. *Black-box importance sampling.* Proceedings of the 20th International Conference on Artificial Intelligence and Statistics, 2017.

Back to overview

zoom 2, Monday, 19:30

Optimal Quantisation of Probability Measures Using Maximum Mean Discrepancy

Onur Teymur
Alan Turing Institute, UK, o@teymur.uk

Several researchers have proposed minimisation of maximum mean discrepancy (MMD) as a method to quantise probability measures, i.e., to approximate a probability distribution by a representative point set. We consider sequential algorithms that greedily minimise MMD over a discrete candidate set. This generalises to MMD the approach of [1], who minimise kernel Stein discrepancy for the same task. We then propose a novel non-myopic algorithm that performs discrete optimisation to select more than one point at each iteration and, in order to both improve statistical efficiency and reduce computational cost, we investigate a variant that applies this technique to a mini-batch of the candidate set at each iteration. This is implemented by recasting the algorithm as an integer quadratic programme (IQP), which allows the use of standard optimisation libraries. When the candidate points are sampled from the target, the consistency of these new algorithms—and their mini-batch variants—is established. We present some examples of the use of these algorithms on a range of important computational problems, including optimisation of nodes in Bayesian cubature and the thinning of Markov chain output. Finally, we consider a method of reducing the cost of the per-iteration optimisations by employing semi-definite relaxation for IQPs, following [2]. This approach has promise but is also challenging to implement—we discuss the reasons why.

[1] Riabiz, Chen, Cockayne, Swietach, Niederer, Mackey and Oates (2020) *Optimal Thinning of MCMC Output.* arXiv 2005.03952

[2] Goemans and Williamson (1995) *Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming.* J. Assoc. Computing Machinery 42(6)1115–1145

Back to overview

zoom 2, Monday, 20:00

Measure Transport with Kernel Stein Discrepancy

Matthew Fisher
Newcastle University, m.fisher1@newcastle.ac.uk

Tui Nolan
Cornell University, University of Technology Sydney, tui.h.nolan@alumni.uts.edu.au

Matthew Graham
Newcastle University, Alan Turing Institute, m.graham@turing.ac.uk

Dennis Prangle
Newcastle University, dennis.prangle@newcastle.ac.uk

Chris Oates
Newcastle University, Alan Turing Institute, chris.oates@newcastle.ac.uk

Measure transport underpins several recent algorithms for posterior approximation in the Bayesian context, wherein a transport map is sought to minimise the Kullback–Leibler divergence (KLD) from the posterior to the approximation. The KLD is a strong mode of convergence, requiring absolute continuity of measures and placing restrictions on which transport maps can be permitted. Here we propose to minimise a kernel Stein discrepancy (KSD) instead, requiring only that the set of transport maps is dense in an L^2 sense and demonstrating how this condition can be validated. The consistency of the associated posterior approximation is established and empirical results suggest that KSD is a competitive and more flexible alternative to KLD for measure transport.

Back to overview

zoom 2, Monday, 20:30

Stochastic Stein Discrepancies

Jackson Gorham
Whisper.ai, Inc, jackson@whisper.ai

Anant Raj
MPI for Intelligent Systems, anant.raj@tuebingen.mpg.de

Lester Mackey
Microsoft Research New England, lmackey@microsoft.com

Stein discrepancies (SDs) monitor convergence and non-convergence in approximate inference when exact integration and sampling are intractable. However, the computation of a Stein discrepancy can be prohibitive if the Stein operator - often a sum over likelihood terms or potentials - is expensive to evaluate. To address this deficiency, we show that stochastic Stein discrepancies (SSDs) based on subsampled approximations of the Stein operator inherit the convergence control properties of standard SDs with probability 1. Along the way, we establish the convergence of Stein variational gradient descent (SVGD) on unbounded domains, resolving an open question of Liu (2017). In our experiments with biased Markov chain Monte Carlo (MCMC) hyperparameter tuning, approximate MCMC sampler selection, and stochastic SVGD, SSDs deliver comparable inferences to standard SDs with orders of magnitude fewer likelihood evaluations.

[1] J. Gorham, A. Raj, and L. Mackey. Stochastic Stein Discrepancies. *Advances in Neural Information Processing Systems*. Pages 17931–17942, 2020.

[2] Q. Liu. Stein variational gradient descent as gradient flow. *Advances in Neural Information Processing Systems*. Pages 3115–3123, 2017.

Special Session Non-reversible Markov Chain / Process MC (part A)

Back to overview

zoom 2, Tuesday, 09:00

High-dimensional scaling limit of some piecewise deterministic Markov processes

Joris Bierkens
Delft University of Technology, Netherlands, joris.bierkens@tudelft.nl

Kengo Kamatani
IMS, Japan, kamatani@ism.ac.jp

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

Recently, piecewise deterministic Markov processes have gained interest in the Monte Carlo community in the context of scalable Monte Carlo integration methods. In this talk, we will discuss high-dimensional scaling limits for some piecewise deterministic Markov processes. We will describe these results using multiscale analysis, which is a useful technique for this purpose. We will also highlight two types of scaling limits corresponding to the tangential direction and the linear direction to the log-density contour.

[1] J. Bierkens, K. Kamatani, and G. O. Roberts. *High-dimensional scaling limits of piecewise deterministic sampling algorithms*. arXiv preprint arXiv:1807.11358.

Back to overview

zoom 2, Tuesday, 09:30

Spectral theory and asymptotic variance of piecewise deterministic samplers

Joris Bierkens
Technische Universiteit Delft, joris.bierkens@tudelft.nl

Sjoerd Verduyn Lunel
Universiteit Utrecht, s.m.verduynlunel@uu.nl

In recent years piecewise deterministic Markov processes (PDMPs) have emerged as a promising alternative to classical MCMC algorithms.

In this talk PDMP based algorithms such as the Zig-Zag Sampler and the Bouncy Particle Sampler will be introduced and recent progress in our understanding of the underlying processes will be presented. My aim is to zoom in on two types of convergence properties, as described by (i) the asymptotic variance and (ii) the spectral properties of the process. Interestingly, these provide an understanding which may appear self-contradictory at a first glance.

[1] Joris Bierkens and Sjoerd M. Verduyn Lunel. Spectral analysis of the zigzag process. *to appear in Annales de l'Institut Henri Poincaré (B)*, 2021.

Back to overview

zoom 2, Tuesday, 10:00

Velocity-jump processes for sampling: from bounces to Hamiltonian dynamics

Pierre Monmarché
Sorbonne Université, pierre.monmarche@sorbonne-universite.fr

Mathias Rousset
INRIA Rennes, mathias.rousset@inria.fr

Pierre-André Zitt
Université Paris Est-Marne la Vallée, pierre-andre.zitt@u-pem.fr

We present a systematic general construction of piecewise deterministic kinetic processes with a given target equilibrium. This leads to a family of processes depending on a parameter ε: for $\varepsilon = +\infty$, we recover the so-called bounce mechanism, i.e. the now standard bouncy particle and zig-zag samplers. As ε goes to zero, we prove that the process converge in law to the Hamiltonian dynamics (or a variation like the kinetic Langevin process when a dissipative part is added). An L^2 quantitative long-time convergence, uniform in ε, is established.

[1] P. Monmarché, M. Rousset, and P.-A. Zitt. *Exact targeting of Gibbs distributions using velocity-jump processes.* arXiv e-prints, arXiv:2008.09360, 2020.

Back to overview

zoom 2, Tuesday, 10:30

The Apogee to Apogee Path Sampler

Chris Sherlock
Lancaster University, c.sherlock@lancaster.ac.uk

Szymon Urbas
Lancaster University, s.urbas@lancaster.ac.uk

Matthew Ludkin
Lancaster University, -

Hamiltonian Monte Carlo (HMC) is often the method of choice when performing inference on complex high-dimensional Bayesian posterior distributions; however, it is notoriously difficult to tune. In particular, slight changes in the choice of integration time can make the difference between an optimally efficient algorithm and a poor one. We define an *apogee* of a Hamiltonian path as a point in the path where the component of momentum in the direction of the gradient of the potential changes from positive to negative; from this we introduce the *Apogee to Apogee Path Sampler* which uses the same leapfrog dynamics as HMC but eschews the integration time parameter in favour of the choice of a number of apogees, and allows for a variety of proposal mechanisms that choose from the whole Hamiltonian path. The resulting algorithm is competitive with optimally tuned HMC, and is more robust to the tuning choice of "number of apogees" than HMC is to choice of integration time. Furthermore, the number of apogees relates directly to intrinsic properties of the posterior, allowing for simple tuning guidelines in terms of a few scalar summaries of the distribution. Finally, the algorithm requires no self-recursions and for certain useful classes of proposal the storage cost through an iteration is $\mathcal{O}(1)$.

Special Session Non-reversible Markov Chain / Process MC (part B)

Back to overview

zoom 2, Tuesday, 14:00

Automatic Zig-Zag Sampling

Alice Corbella
University of Warwick, alice.corbella@warwick.ac.uk

Simon Spencer
University of Warwick, s.e.f.spencer@warwick.ac.uk

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

Recent research showed that Piecewise Deterministic Markov Processes (PDMP) may be exploited to design efficient MCMC algorithms [1]. The Zig-Zag sampler is an example of this: it is based on the simulation of a PDMP whose switching rate $\lambda(t)$ is governed by the derivative of a (log) target density.

While many properties of this sampler have been derived [2, 3], less has been done to explore the applicability of the Zig-Zag sampler to solve Bayesian inference problems. One of the main obstacles towards a wider use of this method by the community is the computation of the derivative of the log-density in the rate $\lambda(t)$. This can be particularly challenging when dealing with likelihoods containing recursive relationships, for which manual differentiation is time-consuming.

To expand the applicability of the Zig-Zag sampler, we incorporate Automatic Differentiation (AD) tools in the Zig-Zag algorithm, to facilitate the computation of $\lambda(t)$ from the functional form of the log-target density. Moreover, to allow the simulation of a PDMP via Poisson thinning, we use univariate optimization routines to find a local upper bound for the bounding rate.

In this talk we present our Automatic Zig-Zag sampler; we discuss the challenges that arise with the simulation via thinning and the need of a new tuning parameter; and we comment on efficiencies and bottlenecks of AD for Zig-Zag.

[1] Fearnhead, P., Bierkens, J., Pollock, M. and Roberts, G.O., 2018. Piecewise deterministic Markov processes for continuous-time Monte Carlo. *Statistical Science*, 33(3), pp.386-412.

[2] Bierkens, J., Fearnhead, P. and Roberts, G., 2019. The zig-zag process and super-efficient sampling for Bayesian analysis of big data. *The Annals of Statistics*, 47(3), pp.1288-1320.

[3] Bierkens, J., Roberts, G.O. and Zitt, P.A., 2019. Ergodicity of the zigzag process. *Annals of Applied Probability*, 29(4), pp.2266-2301.

Back to overview

zoom 2, Tuesday, 14:30

Speed Up Zig-Zag

Giorgos Vasdekis
University of Warwick, Giorgos.Vasdekis.1@warwick.ac.uk

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

Zig-Zag is a Piecewise Deterministic Markov Process, efficiently used for simulation in a Markov Chain Monte Carlo setting. However, it fails to be exponentially ergodic on heavy tailed target distributions. In this talk we introduce an extension of the Zig-Zag process by allowing the process to move with a non-constant speed function s, depending on the current state of the process. We call this process Speed Up Zig-Zag (SUZZ). We provide conditions that guarantee stability properties for the SUZZ process, including non-explosivity, exponential ergodicity on heavy tailed targets and central limit theorem. Interestingly, we find that using speed functions that induce explosive deterministic dynamics may lead to stable algorithms that can even mix faster. We support our findings with simulation results and time permitting, we will discuss the choice of an efficient speed function by providing an efficiency criterion for the one-dimensional process.

[1] G. Vasdekis and G. O. Roberts *Speed Up Zig-Zag.* arxiv preprint, 2021. https://arxiv.org/abs/2103.16620

[2] J. Bierkens and P. Fearnhead and G. O. Roberts. *The Zig-Zag process and super-efficient sampling for Bayesian analysis of big data.* Ann. Statist., 47(3):1288-1320, 06 2019. https://doi.org/10.1214/18-AOS1715

[3] J. Bierkens and G. O. Roberts. and P. A. Zitt *Ergodicity of the zigzag process.* Ann. Appl. Probab., 29(4):2266-2301, 08 2019. https://doi.org/10.1214/18-AAP1453

Back to overview

zoom 2, Tuesday, 15:30

The Restore Process - Practical CFTP by enriching Markov processes

Murray Pollock
Newcastle University, murray.pollock@newcastle.ac.uk

Andi Q. Wang
University of Bristol, andi.wang@bristol.ac.uk

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

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

We develop a new class of Markov processes comprising local dynamics governed by a fixed Markov process, which are enriched with regenerations from a fixed distribution at a state-dependent rate. We give conditions under which such processes possess a given target distribution as their invariant measures, thus making them amenable for use within Monte Carlo methodologies. Enrichment imparts a number of desirable theoretical and methodological properties, which includes straightforward conditions for the process to be uniformly ergodic and possess a coupling from the past construction that enables exact sampling from the invariant distribution. Joint work with David Steinsaltz / Gareth Roberts / Andi Wang.

[1] A.Q. Wang and M. Pollock and G.O. Roberts and D. Steinsaltz *Regeneration-enriched Markov processes with application to Monte Carlo.* Annals of Applied Probability, 31(2): 703–735, 2021

Back to overview

zoom 2, Tuesday, 16:00

Quasi-stationary Monte Carlo methods via stochastic approximation

Andi Q. Wang
University of Bristol, andi.wang@bristol.ac.uk

Murray Pollock
Newcastle University, murray.pollock@newcastle.ac.uk

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

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

Quasi-stationary distributions describe the long-term behaviour of an absorbed stochastic process, conditioned on non-absorption. Surprisingly, such objects have very recently been used to construct scalable MCMC methods to perform exact Bayesian inference on tall datasets [2, 1, 3]. In my talk I will introduce this methodology, and describe an alternative methodological approach based on stochastic approximation techniques to simulating quasi-stationary distributions [4]. This approach is also amenable to subsampling for tall data, and enables further extensions which simplify the procedure considerably.

[1] Kumar, D. On a Quasi-Stationary Approach to Bayesian Computation, with Application to Tall Data. PhD thesis, University of Warwick, 2019.

[2] Pollock, M., Fearnhead, P., Johansen, A. M., Roberts, G. O. Quasi-stationary Monte Carlo and the ScaLE algorithm, with discussion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 82(5), 1167-1221* , 2020.

[3] Wang, A. Q., Kolb, M., Roberts, G. O., Steinsaltz, D. Theoretical properties of quasi-stationary Monte Carlo methods. *The Annals of Applied Probability, 29(1), 434-457,* 2019.

[4] Wang, A. Q., Roberts, G. O., Steinsaltz, D. An approximation scheme for quasi-stationary distributions of killed diffusions. *Stochastic Processes and Their Applications, 130(5), 3193-3219*, 2020.

Contributed talks on MCMC (part A)

Back to overview

zoom 2, Tuesday, 19:00

Parallel Tempering on Optimized Paths

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

Vittorio Romaniello
University of British Columbia, vittorio.romaniello@stat.ubc.ca>

Trevor Campbell
University of British Columbia, trevor@stat.ubc.ca

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

Parallel tempering (PT) is a class of Markov chain Monte Carlo algorithms that constructs a path of distributions annealing between a tractable reference and an intractable target, and then interchanges states along the path to improve mixing in the target. The performance of PT depends on how quickly a sample from the reference distribution makes its way to the target, which in turn depends on the particular path of annealing distributions. However, past work on PT has used only simple paths constructed from convex combinations of the reference and target log-densities. In this talk we will show that this path performs poorly in the common setting where the reference and target are nearly mutually singular. To address this issue, we will present an extension of the PT framework to general families of paths, formulate the choice of path as an optimization problem that admits tractable gradient estimates, and present a flexible new family of spline interpolation paths for use in practice. Theoretical and empirical results will demonstrate that the proposed methodology breaks previously-established upper performance limits for traditional paths.

[1] S. Syed, R. Romaniello, T. Campbell and A. Bouchard-Côté. *Parallel Tempering on Optimized Paths.* arXiv preprint: 2102.07720.

Back to overview

zoom 2, Tuesday, 19:30

Bounding Wasserstein distance with perturbed couplings

Niloy Biswas
Harvard University, niloy_biswas@g.harvard.edu

Lester Mackey
Microsoft Research New England and Stanford University, lmackey@microsoft.com

We propose estimators based on couplings of Markov chains to assess the quality of asymptotically biased sampling methods such as approximate Markov chain Monte Carlo (MCMC). The estimators give empirical upper bounds of the Wasserstein's distance between the limiting distribution of the asymptotically biased sampling method and the original target distribution of interest. We establish theoretical guarantees of our upper bounds, and show that our estimators can remain effective in high dimensions. We apply our measure of sample quality to stochastic gradient MCMC, variational Bayes and Laplace approximation for tall data; to approximate MCMC for high-dimensional linear regression, and to approximate MCMC for high-dimensional logistic regression.

Back to overview

zoom 2, Tuesday, 20:00

Transition kernel couplings of the Metropolis-Hastings algorithm

John O'Leary
Harvard University, joleary@g.harvard.edu

Guanyang Wang
Rutgers University, guanyang.wang@rutgers.edu

Couplings play a central role in the analysis of Markov chain convergence to stationarity and in the construction of novel Markov chain Monte Carlo diagnostics, estimators, and variance reduction techniques. The quality of the resulting bounds or methods typically depends on how quickly the coupling induces meeting between chains, a property sometimes referred to as its efficiency. The design of efficient Markovian couplings remains a difficult open question, especially for discrete time processes. In pursuit of this goal, in this paper we fully characterize the couplings of the Metropolis–Hastings (MH) transition kernel, providing necessary and sufficient conditions in terms of the underlying proposal and acceptance distributions. We apply these results to characterize the set of maximal couplings of the MH kernel, resolving open questions posed in O'Leary et.al. [2] on the structure and properties of these couplings. These results represent an advance in the understanding of the MH kernel and a step toward the formulation of efficient couplings for this popular family of algorithms. We will also discuss several implementable maximal kernel couplings if time permits.

[1] J.O'Leary and G.Wang *Transition kernel couplings of the Metropolis-Hastings algorithm.* https://arxiv.org/abs/2102.00366.

[2] J.O'Leary, G.Wang and P.E.Jacob *Maximal couplings of the Metropolis-Hastings algorithm.* Artificial Intelligence and Statistics, 2021.

Back to overview

zoom 2, Tuesday, 20:30

A computational approach to link cell history with cell fate at the single level

Jessica Forsyth
Faculty of Biology Medicine and Health, University of Manchester, United Kingdom,
jessica.forsyth@postgrad.manchester.ac.uk

Simon Cotter
School of Mathematics, University of Manchester, United Kingdom, simon.cotter@manchester.ac.uk

Berenika Plusa
Faculty of Biology Medicine and Health, University of Manchester, United Kingdom,
berenika.plusa@manchester.ac.uk

During early development of the mammalian embryo, three cell types are specified. How this process is coordinated and controlled remains unknown. We are working to identify key spatio-temporal cues that regulate and control the acquisition of cell fate and the subsequent organisation of cells. In order to study and gain a better understanding of potential spatio-temporal cues, we use different imaging techniques. However the matching of cells across imaging techniques is non-trivial.

We use an MCMC approach to explore a parameter space describing an Affine transformation and non-linear deformation of points between images whilst simultaneously sampling on a discrete permutation vector describing the matching of cells. In addition, we sample on parameters that describe the weighting of specific data points within the dataset to account for the presence of spurious data points due to cell death or division. Due to the complexity of the parameter space we also employ tempering to ensure full exploration of the state space and minimise the risk of solution trapping. This approach provides us with a true probabilistic matching of cells across images.

By using this approach we are able to link cell history with eventual cell fate and start to identify key steps in early mammalian development that have previously not been accessible using classical approaches.

Contributed talks on MCMC (part B)

Back to overview

zoom 2, Wednesday, 09:00

Tuning the refreshment parameter κ of the Discrete Bouncy Particle Sampler

Khai Xiang Au
National University of Singapore, khai@u.nus.edu

Alexandre Thiery
National University of Singapore, a.h.thiery@nus.edu.sg

The Hamiltonian Monte Carlo (HMC) method is well-known for its ability to generate distant proposals and avoid random-walk behaviour. Its sampling efficiency however is highly sensitive to the choice of the number of leapfrog integration steps. Though the No-U-Turn Sampler (NUTS) automates the tuning of this parameter, in practice it is hard to implement and importantly its recursive nature makes sampling parallel chains challenging. In the context of the Discrete Bouncy Particle Sampler (DBPS), the refreshment parameter κ plays an analogous role to that of the number of leapfrog integration steps in

HMC. In this work, we present several simple tuning strategies for κ. Leveraging parallelisation wherever possible, we demonstrate the competitive performance of a Hamiltonian variant of the DBPS against other DBPS variants and expertly-tuned HMC methods.

[1] M. D. Hoffman and A. Gelman. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research,* 15(1):1593–1623, 2014.

[2] R. M. Neal. MCMC Using Hamiltonian Dynamics. *Handbook of Monte Carlo methods.* John Wiley & Sons, 2011.

[3] C. Sherlock and A. H. Thiery. A discrete bouncy particle sampler. *arXiv e-prints.* 2021.

Back to overview

zoom 2, Wednesday, 09:30

More magic numbers: Optimal scaling for Barker's and other acceptances in MCMC

Sanket Agrawal
Indian Institute of Technology Kanpur, agarwalsanket95@gmail.com

Dootika Vats
Indian Institute of Technology Kanpur, dootika@iitk.ac.in

Krzysztof Łatuszyński
University of Warwick, K.G.Latuszynski@warwick.ac.uk

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

The problem of optimally scaling the proposal distribution in a Markov chain Monte Carlo algorithm is critical to the quality of the generated samples. Much work has gone into obtaining such results for various Metropolis-Hastings (MH) algorithms. Recently, acceptance probabilities other than MH are being employed in problems with intractable target distributions. There is little resource available on tuning the proposal distributions for this situation. We obtain optimal scaling results for a general class of acceptance functions, which includes Barker's acceptance function and the Lazy-MH algorithms. For this class, we present a direct expression for obtaining the asymptotically optimal variance of the proposal distribution and the asymptotically optimal acceptance rate for the corresponding algorithms. In particular, optimal values for Barker's algorithm are derived and are found to be significantly different from that obtained for MH algorithms.

Back to overview

zoom 2, Wednesday, 10:00

Discretising a Continuous World: Accelerated Inference for State-Space Models via Hidden Markov Models

Mary Llewellyn
University of Edinburgh, mary.llewellyn@ed.ac.uk

Ruth King
University of Edinburgh, Ruth.King@ed.ac.uk

Gordon Ross
University of Edinburgh, Gordon.Ross@ed.ac.uk

Victor Elvira
University of Edinburgh, victor.elvira@ed.ac.uk

State-space models (SSMs) are often used to model time series data where the observations depend on an unobserved latent process. However, Bayesian inference on the process parameters of a SSM is challenging when the likelihood of the data given these parameters is not available in closed form. This has led to the development of particle MCMC methods which combine sequential Monte Carlo approximations with corrective Metropolis-Hastings steps such that exact samples from the posterior for the parameters can be obtained. Such methods can be inefficient when both the states and parameters need to be estimated because of the sample impoverishment common to sequential Monte Carlo methods. In this talk, we present an alternative framework for state-space inference which does not rely on sequential Monte Carlo and hence avoids sample impoverishment. Our approach discretises the state-space using a simple deterministic approximation which essentially turns the state-space model into a hidden Markov model (HMM). This HMM is then used as a proposal distribution for the states, with Metropolis-Hastings used to correct for the approximation error so that the samples target the correct joint posterior through exact sampling on the state space. We illustrate our approach with several examples.

Back to overview

zoom 2, Wednesday, 10:30

Unbiased inference for discretely observed hidden Markov model diffusions

Neil K. Chada
King Abdullah University of Science and Technology, neilchada123@gmail.com

Jordan Franks
University of Newcastle, jordan.j.franks@jyu.fi

Ajay Jasra
King Abdullah University of Science and Technology, ajay.jasra@kaust.edu.sa

Kody J. H. Law
University of Manchester, kodylaw@gmail.com

Matti Vihola
University of Jyvaskyla, matti.s.vihola@jyu.fi

We develop a Bayesian inference method for diffusions observed discretely and with noise, which is free of discretisation bias. Unlike existing unbiased inference methods, our method does not rely on exact simulation techniques. Instead, our method uses standard time-discretised approximations of diffusions, such as the Euler-Maruyama scheme. Our approach is based on particle marginal Metropolis-Hastings, a particle filter, randomised multilevel Monte Carlo, and importance sampling type correction of approximate Markov chain Monte Carlo. The resulting estimator leads to inference without a bias from the time-discretisation as the number of Markov chain iterations increases. We give convergence results and recommend allocations for algorithm inputs. Our method admits a straightforward parallelisation, and can be computationally efficient. The user-friendly approach is illustrated on three examples, where the underlying diffusion is an Ornstein-Uhlenbeck process, a geometric Brownian motion, and a $2d$ non-reversible Langevin equation.

[1] N. K. Chada, J. Franks, A. Jasra, K. J. H. Law and M. Vihola. Unbiased estimation of discretely observed hidden Markov models. *Accepted by SIAM/ASA Journal on Uncertainty Quantification*, (2021).

Contributed talks on MCMC (part C)

Back to overview

zoom 2, Wednesday, 13:30

Multi-level Markov chain Monte Carlo with maximally coupled proposals

Juan Pablo Madrigal Cianci
EPFL, juan.madrigalcianci@epfl.ch

Fabio Nobile
EPFL, juan.madrigalcianci@epfl.ch

In this work we present a novel class of Multi-Level Markov chain Monte Carlo (ML-MCMC) algorithms based on maximally coupled [1] proposals, and apply them in the context of Bayesian inverse problems. In this context, the likelihood function involves a complex differential model, which is then approximated on a sequence of increasingly accurate discretizations. The key point of this algorithm is to construct highly coupled Markov chains together with the standard Multi-level Monte Carlo argument to obtain a better cost-tolerance complexity than a single level MCMC algorithm [2]. Our approach generates these highly coupled chains by sampling from a maximal coupling of the proposals for each marginal Markov chain. By doing this, we are allowed to create novel ML-MCMC methods for which, contrary to previously used models, the proposals at each iteration can depend on the current state of this chain, while at the same time, creating chains that are highly correlated. The presented method is tested on an array of academic examples which evidence how our extended ML-MCMC method is robust when targeting some *pathological* posteriors, for which some of the previously proposed ML-MCMC algorithms fail.

[1] Torgny Lindvall. *Lectures on the coupling method*. Courier Corporation, 2002.

[2] Tim J Dodwell, Chris Ketelsen, Robert Scheichl, and Aretha L Teckentrup. A hierarchical multilevel Markov Chain Monte Carlo algorithm with applications to uncertainty quantification in subsurface flow. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):1075–1108, 2015.

Back to overview

zoom 2, Wednesday, 14:00

COUPLED MCMC FOR PHYLOGENETIC INFERENCE

Luke Kelly
CEREMADE, Université Paris Dauphine, kelly@ceremade.dauphine.fr

Robin Ryder
CEREMADE, Université Paris Dauphine, ryder@ceremade.dauphine.fr

Many modern phylogenetic methods specify a generative model and take a Bayesian approach to inference. Phylogenetic posterior distributions are typically intractable functions of the tree and model parameters, highly multimodal, and prone to misspecification [1, 2]. Markov chain Monte Carlo is the primary tool for inference — although there have been recent attempts at using sequential Monte Carlo, piecewise deterministic Markov processes and variational approximations — and in practice one resorts to running either extremely long marginal chains or coupled tempered chains. In any case, we lack methods to properly quantify convergence or mixing of Markov schemes on the space of trees, so it is difficult to separate modelling and fitting errors.

We extend recent work using coupled Markov chains to construct unbiased estimators [3] and diagnose convergence [4] to the space of phylogenetic trees. The extension is not straightforward due to the complexities of working in tree space and because we couple existing marginal kernels which only operate on a small subset of the state. We illustrate our convergence diagnostics and unbiased estimators on a variety of problems and discuss their usefulness compared to other methods.

[1] C. Whidden and F.A. Matsen IV. Quantifying MCMC Exploration of Phylogenetic Tree Space. *Syst. Biol.*, 64(3):472–491, 01 2015.

[2] C. Whidden, B.C. Claywell, T. Fisher, A.F. Magee, M. Fourment, and F.A. Matsen IV. Systematic Exploration of the High Likelihood Set of Phylogenetic Tree Topologies. *Syst. Biol.*, 69(2):280–293, 08 2019.

[3] P.E. Jacob, J. O'Leary, and Y.F. Atchadé. Unbiased Markov chain Monte Carlo methods with couplings. *J. Roy. Statist. Soc. B*, 2020.

[4] N. Biswas, P.E. Jacob, and P. Vanetti. Estimating convergence of Markov chains with L-lag couplings. In *NeurIPS*, pages 7389–7399, 2019.

Back to overview

zoom 2, Wednesday, 14:30

Analysis and optimization of certain parallel Monte Carlo methods in the low temperature limit

Paul Dupuis
Brown University, paul_dupuis@brown.edu

Guo-Jhen Wu
KTH Royal Institute of Technology, gjwu@kth.se

Metastability is a formidable challenge to Markov chain Monte Carlo methods. In this talk we present methods for algorithm design to meet this challenge. The design problem we consider is temperature selection for the infinite swapping scheme, which is the limit of the widely used parallel tempering scheme obtained when the swap rate tends to infinity. We use a recently developed tool for the large deviation properties of the empirical measure of a metastable small noise diffusion to transform the variance reduction problem into an explicit graph optimization problem. The nodes in the graph optimization problem correspond to metastable states of the noiseless dynamics. Our first analysis of the optimization problem is in the setting of a double well model, and it shows that the optimal selection of temperature ratios is a geometric sequence except possibly the highest temperature. In the same setting we identify two different sources of variance reduction, and show how their competition determines the optimal highest temperature. In the general multi-well setting we prove that the same geometric sequence of temperature ratios as in the two-well case is always nearly optimal, with a performance gap that decays geometrically in the number of temperatures. Moreover, this optimal placement of temperatures is explicit and independent of the particular functional being integrated or (with mild restrictions) on the potential.

[1] P. Dupuis and G.-J. Wu. *Analysis and optimization of certain parallel Monte Carlo methods in the low temperature limit.* arXiv, 2020.

Back to overview

zoom 2, Wednesday, 15:30

A high order method for underdamped Langevin MCMC

James Foster
University of Oxford, james.foster@maths.ox.ac.uk

Terry Lyons
University of Oxford, terry.lyons@maths.ox.ac.uk

Harald Oberhauser
University of Oxford, harald.oberhauser@maths.ox.ac.uk

Langevin Monte Carlo (LMC) is a popular Markov chain Monte Carlo sampling method, obtained from an Euler discretization of Langevin dynamics. More recently, non-reversable LMC algorithms based on the underdamped Langevin dynamics (ULD) have been studied.

In this talk, I will present a new approach to the numerical approximation of ULD. Our strategy is to first reduce the Langevin SDE to a comparable ODE, before then applying an appropriate ODE solver. For strongly log-concave probability distributions, we show that this ODE approximation achieves a 2-Wasserstein error of ε in $\mathcal{O}(\sqrt{d}/\varepsilon^{\frac{1}{3}})$ steps when the first three derivatives of the target log-density are Lipschitz continuous. Moreover, by discretizing this ODE using a third order Runge-Kutta method, we obtain a practical MCMC method that uses just two additional gradient evaluations per step. In our experiment, where the target comes from a logistic regression, this method shows faster convergence compared to other unadjusted Langevin MCMC algorithms.

[1] J. Foster, T. Lyons and H. Oberhauser *The shifted ODE method for underdamped Langevin MCMC.* https://arxiv.org/abs/2101.03446, 2021.

Back to overview

zoom 2, Wednesday, 16:00

Convergence of unadjusted Hamiltonian Monte Carlo for mean-field models

Katharina Schuh
Universität Bonn, Katharina.Schuh@uni-bonn.de

We consider the unadjusted Hamiltonian Monte Carlo algorithm applied to highdimensional probability distributions of mean-field type. We evolve dimension-free convergence and discretization error bounds. These bounds require the discretization step to be sufficiently small, but do not require strong convexity of either the unary or pairwise potential terms present in the mean-field model. To handle high dimensionality, we use a particlewise coupling that is contractive in a complementary particlewise metric. This talk is based on joint work with Nawaf Bou-Rabee.

Back to overview

zoom 2, Wednesday, 16:30

Ensemble sampler for infinite-dimensional inverse problems

Jeremie Coullon
Lancaster University, jeremie.coullon@gmail.com

Robert J. Webber
New York University, w2515@nyu.edu

We introduce a new Markov chain Monte Carlo (MCMC) sampler for infinite-dimensional inverse problems. Our new sampler is based on the affine invariant ensemble sampler, which uses interacting walkers to adapt to the covariance structure of the target distribution. We extend this ensemble sampler for the first time to infinite-dimensional function spaces, yielding a highly efficient gradient-free MCMC algorithm. Because our new ensemble sampler does not require gradients or posterior covariance estimates, it is simple to implement and broadly applicable.

Special Session Probabilistic Numerical Methods

Back to overview

zoom 2, Thursday, 14:00

Calibration of Probabilistic Numerical Methods for Linear Systems

Jon Cockayne
The Alan Turing Institute, jcockayne@turing.ac.uk

A fundamental task in numerical computation is the solution of large linear systems. Naturally therefore, probabilistic numerical methods for the solution of such systems have received a great deal of attention. In this talk we will introduce Bayesian probabilistic numerical methods for linear systems. A significant challenge in construction of such methods is that they tend to be poorly *calibrated*, that is, that the width of the posterior is not commensurate with the error incurred. The challenges in constructing well-calibrated probabilistic numerical methods for linear systems will be discussed, and recent advances which seek to address this will be presented.

Back to overview

zoom 2, Thursday, 14:30

Reliability of Gaussian Process Approximation

Toni Karvonen
The Alan Turing Institute, tkarvonen@turing.ac.uk

Many probabilistic numerical methods, including standard forms of Bayesian quadrature, are based on the use of Gaussian process (GP) priors. This talk discusses theoretical results about when—and to what extent—such a GP-based method can be expected to be *reliable* in the sense that its conditional standard deviation is commensurate with the true approximation error in a setting where the data-generating function is fixed and deterministic. Specifically, we show that simple *maximum likelihood estimation* of the scaling parameter of the GP prior guarantees at most "slow" asymptotic overconfidence in that, under certain assumptions, the standard deviation decays with at most rate $\mathcal{O}(N^{-1/2})$ (up to

logarithmic factors) faster than the true approximation error, N being the number of distinct data points. The results are derived by exploiting the equivalence of approximation with a GP and worst-case optimal approximation in a reproducing kernel Hilbert space. We also discuss how using *cross-validation* to select the scaling parameter may lead to better results and how the results relate to sample path properties of GPs. The talk is based on the recent articles [1, 2] and an ongoing collaboration with M. Naslidnyk, M. Mahsereci and M. Kanagawa.

[1] T. Karvonen (2021). Small sample spaces for Gaussian processes. *arXiv:2103.03169*.

[2] T. Karvonen, G. Wynne, F. Tronarp, C. J. Oates & S. Särkkä (2020). Maximum likelihood estimation and uncertainty quantification for Gaussian process approximation of deterministic functions. *SIAM/ASA Journal on Uncertainty Quantification*, 8(3):926–958.

Back to overview

zoom 2, Thursday, 15:30

Bayesian Probabilistic Numerical Integration with Tree-Based Models

Harrison Zhu
Imperial College London, hbz15@ic.ac.uk

Xing Liu
Imperial College London, xl6116@ic.ac.uk

Ruya Kang
Brown University, ruya_kang@brown.edu

Zhichao Shen
University of Oxford, zhichao.shen@new.ox.ac.uk

Seth Flaxman
Imperial College London, s.flaxman@imperial.ac.uk

François-Xavier Briol
University College London, f.briol@ucl.ac.uk

Bayesian quadrature (BQ) is a method for solving numerical integration problems in a Bayesian manner, which allows users to quantify their uncertainty about the solution. The standard approach to BQ is based on a Gaussian process (GP) approximation of the integrand. As a result, BQ is inherently limited to cases where GP approximations can be done in an efficient manner, thus often prohibiting very high-dimensional or non-smooth target functions. This paper proposes to tackle this issue with a new Bayesian numerical integration algorithm based on Bayesian Additive Regression Trees (BART) priors, which we call BART-Int. BART priors are easy to tune and well-suited for discontinuous functions. We demonstrate that they also lend themselves naturally to a sequential design setting and that explicit convergence rates can be obtained in a variety of settings. The advantages and disadvantages of this new methodology are highlighted on a set of benchmark tests including the Genz functions, on a rare-event simulation problem and on a Bayesian survey design problem.

Back to overview

zoom 2, Thursday, 16:00

Probabilistic Numerical Methods – An Algorithmic Perspective.

Jonathan Wenger
Universität Tübingen, jonathan.wenger@uni-tuebingen.de

Probabilistic numerical methods (PNMs) [1, 2, 3] aim to explicitly represent uncertainty resulting from limited computational resources and stochastic input in numerical computation. This emerging paradigm treats a numerical problem as one of statistical inference instead. It holds the promise to improve upon classic methods in areas where adaptivity is needed, related problems are solved repeatedly, stochasticity plays a role and solutions are not unique. However, one key barrier to the wider adoption of such methods is the lack of probabilistic numerics software.

In this talk, we demonstrate how the theoretical formulation of PNMs can be translated into an algorithmic framework and introduce ProbNum – a software stack for probabilistic numerics. It enables the composition of problem-specific PNMs and paves the way for consistent propagation of uncertainty through the entirety of a computational pipeline.

[1] P. Hennig, M. A. Osborne, and M. Girolami. Probabilistic numerics and uncertainty in computations. *Proc. A.*, 471(2179):20150142, 17, 2015.

[2] C. J. Oates and T. J. Sullivan. A modern retrospective on probabilistic numerics. *Stat. Comput.*, 29(6):1335–1351, 2019.

[3] J. Cockayne, C. J. Oates, T. J. Sullivan, and M. Girolami. Bayesian probabilistic numerical methods. *SIAM Rev.*, 61(4):756–789, 2019a.

6.3 Zoom Channel 3

back to table of contents

Clicking on the name of the speaker will lead you to the abstract. NOTE THAT ALL TIMES ARE UTC+2.

	Monday	Tuesday	Wednesday	Thursday	Friday
		Special Session **Numerics for SDEs with boundary issues (part B)**	Special Session **MC methods for discontinuous functions**	Special Session **Theory and Application of Particle Systems**	Contributed talks on **Regression and Learning**
09:00			Mike Giles	Boualem Djehiche	
09:30		Juan A. Acebrón	Abdul-Lateef Haji-Ali	Pierre Del Moral	Vesa Kaarnioja
10:00		Mike Giles	Simon Weissmann	Dan Crisan	Christian Weiss
10:30		Madalina Deaconu	Per Pettersson	Raúl Tempone	Daniel Roth
	Special Session **Numerics for SDEs with boundary issues (part A)**	Special Session **Variance Reduction: Rare events**	Special Session **Variance Reduction: Pure Jumps and Stochastic Reaction Networks**	Special Session **Variance Reduction: Computational Finance**	
13:30			Sophia Wiechert		
14:00	Francisco Bernal	Bruno Tuffin	Michael Backenköhler	Zhijian He	
14:30	Michael Tretyakov	Abdul-Lateef Haji-Ali	Chiheb Ben Hammouda	Jérôme Lelong	
15:30	Antoine Lejay	Nadhir Ben Rached	Florian Puchhammer	Christian Bayer	
16:00	Samuel Herrmann	Elisabeth Ullmann	David F. Anderson	Ahmed Kebaier	
16:30	Michael Mascagni	Jose Blanchet	Quan Long		
19:00					
19:30					
20:00					
20:30					

Numerics for SDEs with boundary issues (part A)

Back to overview

zoom 3, Monday, 14:00

A Walk on Cylinders algorithm

Francisco Bernal
Universidad Carlos III de Madrid, franciscomanuel.bernal@uc3m.es

Muller's classical Walk on Spheres (WoS) algorithm is arguably the method of choice to numerically integrate, in the weak sense, a Brownian motion bounded by Dirichlet data. For polygonal and box-like domains, Deaconu and Lejay introduced the Walk on Rectangles (WoR) method which takes the specific boundary geometry into account, resulting in fewer hops to the boundary. Motivated by WoS and WoR, I will introduce the Walk on Cylinders algorithm. It is intended for three-dimensional irregular domains whose boundary is "mostly flat". At every step, WoC chooses between a regular WoS step, or a hop onto the surface of the largest base cylinder standing on its boundary. Critically, the latter is constructed in an unsupervised way, relying on the distance map which implicitly defines the boundary. The potential for higher efficiency stems from the trade-off between fewer hops to the boundary on average, and the cost of sampling first-exit-point distributions out of cylinders.

Back to overview

zoom 3, Monday, 14:30

Weak approximation of reflected diffusions and sampling from distributions with compact support

Benedict Leimkuhler
University of Edinburgh,

Akash Sharma
University of Nottingham,

Michael Tretyakov
University of Nottingham, Michael.Tretyakov@nottingham.ac.uk

A simple-to-implement weak-sense numerical method to approximate reflected stochastic differential equations (RSDEs) is proposed and analysed. It is proved that the method has the first order of weak convergence. Together with the Monte Carlo technique, it can be used to numerically solve linear parabolic and elliptic PDEs with Robin boundary condition. One of the key results of this work is the use of the proposed method for computing ergodic limits, i.e. expectations with respect to the invariant law of RSDEs, both inside a domain and on its boundary. This allows to efficiently sample from distributions with compact support. Both time-averaging and ensemble-averaging estimators are considered and analysed. A number of extensions are considered, including a second-order weak approximation and the case of arbitrary oblique direction of reflection. The presented theoretical results are supported by several numerical experiments.

Back to overview

zoom 3, Monday, 15:30

A numerical scheme for diffusions processes in presence of interfaces

Antoine Lejay
Institut Élie Cartan de Lorraine & Inria Nancy Grand-Est, antoine.lejay@inria.fr

Lionel Lenôtre
Université Haute-Alsace, lionel.lenotre@uha.fr

Géraldine Pichot
Inria Paris, Geraldine.Pichot@inria.fr

In this talk, we present a new numerical scheme — presented in [1] — to deal with second-order diffusion processes evolving in a media with discontinuous coefficients and in presence of interfaces. In such situations, Euler schemes fails as the distribution of the position after a single step is far from the Gaussian distribution. The main idea of the scheme is to approximate the exact density with the Green kernel which may be simpler to compute thanks to Sturm-Liouville's theory. Numerical experiments show a good behavior of our scheme, which extends the one proposed in [2].

[1] A. Lejay, L. Lenôtre and G. Pichot. An exponential timestepping algorithm for diffusion with discontinuous coefficients *Journal of Computational Physics*, vol. 396, pp.888-904, 2019.

[2] K. M. Jansons and G. D. Lythe. Efficient numerical solution of stochastic differential equations using exponential timestepping. *J. Statist. Phys.*, 100(5-6):1097–1109, 2000.

Back to overview

zoom 3, Monday, 16:00

Exact simulation of the first time a diffusion process overcomes a given threshold

Samuel Herrmann
Université de Bourgogne, samuel.herrmann@u-bourgogne.fr

Nicolas Massin
Université de Bourgogne, nicolas.massin@u-bourgogne.fr

Cristina Zucca
University of Torino, cristina.zucca@unito.it

The aim of our study is to propose a new exact simulation method for the first passage time (FPT) of a diffusion process $(X_t,\, t \geq 0)$. We shall consider either a continuous diffusion process (in collaboration with Cristina Zucca, University of Turino) either a jump diffusion (in collaboration with Nicolas Massin, Université de Bourgogne). We define τ_L the first passage time through the level L:

$$\tau_L := \inf\{t \geq 0 : X_t \geq L\}.$$

In order to exactly simulate τ_L, we cannot use an explicit expression of its density. The classical way to overcome this difficulty is to use efficient algorithms for the simulation of sample paths, like discretization schemes. Such methods permit to obtain approximations of the first-passage times as a by-product.

For efficiency reasons, it is particularly challenging to simulate directly this hitting time by avoiding to construct the whole paths. The authors introduce a new rejection sampling algorithm which permits to perform an exact simulation of the first-passage time for general one-dimensional diffusion processes.

The main ideas are based both on a previous algorithm pointed out by A. Beskos et G. O. Roberts which uses Girsanov's transformation and on properties of Bessel paths. The efficiency of the method is described through theoretical results and numerical examples.

Back to overview

zoom 3, Monday, 16:30

The ZENO Code Uses Probabilistic Potential Theory to Compute Material Properties

Michael Mascagni
Florida State University and NIST, mascagni@fsu.edu

Walid Keyrouz
National Institute of Standards and Technology (NIST), Walid.Keyrouz@nist.gov

Derek Juba
National Institute of Standards and Technology (NIST), Derek.Juba@nist.gov

Computing material properties from molecular structure is of great use in many areas of study. To that end, the National Institute of Standards and Technology (NIST) developed the ZENO code to compute material properties using stochastic methods. A good example of such a property is the capacitance of an object. We think of capacitance as an electrical property, but it is in fact a very basic geometrical property, and can be used to compute many related quantities such as the permeability of porous media, [2].

We describe how capacitance can be computed using Brownian motion probabilities, and the show how this computation can be accelerated considerably using the walk-on-spheres (WOS) technique, [3, 1]. WOS is used in ZENO, and this shifts to computational bottleneck to the computational geometrical problem. A modern data structure considerably speeds up this computation and provided a substantial performance improvement in ZENO.

[1] P. Hamilin, W. J. Thrasher, W. Keyrouz and M. Mascagni (2019), "Geometry Entrapment in Walk-on-Subdomains," *Monte Carlo Methods and Applications*, **24(4)**: 178–193.

[2] C.-O. Hwang, J. A. Given and M. Mascagni (2000),"On the Rapid Calculation of Permeability for Porous Media Using Brownian Motion Paths," *Physics of Fluids*, **12**: 1699–1709.

[3] D. Juba, W. Keyrouz, M. Mascagni, M. Brady (2016), "Acceleration and Parallelization of ZENO/Walk-on-Spheres," *Procedia Computer Science*, **80**: 269–278.

Numerics for SDEs with boundary issues (part B)

Back to overview

zoom 3, Tuesday, 09:30

A multilevel Monte Carlo method for computing the action of a matrix exponential on a vector

Juan A. Acebrón
ISCTE-University Institute of Lisbon, juan.acebron@iscte-iul.pt

José R. Herrero
Universitat Politecnica de Catalunya, josepr@ac.upc.edu

José Monteiro
Instituto Superior Tecnico, jcm@inesc-id.pt

We introduce a novel probabilistic algorithm for computing the action of a matrix exponential over a vector. The algorithm is based on a multilevel Monte Carlo method, and the vector solution is computed generating suitable random paths which evolve through the indices of the matrix according to a suitable probability law. Among other potential applications, this algorithm has been applied for solving boundary-value parabolic PDEs by means of the method of lines, using an exponential integrator. When the method of lines is applied to the PDE problem discretizing the spatial variable, a system of ordinary differential equations, with time as the independent variable, is obtained. Finally, the system can be solved resorting to the computation of a matrix exponential which acts on the discretized initial value function. Since the boundary data are implicitly embedded in the discretization matrix, the method is free of the typical numerical boundary issues affecting any probabilistic numerical schemes for solving diffusion equations. Furthermore, the positive features of the algorithm in terms of parallelism were exploited in practice to develop a highly scalable implementation capable of solving some test problems very efficiently using high performance supercomputers equipped with a large number of cores.

Back to overview

zoom 3, Tuesday, 10:00

MLMC for Reflected Brownian Diffusions

Mike Giles
University of Oxford, mike.giles@maths.ox.ac.uk

In this talk we discuss the challenges in developing and analysing an efficient Multilevel Monte Carlo (MLMC) method for reflected Brownian diffusions. The difficulty is in treating the boundary reflections, where the current numerical approximations using projection [1, 2, 3, 5], reflection [1, 3], or penalization [4, 5, 6] are of strong order 1/2, at best, even when the discretisation of the SDE in the interior is first order. This suggests using a smaller timestep near the boundaries, and it is found computationally that this works very well as part of a MLMC construction. Initial steps towards the numerical analysis will be presented, but much more remains to be done.

[1] M. Bossy, E. Gobet, and D. Talay. Symmetrized Euler scheme for an efficient approximation of reflected diffusions. *Journal of Applied Probability*, 41(3):877–889, 2004.

[2] C. Constantini, B. Pacchiarotti and F. Sartoretto. Numerical approximation for functionals of reflecting diffusion processes. *SIAM Journal of Applied Mathematics*, 58(1):73-1-2, 1998.

[3] E. Gobet. Euler schemes and half-space approximations for the simulation of diffusion in a domain. *ENSIAM Probability and Statistics*, 5:261–297, 2001.

[4] L. Słomiński. On approximation of solutions of multidimensional SDE's with reflecting boundary conditions. *Stochastic Processes and their Applications*, 50(2):197–219, 1994.

[5] L. Słomiński. Euler's approximations of solutions of SDEs with reflecting boundary. *Stochastic Processes and their Applications*, 94(2):317–337, 2001.

[6] L. Słomiński. Weak and strong approximations of reflected diffusions via penalization methods. *Stochastic Processes and their Applications*, 123(3):752–763, 2013.

Back to overview

zoom 3, Tuesday, 10:30

Strong approximation for some stochastic processes

Madalina Deaconu
Inria Nancy and IECL, madalina.deaconu@inria.fr

Samuel Herrmann
University of Burgundy, Samuel.Herrmann@u-bourgogne.fr

In this talk we develop new techniques for the path approximation of stochastic processes, more precisely the Brownian motion, a family of stochastic differential equations sharply linked to the Brownian motion (usually known as L and G-classes) and the Bessel process, whatever the dimension. We are studying here the ε-strong approximation.

Our approach constructs jointly the sequences of exit times and corresponding exit positions of some well-chosen domains. We emphasize a numerical scheme which is easy to implement. The main results control the number of steps to cover a fixed time interval and the convergence theorems for our scheme. We expand the theoretical part by a series of numerical developments. This talk is based on several recent works by the authors.

Special Session Variance reduction techniques: Rare events

Back to overview

zoom 3, Tuesday, 14:00

Insights on Regenerative Simulation Estimators of Hitting Times to Rarely Visited Sets Based on Exponential Approximation

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

Marvin K. Nakayama
New Jersey Institute of Technology, marvin@njit.edu

Bruno Tuffin
Inria, Univ Rennes, CNRS, IRISA, bruno.tuffin@inria.fr

Determining the hitting time of a rarely visited set is of primary importance in many areas, such as dependability and finance. Estimating the full distribution of this hitting time is challenging and more informative than simply estimating its mean, as most of rare-events studies do. In a regenerative context where the process probabilistically restarts at regeneration times, and under the condition that the probability to reach the rare set before a regeneration occurs goes to 0 as a rarity parameter $\epsilon \to 0$, the distribution of the hitting time divided by its mean is known to converge to an exponential [2]. The estimation of the distribution is then reduced to estimating the mean, for which efficient estimators exist.

In a previous work [1], we designed regenerative estimators based on this exponential property, which resulted in so-called exponential and convolution estimators. The exponential estimator replaces the unknown mean in the limiting exponential distribution with a simulation estimator of the mean. The convolution estimator on the other hand exploits a similar weak convergence of the scaled sum of lengths of cycles between regenerations before the hitting the rare set to an exponential, and convolves it with the empirical distribution of hitting times within a cycle estimated by rare-event simulation techniques.

While our methods are statistically efficient, there remained questions about how widely the estimators could be used as well as about their bias (since the considered exponential distributions are approximations). Our goal during this talk is to provide insights on the quality of the estimators from a rare-event theoretical study of an analytically-tractable model. The example illustrates when the weak convergence to an exponential of hitting times holds, ensuring the validity of the exponential estimator, and why the convolution estimator can be more generally applied. We next perform on the same example a theoretical analysis of the bias and mean-squared relative error of the exponential estimator, something intractable in general, to illustrate its limiting behavior.

[1] P. W. Glynn, M. K. Nakayama, and B. Tuffin. Using simulation to calibrate exponential approximations to tail-distribution measures of hitting times to rarely visited sets. In M. Rabe, A. A. Juan, N. Mustafee, A. Skoogh, S. Jain, and B. Johansson, editors, *Proceedings of the 2018 Winter Simulation Conference*, pages 1802–1813, Piscataway, NJ, 2018. Institute of Electrical and Electronics Engineers.

[2] V. Kalashnikov. *Topics on Regenerative Processes*. CRC Press, Boca Raton, 1994.

Back to overview

zoom 3, Tuesday, 14:30

Multilevel Monte Carlo for Computing Probabilities

Abdul-Lateef Haji-Ali
Heriot-Watt University, a.hajiali@hw.ac.uk

Jonathan Spence
Heriot-Watt University, jws5@hw.ac.uk

Aretha Teckentrup
University of Edinburgh, a.teckentrup@ed.ac.uk

In this talk, I will discuss the challenges in computing probabilities of the form $\mathrm{P}[X \in \Omega]$ where X is a random variable and Ω is a d-dimensional set. Computing such probabilities is important in many contexts, e.g., risk assessment and finance. A frequent challenge is encountered when only a costly approximation of the random variable X can be sampled. For example, when X depends on an inner expectation that has to be approximated with Monte Carlo or when X depends on a non-trivial (stochastic) differential equations and a numerical discretization must be employed. A naive Monte Carlo method has a prohibitive complexity that compounds the slow convergence of Monte Carlo with the complexity of approximation. Instead, I will present a variant of Multilevel Monte Carlo (MLMC)

with adaptive levels that can, under certain conditions, have a complexity that is independent of the complexity of approximation. This work is an extension and a generalization of [5].

[1] M.B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 2015.

[2] M.B. Giles and A.-L. Haji-Ali. Multilevel nested simulation for efficient risk estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):497–525, 2019.

Back to overview

zoom 3, Tuesday, 15:30

Efficient importance sampling for large sums of independent and identically distributed random variables

Nadhir Ben Rached
RWTH Aachen University, benrached@uq.rwth-aachen.de

Abdul-Lateef Haji-Ali
Heriot-Watt University, a.hajiali@hw.ac.uk

Gerardo Rubino
INRIA, Rennes - Bretagne Atlantique, Gerardo.Rubino@inria.fr

Raúl Tempone
RWTH Aachen University and KAUST, raul.tempone@kaust.edu.sa

We consider the probability that the sum of nonnegative i.i.d random variables falls below a given threshold. We are particularly interested in the rare event regime corresponding to large sums and/or small threshold. Exponential twisting is a popular importance sampling (IS) technique that, in most cases, compares favorably to existing estimators. However, it has several limitations such as sampling under the new measure is not straightforward and might be expensive. We propose an alternative change of measure that yields at least the same performance as the exponential twisting technique and, at the same time, does not introduce serious limitations. For distributions whose probability density functions (PDFs) tend to zero polynomially, we prove that the Gamma IS PDF with appropriately chosen parameters retrieves asymptotically, in the rare event regime, the same performance of the estimator based on the use of the exponential twisting technique. Moreover, in the Log-normal setting, we numerically show that a Gamma IS PDF with optimized parameters clearly outperforms the exponential twisting change of measure. Numerical experiments validate the efficiency of the proposed estimator in delivering a highly accurate estimate in the regime of large sums and/or small threshold.

Back to overview

zoom 3, Tuesday, 16:00

Analysis and multilevel methods for rare event estimation with approximate models

Fabian Wagner
Technische Universität München, fabian.wagner@ma.tum.de

Jonas Latz
University of Cambridge, jl2160@cam.ac.uk

Iason Papaioannou
Technische Universität München, iason.papaioannou@tum.de

Elisabeth Ullmann
Technische Universität München, elisabeth.ullmann@tum.de

The estimation of the probability of rare events is an important task in reliability and risk assessment of critical societal systems, for example, groundwater flow and transport, and engineering structures. In this talk we consider rare events that are expressed in terms of a limit state function which depends on the solution of a partial differential equation (PDE). We present recent progress on mathematical and computational aspects of this problem: the impact of the PDE approximation error on the failure probability estimate [2], and a Multilevel Sequential Importance Sampling approach for the estimation of failure probabilities [1].

[1] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann. *Multilevel sequential importance sampling for rare event estimation.* SIAM J. Sci. Comput. 42(4), pp. A2062–A2087, 2020.

[2] F. Wagner, J. Latz, I. Papaioannou, and E. Ullmann. *Error analysis for probabilities of rare events with approximate models.* To appear in SIAM J. Numer. Anal., arXiv:2008.06368.

Back to overview

zoom 3, Tuesday, 16:30

Efficient Sampling for Worst-Case Rare Event Distributions

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

Most rare-event estimates are difficult to carry out in a purely non-parametric way. This is because the definition of a rare event inherently implies a lack of enough historical data which allows a good relative precision estimation. So, in the Monte Carlo analysis of rare events, one often uses a parametric family of benchmark/proxy models, leading to a potential problem in which model errors may overwhelm the Monte Carlo error incurred in estimating low probability events. To mitigate this problem, we consider introducing a systematic approach to robustify rare event probability estimates by maximizing a probability of interest over probability models that belong to an optimal transport neighborhood centered at a benchmark model. We characterize the worst-case probability model in several risk-sensitive examples motivated by finance and insurance applications and discuss efficient algorithms for sampling such worst-case distributions.

Back to overview

zoom 3, Wednesday, 09:00

MLMC treatments for discontinuous functionals

Mike Giles
University of Oxford, mike.giles@maths.ox.ac.uk

Relatively simple Multilevel Monte Carlo (MLMC) treatments often work well when the output quantity of interest is a Lipschitz function of intermediate variables, such as the solution of an SDE or a PDE with stochastic coefficients. In such cases, strong convergence results for the intermediate quantities lead naturally to MLMC variance bounds for the correction between one level of approximation and the next. However, when the output functional is discontinuous, a small difference in the intermediate quantities computed on successive levels can give an $O(1)$ difference in the output quantity of interest, leading to a much larger MLMC variance and a poorer overall complexity.

In this talk I will survey applications which have this challenge, and the various treatments which have been used to address it and obtain a good rate of decay of the MLMC variance. These include smoothing [6], conditional expectations [2], a change of measure [7], splitting [1, 4], and adaptive sampling [5], but time may not allow me to discuss all of these. Further details are available in the papers listed below.

[1] S. Burgos and M.B. Giles. Computing Greeks using multilevel path simulation. In L. Plaskota and H. Woźniakowski, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2010*, pages 281–296. Springer, 2012.

[2] M.B. Giles. Improved multilevel Monte Carlo convergence using the Milstein scheme. In A. Keller, S. Heinrich, and H. Niederreiter, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2006*, pages 343–358. Springer, 2008.

[3] M.B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 2015.

[4] M.B. Giles and F. Bernal. Multilevel estimation of expected exit times and other functionals of stopped diffusions. *SIAM/ASA Journal on Uncertainty Quantification*, 6(4):1454–1474, 2018.

[5] M.B. Giles and A.-L. Haji-Ali. Multilevel nested simulation for efficient risk estimation. *SIAM/ASA Journal on Uncertainty Quantification*, 7(2):497–525, 2019.

[6] M.B. Giles, T. Nagapetyan, and K. Ritter. Multilevel Monte Carlo approximation of distribution functions and densities. *SIAM/ASA Journal on Uncertainty Quantification*, 3(1):267–295, 2015.

[7] Y. Xia and M.B. Giles. Multilevel path simulation for jump-diffusion SDEs. In L. Plaskota and H. Woźniakowski, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2010*, pages 695–708. Springer, 2012.

Back to overview

zoom 3, Wednesday, 09:30

Computing Digital Options using Multilevel Monte Carlo and Nested Simulations

Mike Giles
University of Oxford, mike.giles@maths.ox.ac.uk

Abdul-Lateef Haji-Ali
Heriot-Watt University, a.hajiali@hw.ac.uk

In this talk we consider estimators for a digital option of the form $\mathrm{E}[g(X_T)\,]$ where X_T is the solution to a d-dimensional stochastic differential equation (SDE) at time $T > 0$ and $g(\cdot)$, the payoff function, is a *discontinuous* function such as an indicator function. For all but the most basic SDEs, samples of X_T must be approximated using numerical schemes like Euler-Maruyama and Milstein. Using Monte Carlo to estimate the value of a digital option is prohibitively expensive because it compounds the costs of sampling and approximation. Multilevel Monte Carlo (MLMC) [1] reduces the overall computational complexity by relying on correlated approximations to the solution of the SDE such that the variance of a difference of a fine and a coarse approximation converges as the cost of the approximations increases. Unfortunately, the efficiency MLMC is reduced when the payoff is discontinuous [2].

In this talk, we look at a new estimator based on nested expectations and which relies on the correlation of many approximate *paths* of X_T that share a part of the underlying Brownian paths. The result is that, for most cases, using MLMC for computing options with discontinuous payoffs can be made as efficient as when computing options with Lipschitz payoffs.

[1] M.B. Giles. Multilevel Monte Carlo methods. *Acta Numerica*, 24:259–328, 2015.

[2] M.B. Giles, D.J. Higham and X. Mao. Analysing multilevel Monte Carlo for options with non-globally Lipschitz payoff. Finance and Stochastics, 13(3):403-413, 2009.

Back to overview

zoom 3, Wednesday, 10:00

A multilevel subset simulation for estimating rare events via shaking transformations

Simon Weissmann
Universität Heidelberg, simon.weissmann@uni-heidelberg.de

In this talk, we analyse a multilevel version of subset simulation to estimate the probability of rare events for complex physical systems. Given a sequence of nested failure domains of increasing size, the rare event probability is expressed as a product of conditional probabilities. The proposed estimator uses different model resolutions and varying numbers of samples across the hierarchy of nested failure sets. The key idea in our proposed estimator is the use of a selective refinement strategy that guarantees the critical subset property which may be violated when changing model resolution from one failure set to the next. In order to estimate the probabilities of the underlying subsets we formulate and analyse a parallel one-path algorithm based on shaking transformations. Considering a physical model based on Gaussian transformation we can verify the ergodicity of the resulting Markov chain. Additionally, we present a detailed complexity analysis of the considered subset simulation. This is joint work with Daniel Elfverson, Robert Scheichl and Francisco Alejandro DiazDelaO.

Back to overview

zoom 3, Wednesday, 10:30

Adaptive Stratified Sampling for Non-smooth Problems

Per Pettersson
NORCE Norwegian Research Centre, per.pettersson@norceresearch.no

Sebastian Krumscheid
RWTH Aachen University, krumscheid@uq.rwth-aachen.de

Sampling based variance reduction techniques, such as multilevel Monte Carlo methods, have been established as a general-purpose procedure for quantifying uncertainties in computational models. It is known however, that these techniques may not provide performance gains when there is a non-smooth parameter dependence. Moreover, in many applications (e.g. transport problems in fractured porous media of relevance to carbon storage and wastewater injection) the key idea of MLMC cannot be fully exploited since a hierarchy of computational models cannot be easily constructed. An alternative means to obtain variance reduction in these cases is offered by stratified sampling methods. In this talk we will discuss various ideas on adaptive stratified sampling methods tailored to applications with a discontinuous parameter dependence. Specifically, we will build upon ideas from adaptive PDE mesh refinement strategies applied to the stochastic instead of the physical domain. The stochastic domain is adaptively stratified using local variance estimates, and the samples are sequentially allocated to the strata for asymptotically optimal variance reduction. The proposed methodology is demonstrated on discontinuous test cases from computational fluid mechanics and CO_2 storage in subsurface reservoirs, and computational speedup compared to standard Monte Carlo is obtained.

Variance Reduction Techniques: Pure Jumps and Stochastic Reaction Networks

Back to overview

zoom 3, Wednesday, 13:30

Efficient Importance Sampling via Optimal Control for Stochastic Reaction Networks

Chiheb Ben Hammouda
RWTH Aachen University, benhammouda@uq.rwth-aachen.de

Nadhir Ben Rached
RWTH Aachen University, benrached@uq.rwth-aachen.de

Raúl Tempone
RWTH Aachen University and KAUST, Saudi Arabia, raul.tempone@kaust.edu.sa

Sophia Wiechert
RWTH Aachen University, wiechert@uq.rwth-aachen.de

We are interested in the efficient estimation of statistical quantities, particularly rare event probabilities, for stochastic reaction networks (SRNs) and biochemical systems. To this aim, we propose a novel importance sampling (IS) approach to improve the efficiency of Monte Carlo (MC) estimators when based on an approximate tau-leap (TL) scheme. The IS framework's crucial step is to choose an appropriate change of probability measure to achieve a substantial variance reduction. In general, this is challenging, and it often requires insights into the given problem. Based on an original connection between finding the

optimal IS parameters, within a class of probability measures, and a stochastic optimal control (SOC) formulation, we propose an automated approach to derive a highly efficient path-dependent measure change. Our optimal IS parameters are obtained by solving a variance minimization problem. Given that it is challenging to solve the SOC problem analytically, we propose an efficient numerical algorithm to approximate the optimal control parameters. Our analysis and the conducted numerical experiments demonstrate that our novel IS approach substantially reduces the MC estimator's variance. Moreover, our numerical results show that the variance of our proposed estimator decays with rate $\mathcal{O}(\Delta t)$ for a step size of Δt, compared to being $\mathcal{O}(1)$ for the standard MC-TL estimator. Given a prescribed error tolerance, TOL, this implies an improvement of the computational complexity to reach $\mathcal{O}(TOL^{-2})$ instead of $\mathcal{O}(TOL^{-3})$ when using the standard MC-TL estimator. Previously, this optimal complexity was only achieved using a scheme with exact steps. However, thanks to our original IS approach, we reach this optimal complexity with a much lower constant, using the TL approximate scheme.

Back to overview

zoom 3, Wednesday, 14:00

Control Variates for Stochastic Reaction Networks

Michael Backenköhler
Universität des Saarlandes, michael.backenkoehler@uni-saarland.de

Luca Bortolussi
Trieste University, lbortolussi@units.it

Verena Wolf
Universität des Saarlandes, wolf@cs.uni-saarland.de

Stochastic simulation is a widely used method for estimating quantities in models of chemical reaction networks where uncertainty plays a crucial role. However, reducing the statistical uncertainty of the corresponding estimators requires the generation of a large number of simulation runs, which is computationally expensive. To reduce the number of necessary runs, we propose to use the variance reduction technique of linear control variates. We derive control variates from constraints on the statistical moments of the stochastic process. These are then computed alongside each Monte-Carlo simulation and are used for an estimator with reduced variance. We develop an algorithm that selects appropriate control variates in an online fashion. We demonstrate the efficiency (variance reduction and cost overhead trade-off) of our approach on several case studies.

[1] M. Backenköhler, L. Bortolussi, V. Wolf *Control Variates for Stochastic Simulation of Chemical Reaction Networks*. Springer LNCS 11773 pp. 42-59, 2019.

Back to overview

zoom 3, Wednesday, 14:30

Importance sampling for a robust and efficient multilevel Monte Carlo estimator for stochastic reaction networks

Chiheb Ben Hammouda
RWTH Aachen University, benhammouda@uq.rwth-aachen.de

Nadhir Ben Rached
RWTH Aachen University, benrached@uq.rwth-aachen.de

Raúl Tempone
RWTH Aachen University and KAUST, Saudi Arabia, tempone@uq.rwth-aachen.de

The multilevel Monte Carlo (MLMC) method for continuous-time Markov chains, first introduced by Anderson and Higham (SIAM Multiscal Model Simul 10(1):146–179, 2012), is a highly efficient simulation technique that can be used to estimate various statistical quantities for stochastic reaction networks (SRNs), in particular for stochastic biological systems. Unfortunately, the robustness and performance of the multilevel method can be affected by the high kurtosis, a phenomenon observed at the deep levels of MLMC, which leads to inaccurate estimates of the sample variance. In [1], we address cases where the high-kurtosis phenomenon is due to *catastrophic coupling* (characteristic of pure jump processes where coupled consecutive paths are identical in most of the simulations, while differences only appear in a tiny proportion) and introduce a pathwise-dependent importance sampling (IS) technique that improves the robustness and efficiency of the multilevel method. Our theoretical results, along with the conducted numerical experiments, demonstrate that our proposed method significantly reduces the kurtosis of the deep levels of MLMC, and also improves the strong convergence rate from $\beta = 1$ for the standard case (without IS), to $\beta = 1 + \delta$, where $0 < \delta < 1$ is a user-selected parameter in our IS algorithm. Due to the complexity theorem of MLMC, and given a pre-selected tolerance, TOL, this results in an improvement of the complexity from $\mathcal{O}\left(\mathrm{TOL}^{-2}\log(\mathrm{TOL})^2\right)$ in the standard case to $\mathcal{O}\left(\mathrm{TOL}^{-2}\right)$, which is the optimal complexity of the MLMC estimator. We achieve all these improvements with a negligible additional cost since our IS algorithm is only applied a few times across each simulated path.

[1] Ben Hammouda, Chiheb and Ben Rached, Nadhir and Tempone, Raúl. *Importance sampling for a robust and efficient multilevel Monte Carlo estimator for stochastic reaction networks* . Statistics and Computing, Springer, Volume 30, 2020.

Back to overview

zoom 3, Wednesday, 15:30

Simulating Chemical Reaction Networks with Tau-Leaping and Array-RQMC

Florian Puchhammer
University of Waterloo, florian.puchhammer@uwaterloo.ca

Amal Ben Abdellah
Université de Montréal, amal.ben.abdellah@umontreal.ca

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

The τ-leap algorithm by Gillespie is used to simulate stochastic biological systems and chemical reaction networks as a discrete time Markov chain. The objective is to estimate the expectation of a function of the molecule numbers at a given future time. We combine this approach with a powerful randomized quasi-Monte Carlo (RQMC) technique, Array-RQMC, which has been proven to outperform Monte Carlo and conventional RQMC in many applications [2, 3, 4]. For chemical reaction networks, this combination can reduce the variance by factors in the thousands, compared to Monte Carlo. Even more than it was observed for ordinary RQMC in [1].

Array-RQMC efficiently simulates an array of realizations of the Markov chain in parallel, but requires to sort the chains at each step by their (multi-dimensional) states. In our simulations, Array-RQMC always performed better than ordinary MC, regardless of the sort. Yet, the choice of the sorting function can have a significant impact on both variance reduction and simulation speed. We demonstrate how different sorting algorithms can increase the efficiency of the method and discuss how they can be customized for specific applications and for different performance measures.

[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] 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.

[3] P. L'Ecuyer, C. Lécot, and A. L'Archevêque-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.

[4] 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.

Back to overview

zoom 3, Wednesday, 16:00

Variance of finite difference methods for reaction networks with non-Lipschitz rate functions

David F. Anderson
University of Wisconsin-Madison, anderson@math.wisc.edu

Chaojie Yuan
Indiana Unversity Bloomington, yuan13@iu.edu

In this talk, I will discuss recent work that provides sharp bounds on the variance of finite difference methods in the context of stochastic reaction networks. The analysis is valid for most models from the literature, even those with non-Lipschitz rate functions. This is joint work with Chaojie Yuan at the University of Indiana.

[1] D. F. Anderson and C. Yuan, Variance of Finite Difference Methods for Reaction Networks with non-Lipschitz Rate Functions, *SIAM Journal on Numerical Analysis*, Vol. 58, No. 6, 2020.

Back to overview

zoom 3, Wednesday, 16:30

Efficient Multilevel Supply Chain Management Under Uncertainties

Quan Long
Honeywell International, quan.spartanlq@gmail.com

Uncertainty propagation of large scale discrete supply chains can be prohibitive when a large number of events occur during the simulated period and discrete event simulations (DES) are costly. We present a time bucket method to approximate and accelerate the DES of supply chains. Its stochastic version, which we call the L(logistic)-leap method, can be viewed as an extension of the leap methods developed in the chemical engineering community for the acceleration of stochastic DES of chemical reactions. The L-leap method instantaneously updates the system state vector at discrete time points and the production rates and policies of a supply chain are assumed to be stationary during each time bucket. We propose to use Multilevel Monte Carlo (MLMC) to efficiently propagate the uncertainties in a supply chain network, where the levels are naturally defined by the sizes of the time buckets of the simulations.

We demonstrate the efficiency and accuracy of our methods using numerical examples derived from a real world manufacturing material flow. In these examples, our multilevel L-leap approach can be faster than the standard Monte Carlo (MC) method by one or two orders of magnitude without compromising the accuracy.

[1] NY Chiang, Y. Lin and Q. Long. *Efficient propagation of uncertainties in manufacturing supply chains: Time buckets, L-leap, and multilevel Monte Carlo methods.* Operations Research Perspective, Volume 7, 2020.

Special Session Theory and Application of Particle Systems

Back to overview

zoom 3, Thursday, 09:00

Propagation of chaos for mean-field reflected BSDE with jumps

Boualem Djehiche
KTH Royal Institute of Technology, boualem@kth.se

We establish a propagation of chaos of a system of weakly interacting processes which solve a system of mean field reflected backward stochastic differential equations with jumps, with interconnected obstacles.

Back to overview

zoom 3, Thursday, 09:30

A variational approach to nonlinear and interacting diffusions

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

This talk presents a novel variational calculus to analyze the stability and the propagation of chaos properties of nonlinear and interacting diffusions.

This differential methodology combines gradient flow estimates with backward stochastic interpolations, Lyapunov linearization techniques as well as spectral theory.

This framework applies to a large class of stochastic models including non homogeneous diffusions, as well as stochastic processes evolving on differentiable manifolds, such as constraint-type embedded manifolds on Euclidian spaces and manifolds equipped with some Riemannian metric.

We present uniform as well as almost sure exponential contraction inequalities at the level of the nonlinear diffusion flow, as well as, uniform propagation of chaos properties w.r.t. the time parameter are also provided. Illustrations are provided in the context of a class of gradient flow diffusions arising in fluid mechanics and granular media literature.

Back to overview

zoom 3, Thursday, 10:00

Particle filters with nudging. Applications to Data Assimilation

Dan Crisan
Imperial College London, d.crisan@ic.ac.uk

Igor Shevchenko
Imperial College London, i.shevchenko@ic.ac.uk

This talk covers some recent work of the authors' project on developing particle filters based data assimilation methodology for high dimensional fluid dynamics models. The algorithm presented here is a particle filter with a so-called "nudging" mechanism. The nudging procedure is used in the prediction step. In the absence of nudging, the particles have trajectories that are independent solutions of the model equations. The nudging presented here consists in adding a drift to the trajectories of the particles with the aim of maximising the likelihood of their positions given the observation data. This introduces a bias in the system that is corrected during the resampling step. The nudging procedure is theoretically justified through a standard convergence argument. The Data Assimilation algorithm presented in this paper gives an asymptotically (as the number of particles increases) consistent approximation of the posterior distribution of the state given the data. The methodology is tested on a two-layer quasi-geostrophic model for a β-plane channel flow with $O(10^6)$ degrees of freedom out of which only a minute fraction are noisily observed. I will present the effect of the nudging procedure on the performance of the data assimilation procedure for a reduced model in terms of the accuracy and uncertainty of the results. The results presented here are incorporated in [1] and [2].

[1] C Cotter, D Crisan, D Holm, W Pan, I Shevchenko, *Data assimilation for a quasi-geostrophic model with circulation-preserving stochastic transport noise*, Journal of Statistical Physics, 1-36, 2020.

[2] D Crisan, I Shevchenko, *Particle filters with nudging*, work in progress.

Back to overview

zoom 3, Thursday, 10:30

A simple approach to proving the existence, uniqueness, and strong and weak convergence rates for a broad class of McKean–Vlasov equations

Abdul-Lateef Haji-Ali
Heriot-Watt University, a.hajiali@hw.ac.uk

Håkon Hoel
RWTH Aachen University, hoel@uq.rwth-aachen.de

Raúl Tempone
RWTH Aachen University and KAUST, raul.tempone@kaust.edu.sa

By employing a system of interacting stochastic particles as an approximation of the McKean–Vlasov equation and utilizing classical stochastic analysis tools, namely Ito's formula and Kolmogorov–Chentsov continuity theorem, we prove the existence and uniqueness of strong solutions for a broad class of McKean–Vlasov equations as a limit of the conditional expectation of exchangeable particles. Considering an increasing number of particles in the approximating stochastic particle system, we also prove the L^p strong convergence rate and derive the weak convergence rates using the Kolmogorov backward equation and variations of the stochastic particle system. Our convergence rates were verified by numerical experiments which also indicate that the assumptions made here and in the literature can be relaxed.

Back to overview

zoom 3, Thursday, 14:00

Efficient risk estimation via nested multilevel quasi-Monte Carlo simulation

Zhenghang Xu
Tsinghua University, xzh17@mails.tsinghua.edu.cn

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

Xiaoqun Wang
Tsinghua University, wangxiaoqun@mail.tsinghua.edu.cn

We consider the problem of estimating the probability of a large loss from a financial portfolio, where the future loss is expressed as a conditional expectation. Since the conditional expectation is intractable in most cases, one may resort to nested simulation. To reduce the complexity of nested simulation, we present a method that combines multilevel Monte Carlo (MLMC) and quasi-Monte Carlo (QMC). In the outer simulation, we use Monte Carlo to generate financial scenarios. In the inner simulation, we use QMC to estimate the portfolio loss in each scenario. We prove that using QMC can accelerate the convergence rates in both the crude nested simulation and the multilevel nested simulation. Under certain conditions, the complexity of MLMC can be reduced to $O(\epsilon^{-2}(\log \epsilon)^2)$ by incorporating QMC. On the other hand, we find that MLMC encounters catastrophic coupling problem due to the existence of indicator functions. To remedy this, we propose a smoothed MLMC method which uses logistic sigmoid functions to approximate indicator functions. Numerical results show that the optimal complexity $O(\epsilon^{-2})$ is almost attained when using QMC methods in both MLMC and smoothed MLMC, even in moderate high dimensions.

Back to overview

zoom 3, Thursday, 14:30

Automatic Control Variates for option pricing using Neural Networks

Jérôme Lelong
Université Grenoble Alpes, jerome.lelong@univ-grenoble-alpes.fr

Zineb El Filali Ech-Chafiq
Université Grenoble Alpes and Natixis, zinebelfilaliechchafique@gmail.com

Adil Reghai
Natixi, adil.reghai@natixis.com

Many pricing problems boil down to the computation of a high dimensional integral, which is usually estimated using Monte Carlo. In fact, the accuracy of a Monte Carlo estimator with M simulations is given by $\frac{\sigma}{\sqrt{M}}$. Meaning that its convergence is immune to the dimension of the problem. However, this convergence can be relatively slow depending on the variance σ of the function to be integrated. To resolve such a problem, one would perform some variance reduction techniques such as importance sampling, stratification, or control variates. In this paper, we will study two approaches for improving the convergence of Monte Carlo using Neural Networks. The first approach relies on the fact that many

high dimensional financial problems are of low effective dimensions. We expose a method to reduce the dimension of such problems in order to keep only the necessary variables. The integration can then be done using fast numerical integration techniques such as Gaussian quadrature. The second approach consists in building an automatic control variate using neural networks. We learn the function to be integrated (which incorporates the diffusion model plus the payoff function) in order to build a network that is highly correlated to it. As the network that we use can be integrated exactly, we can use it as a control variate.

Back to overview

zoom 3, Thursday, 15:30

Numerical smoothing and hierarchical approximations for efficient option pricing and density estimation

Christian Bayer
Weierstrass Institute, christian.bayer@wias-berlin.de

Chiheb Ben Hammouda
RWTH Aachen, benhammouda@uq.rwth-aachen.de

Raúl Tempone
RWTH Aachen & KAUST, tempone@uq.rwth-aachen.de

When approximating the expectation of a functional of a certain stochastic process, the efficiency and performance of deterministic quadrature methods such as sparse grids, and hierarchical variance reduction methods such as multilevel Monte Carlo (MLMC), may be highly deteriorated in different ways by the low regularity of the integrand with respect to the input parameters. To overcome this issue, a smoothing procedure is needed to uncover the available regularity and improve the performance of the aforementioned numerical methods. In this work, we consider cases where we cannot perform an analytic smoothing. Thus, we introduce a novel numerical smoothing technique based on root-finding combined with a one dimensional integration with respect to a single well-chosen variable. We prove that under appropriate conditions, the resulting function of the remaining variables is a highly smooth function, potentially allowing a higher efficiency of adaptive sparse grids quadrature (ASGQ), in particular when combined with hierarchical transformations (Brownian bridge and Richardson extrapolation on the weak error) to treat high dimensionality effectively. Our study is motivated by option pricing problems and our main focus is on dynamics where a discretization of the asset price is needed. Through our analysis and numerical experiments, we illustrate the advantage of combining numerical smoothing with ASGQ compared to the Monte Carlo (MC) approach. Furthermore, we demonstrate how numerical smoothing significantly reduces the kurtosis at the deep levels of MLMC, and also improves the strong convergence rate, when using Euler scheme. Due to the complexity theorem of MLMC, and given a pre-selected tolerance, TOL, this results in an estimated improvement of the complexity from $\mathcal{O}(\mathrm{TOL}^{-2.5})$ in the standard case to $\mathcal{O}(\mathrm{TOL}^{-2}\log(\mathrm{TOL})^2)$. Finally, we show how our numerical smoothing combined with MLMC enables us also to estimate density functions.

Back to overview

zoom 3, Thursday, 16:00

Approximation of Stochastic Volterra Equations

Aurélien Alfonsi
CERMICS, École des Ponts, aurelien.alfonsi@enpc.fr

Ahmed Kebaier
Université Sorbonne Paris Nord, kebaier@math.univ-paris13.fr

In this talk, we present a multi-factor approximation for Stochastic Volterra Equations with Lipschitz coefficients and kernels of completely monotone type that may be singular. Our approach consists in truncating and then discretizing the integral defining the kernel, which corresponds to a classical Stochastic Differential Equation. We prove strong convergence results for this approximation. For the particular rough kernel case with Hurst parameter lying in $(0, 1/2)$, we propose various discretization procedures and give their precise rates of convergence. We illustrate the efficiency of our approximation schemes with numerical tests for the rough Bergomi model.

Contributed talks on Regression and Learning

Back to overview

zoom 3, Friday, 09:30

On least squares approximation in linear inverse statistical learning problems

Tapio Helin
LUT University, tapio.helin@lut.fi

Vesa Kaarnioja
LUT University, vesa.kaarnioja@lut.fi

In linear inverse problems, the goal is to recover an unknown function f based on a measurement model

$$y = Af,$$

where y is the exact datum and $A\hat{A}$ is a linear bounded operator. In general, this problem is ill-posed, i.e., a solution f may not exist, be unique, or it does not depend continuously on the datum y. In the framework of learning problems, one is interested in recovering f based on an ensemble of noisy observations $(y_i)_{i=1}^n$ such that

$$y_i = (Af)(x_i) + \text{noise}, \quad i \in \{1, \dots, n\},$$

where the sample $(x_i)_{i=1}^n$ follows a random distribution. In this work, we consider the least squares approximation of the unknown function f based on a set of noisy observations, prove rigorous error bounds for the approximate least squares solution of the inverse problem, and present numerical results that assess the theoretical convergence rates.

Back to overview

zoom 3, Friday, 10:00

Low-discrepany-sequences and Regression Design

Zoran Nikolic
Universität Köln, znikolic@uni-koeln.de

Christian Weiß
Hochschule Ruhr West, christian.weiss@hs-ruhrwest.de

Practitioners sometimes suggest to use a combination of Sobol sequences and orthonormal polynomials when applying an regression algorithms. A typical example would be LSMC approaches for evaluation of option prices or risk capital calculation under the Solvency II regime. In this talk, we give a theoretical justification why good implementations of an LSMC algorithm should indeed combine these two features in order to assure numerical stability. Moreover, an explicit bound for the number of outer scenarios necessary to guarantee a prescribed degree of numerical stability is derived. Numerical experiments supporting our theoretical results are presented.

[1] C. Weiß, Z. Nikolic, *An aspect of optimal regression design for LSMC*. Monte Carlo Methods Appl., vol. 25(4), pp. 283–291 (2019).

Back to overview

zoom 3, Friday, 10:30

Multilevel Monte Carlo learning

Thomas Gerstner
University Frankfurt, gerstner@math.uni-frankfurt.de

Bastian Harrach
University Frankfurt, harrach@math.uni-frankfurt.de

Daniel Roth
University Frankfurt, roth@math.uni-frankfurt.de

Martin Simon
Deka Investment GmbH, -

In this work, we study the approximation of expected values of functional quantities on the solution of a stochastic differential equation (SDE), where we replace the Monte Carlo estimation with the evaluation of a deep neural network. Once the neural network training is done, the evaluation of the resulting approximating function is computationally highly efficient so that using deep neural networks to replace costly Monte Carlo integration is appealing, e.g., for near real-time computations in quantitative finance. However, the drawback of these nowadays widespread ideas lies in the fact that training a suitable neural network is likely to be prohibitive in terms of computational cost. We address this drawback here by introducing a multilevel approach to the training of deep neural networks. More precisely, we combine the deep learning algorithm introduced by Beck et al. [1] with the idea of multilevel Monte Carlo path simulation of Giles [2]. The idea is to train several neural networks, each having a certain approximation quality and computational complexity, with training data computed from so-called *level estimators*, introduced by Giles [2]. We show that under certain assumptions, the variance in the training process can be reduced by shifting most of the computational workload to training neural nets at coarse levels where producing the training data sets is comparably cheap, whereas training the neural nets corresponding to the fine levels requires only a limited number of training data sets. We formulate a complexity theorem showing that the multilevel idea can indeed reduce computational complexity.

[1] Beck, Christian and Becker, Sebastian and Grohs, Philipp and Jaafari, Nor and Jentzen, Arnulf *Solving stochastic differential equations and Kolmogorov equations by means of deep learning.* arXiv preprint arXiv:1806.00421, 2018

[2] Giles, Michael B *Multilevel monte carlo path simulation.* Operations Research, Volume 56, 2008.

6.4 Zoom Channel 4

back to table of contents

Clicking on the name of the speaker will lead you to the abstract. NOTE THAT ALL TIMES ARE UTC+2.

	Monday	Tuesday	Wednesday	Thursday	Friday
		Special Session **(Part B) Stochastic Computation and Complexity**	Special Session **(Part D) Stochastic Computation and Complexity**	Special Session **(Part A) SPDEs Simulation and Modelling**	Special Session **(Part B) SPDEs Simulation and Modelling**
09:00		Stefan Heinrich	Dai Taguchi		Christoph Reisinger
09:30		Tomasz Bochacik	Yue Wu	Ricarda Missfeldt	Annika Lang
10:00		David Krieg	Máté Gerencsér	Felix Lindner	Kristin Kirchner
10:30		Mathias Sonnleitner	Michaela Szölgyenyi	Ludovic Goudenège	Erika Hausenblas
	Special Session **(Part A) Stochastic Computation and Complexity**	Special Session **(Part C) Stochastic Computation and Complexity**	Special Session **(Part E) Stochastic Computation and Complexity**	Special Session **Quantitative Aspects of SDEs**	
13:30	Arnulf Jentzen	Sotirios Sabanis	Wolfgang Stockinger	Sonja Cox	
14:00	Steffen Dereich	Konstantinos Dareiotis	Christoph Reisinger	Stefan Geiss	
14:30	Diyora Salimova	Paweł Przybyłowicz	Raphael Kruse	Emmanuel Gobet	
15:30	Monika Eisenmann	Thomas Müller-Gronbach		Dan Crisan	
16:00	Thomas Kruse	Annalena Mickel		Céline Labart	
16:30	Emmanuel Gobet	Andreas Neuenkirch		Till Massing	
19:00					
19:30					
20:00					
20:30					

Back to overview

zoom 4, Monday, 13:30

Convergence analysis for the gradient descent optimization method in the training of artificial neural networks with ReLU activation for piecewise linear target functions

Arnulf Jentzen
University of Muenster, ajentzen@uni-muenster.de

Adrian Riekert
University of Muenster, ariekert@uni-muenster.de

Gradient descent (GD) type optimization methods are the standard instrument to train artificial neural networks (ANNs) with rectified linear unit (ReLU) activation. Despite the great success of GD type optimization methods in numerical simulations for the training of ANNs with ReLU activation, it remains – even in the simplest situation of the plain vanilla GD optimization method with random initializations – an open problem to prove (or disprove) the conjecture that the true risk of the GD optimization method converges in the training of ANNs with ReLU activation to zero as the width of the ANNs, the number of independent random initializations, and the number of GD steps increase to infinity. In this talk we prove this conjecture in the special situation where the probability distribution of the input data is absolutely continuous with respect to the continuous uniform distribution on a compact interval and where the target function under consideriation is piecewise linear.

Back to overview

zoom 4, Monday, 14:00

Convergence of stochastic gradient descent schemes for Łojasiewicz-landscapes

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

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

We consider convergence of stochastic gradient descent schemes (SGD) under weak assumptions on the underlying landscape. More explicitly, we show that on the event that the SGD stays local we have convergence of the SGD if there is only a countable number of critical points or if the target function/landscape satisfies Łojasiewicz-inequalities around all critical levels as all analytic functions do. In particular, we show that for neural networks with analytic activation function such as softplus, sigmoid and the hyperbolic tangent, SGD converges on the event of staying local, if the random variables modelling the signal and response in the training are compactly supported.

[1] S. Dereich and S. Kassing. Convergence of stochastic gradient descent schemes for Łojasiewicz-landscapes. arXiv:2102.09385 [cs.LG], 2021.

Back to overview

zoom 4, Monday, 14:30

Weak error analysis for stochastic gradient descent optimization algorithms

Diyora Salimova
ETH Zurich, sdiyora@mins.ee.ethz.ch

Aritz Bercher
ETH Zurich, abercher@outlook.com

Lukas Gonon
LMU Mathematics Institute, gonon@math.lmu.de

Arnulf Jentzen
University of Münster, ajentzen@uni-muenster.de

Stochastic gradient descent (SGD) type optimization schemes are fundamental ingredients in a large number of machine learning based algorithms. In mathematical convergence results for SGD type optimization schemes there are usually two types of error criteria studied in the scientific literature, that is, the error in the strong sense and the error with respect to the objective function. In applications one is often not only interested in the size of the error with respect to the objective function but also in the size of the error with respect to a test function which is possibly different from the objective function. The analysis of the size of this error is the subject of this talk. In particular, we show that under suitable assumptions that the size of this error decays at the same speed as in the special case where the test function coincides with the objective function.

Back to overview

zoom 4, Monday, 15:30

Sub-linear convergence of a tamed stochastic optimization method

Monika Eisenmann
Lund University, monika.eisenmann@math.lth.se

Tony Stillfjord
Lund University, tony.stillfjord@math.lth.se

In order to solve a minimization problem, a possible approach is to find the steady state of the corresponding gradient flow initial value problem through a long time integration. The well-known stochastic gradient descent (SGD) method then corresponds to the forward Euler scheme with a stochastic approximation of the gradient. Our goal is to find more suitable schemes that work well in the stochastic setting.

In the talk, we present a stochastic version of the tamed Euler scheme in this context. This method is fully explicit but is more stable for larger step sizes compared to the standard SGD method. We provide convergence results with a sub-linear rate also in an infinite-dimensional setting. We will illustrate the theoretical results on numerical examples. A typical application for such optimization problems is supervised learning.

Back to overview

zoom 4, Monday, 16:00

Multilevel Picard approximations for high-dimensional semilinear parabolic partial differential equations

Weinan E
Princeton University, weinan@math.princeton.edu

Martin Hutzenthaler
Universität Duisburg-Essen, martin.hutzenthaler@uni-due.de

Arnulf Jentzen
Universität Münster, a.j@uni-muenster.de

Thomas Kruse
Universität Giessen, thomas.kruse@math.uni-giessen.de

Tuan Nguyen
Universität Duisburg-Essen, tuan.nguyen@uni-due.de

Philippe von Wurstemberger
ETH Zurich, philippe.vonwurstemberger@math.ethz.ch

We present new approximation methods for high-dimensional semilinear parabolic PDEs. A key idea of our methods is to combine multilevel approximations with Picard fixed-point approximations. We prove in the case of semilinear heat equations with Lipschitz continuous nonlinearities that the computational effort of one of the proposed methods grows polynomially both in the dimension and in the reciprocal of the required accuracy. We illustrate the efficiency of the approximation methods by means of numerical simulations.

Back to overview

zoom 4, Monday, 16:30

A generative model for fBm with deep ReLU neural networks

Michaël Allouche
Ecole polytechnique, michael.allouche@polytechnique.edu

Stéphane Girard
INRIA Grenoble, stephane.girard@inria.fr

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

We provide a large probability bound on the uniform approximation of fractional Brownian motion with Hurst parameter H, by a deep-feedforward ReLU neural network fed with a N-dimensional Gaussian vector, with bounds on the network construction (number of hidden layers and total number of neurons). Essentially, up to log terms, achieving an uniform error of $\mathcal{O}(N^{-H})$ is possible with $\log(N)$ hidden layers and $\mathcal{O}(N)$ parameters. Our analysis relies, in the standard Brownian motion case ($H = 1/2$), on the Levy construction and in the general fractional Brownian motion case ($H \neq 1/2$), on the Lemarié-Meyer wavelet representation. This work gives theoretical support on new generative models based on neural networks for simulating continuous-time processes.

Back to overview

zoom 4, Tuesday, 09:00

Randomized Bit Complexity of Parametric Integration in Sobolev Spaces

Stefan Heinrich
Department of Computer Science, University of Kaiserslautern, heinrich@informatik.uni-kl.de

We study the n-th minimal errors (and hence the complexity) of randomized computation of integrals depending on a parameter, with integrands from Sobolev spaces. That is, for $r, d_1, d_2 \in \mathbf{N}$, $1 \le p \le \infty$, $D_1 = [0,1]^{d_1}$, and $D_2 = [0,1]^{d_2}$ we are given $f \in W_p^r(D_1 \times D_2)$ and we seek to approximate

$$Sf = \int_{D_2} f(s,t)dt \quad (s \in D_1),$$

with error measured in the $L_p(D_1)$-norm. We work in both the restricted and unrestricted randomized setting. Restricted means that the considered algorithms have only access to random bits, while in the unrestricted setting they can use arbitrary randomness.

Wiegand [1] has determined the n-th minimal errors in the unrestricted setting, under the assumption that $W_p^r(D_1 \times D_2)$ is continuously embedded in $C(D_1 \times D_2)$ (embedding condition). We show that in the restricted setting the same rates can be obtained with only a logarithmic number of random bits.

We also determine the n-th minimal errors in the unrestricted setting for the case that the embedding condition does not hold and show that in this case the restricted setting gives no rate whatsoever.

[1] C. Wiegand, Optimal Monte Carlo and Quantum Algorithms for Parametric Integration, Shaker Verlag, 2006.

Back to overview

zoom 4, Tuesday, 09:30

Randomized numerical algorithms approximating solutions of ODEs

Tomasz Bochacik
AGH University of Science and Technology, Krakow, bochacik@agh.edu.pl

Maciej Goćwin
AGH University of Science and Technology, Krakow, gocwin@agh.edu.pl

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

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

We will present the results from [1, 2] concerning randomized versions of three classical algorithms approximating solutions of ODEs: explicit Euler, implicit Euler and two-stage Runge-Kutta. Firstly, we will establish error bounds of the aforementioned methods in the setting of inexact information and under mild assumptions about the right-hand side function (including local Lipschitz continuity with respect to the space variable). Similar analysis for exact information and slightly different assumptions has been performed in [3, 5]. Furthermore, we will provide lower bounds and discuss optimality of the considered methods (in the sense of attaining the n-th minimal error). The last part of the talk will be devoted to stablility of these algorithms. We will adopt the concept of probabilistic stability regions investigated in [4] in the context of SDEs to randomized numerical methods for ODEs.

[1] T. Bochacik, M. Goćwin, P.M. Morkisz, P. Przybyłowicz. *Randomized Runge-Kutta method – stability and convergence under inexact information.* J. Complex. **65** (2021), 101554.

[2] T. Bochacik, P. Przybyłowicz. *On the randomized Euler schemes for ODEs under inexact information* (2021), preprint: https://arxiv.org/pdf/2104.15071.pdf.

[3] M. Eisenmann, M. Kovács, R. Kruse, S. Larsson. *On a randomized backward Euler method for nonlinear evolution equations with time-irregular coefficients.* Found. Comp. Math. **19** (2019), 1387–1430.

[4] D.J. Higham. *Mean-square and asymptotic stability of the stochastic theta method.* Siam J. Numer. Anal. **38** (2000), 753–769.

[5] R. Kruse, Y. Wu. *Error analysis of randomized Runge-Kutta methods for differential equations with time-irregular coefficients.* Comput. Methods Appl. Math. **17** (2017), 479–498.

Back to overview

zoom 4, Tuesday, 10:00

Sampling recovery in L_2

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

Mario Ullrich
JKU Linz, Moscow State University, mario.ullrich@jku.at

We consider the recovery of functions coming from a compact subset F of L_2. We are interested in the worst case error of algorithms using a finite number of samples. We only assume that function evaluation is continuous on F.

The following result was recently obtained by the authors: If the Kolmogorov widths of F show a polynomial decay of order $s > 1/2$, then there is a least squares estimator that achieves the same rate of convergence.

We discuss this result and address the following questions: What does the algorithm look like? What can be said in the case $s \leq 1/2$? What results do we obtain for the tractability of the problem in high dimensional settings?

We also relate to recent results of Nagel/Schäfer/Ullrich, Temlyakov and Cohen/Dolbeault.

Back to overview

zoom 4, Tuesday, 10:30

Recovery of Sobolev functions restricted to iid sampling

David Krieg
Johannes Kepler Universität Linz, david.krieg@jku.at

Erich Novak
Friedrich-Schiller-Universität Jena, erich.novak@uni-jena.de

Mathias Sonnleitner
Johannes Kepler Universität Linz, mathias.sonnleitner@jku.at

This talk is about reconstructing Sobolev functions on domains using function values at independent and uniformly distributed random points. The domain is only required to satisfy an interior cone condition. We present recent results showing that in many cases the same rate of convergence as with optimally chosen points is attained. In the remaining cases there is a logarithmic loss which cannot be avoided. This is made precise for a uniform and a randomized error criterion. The algorithm obtaining all the presented upper bounds is based on local polynomial approximation and will be sketched briefly. Additionally, consequences for the integration problem are presented.

Special Session Stochastic Computation and Complexity (part C)

Back to overview

zoom 4, Tuesday, 13:30

A review of recent advances of Euler's polygonal approximations

Sotirios Sabanis
University of Edinburgh / Alan Turing Institute, s.sabanis@ed.ac.uk

The idea of using a new form of Euler's polygonal approximations, one which allows coefficients to depend directly on the step size, was highlighted in [1] and [2]. It was used to simplify proofs and extend results on existence of SDEs with monotone coefficients. Since then, it has been applied in several other directions, which include solvability questions for stochastic delay differential equations (SDDEs), see [3], new explicit numerical schemes for SDEs with superlinear coefficients, see e.g. [4], [5] and references therein, as well as in the construction of MCMC algorithms, see e.g. [6] and [7]. More recently, this new form of Euler's polygonal approximations has been used to create new, stochastic (adaptive) optimization algorithms with superior performance, in many cases, than other leading optimization algorithms within the context of fine tuning of artificial neural networks, see [8] and [9]. We will review some key developments of this new methodology, in particular with regards to data science and ML.

[1] N. V. Krylov. *Extremal properties of the solutions of stochastic equations. Theory of Probability and its Applications.* 29(2):205-217, 1985.

[2] N. V. Krylov. *A simple proof of the existence of a solution to the Itô's equation with monotone coefficients. Theory of Probability and its Applications.* 35(3):583–587, 1990.

[3] I. Gyöngy and S. Sabanis. *A Note on Euler Approximations for Stochastic Differential Equations with Delay. Appl. Math. Optim.* 68, 391–412, 2013.

[4] S. Sabanis. *Euler approximations with varying coefficients: the case of superlinearly growing diffusion coefficients. Annals of Applied Probability.* 26(4), 2083–2105, 2016.

[5] S. Sabanis and Y. Zhang. *On explicit order 1.5 approximations with varying coefficients: the case of super-linear diffusion coefficients. Journal of Complexity.* 50, 84–115, 2019.

[6] N. Brosse, A. Durmus, É. Moulines, and S. Sabanis. *The tamed unadjusted Langevin algorithm. Stochastic Processes and their Applications.* 129(10):3638–3663, 2019.

[7] S. Sabanis and Y. Zhang. *Higher order Langevin Monte Carlo algorithm. Electronic Journal of Statistics.* 13(2):3805-3850, 2019.

[8] A. Lovas, I. Lytas, M. Rasonyi, and S. Sabanis. *Taming neural networks with tusla: Non-convex learning via adaptive stochastic gradient langevin algorithms. arXiv preprint arXiv:2006.14514*, 2020.

[9] D.-Y. Lim and S. Sabanis *Polygonal Unadjusted Langevin Algorithms: Creating stable and efficient adaptive algorithms for neural networks. arXiv preprint arXiv:2105.13937*, 2021.

Back to overview

zoom 4, Tuesday, 14:00

A closer look to stochastic sewing and its applications on quadrature approximations

Oleg Butkovsky
Weierstrass Institute, oleg.butkovskiy@gmail.com

Konstantinos Dareiotis
University of Leeds, K.Dareiotis@leeds.ac.uk

Máté Gerencsér
TU Vienna, mate.gerencser@tuwien.ac.at

Khoa Lê
TU Berlin, le@math.tu-berlin.de

Stochastic sewing is a technique which allowed us to obtain sharp results for the rate convergence of numerical approximations of stochastic equations with drifts of poor regularity, in different settings. An overview of these results will be given in the talk of Máté Gerenscér. In the present talk, we will have a closer look to the stochastic sewing lemma and we will try to demonstrate its usefulness in such situations. In particular, we will see how it can be applied in order to obtain quadrature-type estimates, which is one of the two basic steps when analysing the rate of convergence of numerical schemes.

[1] O. Butkovsky, K. Dareiotis, and M. Gerencsér. "Approximation of SDEs–a stochastic sewing approach". arXiv preprint arXiv:1909.07961 (2019), Accepted in *Probability Theory and Related fields*

[2] K. Dareiotis, M. Gerencsér, and K. Lê. "Quantifying a convergence theorem of Gyöngy and Krylov." arXiv preprint arXiv:2101.12185 (2021).

Back to overview

zoom 4, Tuesday, 14:30

On efficient aproximation of solutions of SDEs driven by countably dimensional Wiener process and Poisson random measure

Paweł Przybyłowicz
AGH UST in Cracow, pprzybyl@agh.edu.pl

Michał Sobieraj
AGH UST in Cracow, sobieraj@agh.edu.pl

Łukasz Stępień
AGH UST in Cracow, lstepie@agh.edu.pl

In the talk we refer recent results ([3]) on optimal pointwise approximation of SDEs of the following form

$$\begin{cases} dX(t) = a(t, X(t))dt + b(t, X(t))dW(t) + \int\limits_{\mathcal{E}} c(t, X(t-), y)N(dy, dt), \ t \in [0, T], \\ X(0) = \eta, \end{cases} \tag{6.1}$$

where $T > 0$, $\mathcal{E} = \mathbf{R}^{d'} \setminus \{0\}$, $d' \in \mathbf{N}$, $W = [W_1, W_2, \ldots]^T$ is a countably dimensional Wiener process, and $N(dy, dt)$ is a Poisson random measure with an intensity measure $\nu(dy)dt$ (see [1], [2]). We assume that $\nu(dy)$ is a finite Lévy measure on $(\mathcal{E}, \mathcal{B}(\mathcal{E}))$.

In a certain class of coefficients $a : [0, T] \times \mathbf{R}^d \to \mathbf{R}^d$, $b : [0, T] \times \mathbf{R}^d \to l^2(\mathbf{R}^d)$, $c : [0, T] \times \mathbf{R}^d \to \mathbf{R}^d$ we investigate error of a truncated dimension randomized Euler scheme, which uses evaluations of finite number N of components of the Wiener process W. We establish upper bound on its error in the terms of the discretization parameter n and the truncation parameter N. In suitable subclasses we show also corresponding lower bounds on the error of an arbitrary algorithm that is based on (finite dimensional) evaluations of (a, b, c, W).

At the end we present results of numerical experiments performed on GPUs, where we used a suitable implementation (in CUDA C) of the truncated dimension randomized Euler algorithm.

[1] S. N. Cohen, R. J. Elliott. *Stochastic Calculus and Applications, 2nd. ed.* Springer, 2015.

[2] I. Gyöngy, N. V. Krylov. *On stochastic equations with respect to semimartingales I.* Stochastics, Volume 4, 1980,1–21.

[3] P. Przybyłowicz, M. Sobieraj, Ł. Stępień. *Efficient approximation of SDEs driven by countably dimensional Wiener process and Poisson random measure.* in preparation

Back to overview

zoom 4, Tuesday, 15:30

Error rates of the Euler scheme for pathwise approximation of SDEs with piecewise superlinear drift coefficient

Thomas Müller-Gronbach
University of Passau, thomas.mueller-gronbach@uni-passau.de

Sotirios Sabanis
University of Edinburgh, S.Sabanis@ed.ac.uk

Larisa Yaroslavtseva
University of Passau, larisa.yaroslavtseva@uni-passau.de

We study strong approximation of scalar SDEs on the time-interval $[0,1]$ in the case of a superlinearly growing diffusion coefficient and a piecewise superlinearly growing drift coefficient. We study the performance of a tamed Euler scheme and provide corresponding upper error bounds for approximation at the final time $t = 1$ as well as global approximation on the full time-interval $[0,1]$.

Back to overview

zoom 4, Tuesday, 16:00

The weak convergence order of two Euler-type discretization schemes for the log-Heston model

Annalena Mickel
Universität Mannheim and DFG Research Training Group 1953,
annalena.mickel@math.uni-mannheim.de

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

We study the weak convergence order of two Euler-type discretizations of the log-Heston Model where we use symmetrization and absorption, respectively, to prevent the discretization of the underlying CIR process from becoming negative. If the Feller index ν of the CIR process satisfies $\nu > 1$, we establish weak convergence order one, while for $\nu \leq 1$, we obtain weak convergence order $\nu - \epsilon$ for $\epsilon > 0$ arbitrarily small. Our analysis relies on results from [1, 2, 3]. We illustrate our theoretical findings by several numerical examples. This talk is based on the paper [4].

[1] M. Briani, L. Caramellino and G. Terenzi *Convergence rate of Markov chains and hybrid numerical schemes to jump-diffusions with application to the Bates model.* arXiv:1809.10545v2, 2020.

[2] M. Bossy and A. Diop *Weak convergence analysis of the symmetrized Euler scheme for one dimensional SDEs with diffusion coefficient* $|x|^a$, $a \in [1/2, 1)$. arXiv:1508.04573v1, 2015

[3] A. Cozma and C. Reisinger *Strong order 1/2 convergence of full truncation Euler approximations to the Cox-Ingersoll-Ross process.* IMA Journal of Numerical Analysis, 40(1): 1-19, 2018.

[4] A. Mickel and A. Neuenkirch *The weak convergence order of two Euler-type discretization schemes for the log-Heston model.* arXiv:2106.10926, 2021

Back to overview

zoom 4, Tuesday, 16:30

Optimal L^1-Approximation of the log-Heston SDE by Euler-type methods

Annalena Mickel
University of Mannheim and DFG Research Training Group 1953,
annalena.mickel@math.uni-mannheim.de

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

We study the L^1-approximation of the log-Heston SDE at discrete time points by equidistant Euler-type methods. We establish the convergence order $1/2 - \epsilon$ for $\epsilon > 0$ arbitrarily small, if the Feller index ν of the underlying CIR process satisfies $\nu > 1$. For the L^1-approximation at the final time point this convergence order is optimal, since we also show that arbitrary methods which use an equidistant discretization of the driving Brownian motion can achieve at most order $1/2$ in this case. Moreover, we discuss the case $\nu \leq 1$ and illustrate our findings by several numerical examples.

Back to overview

zoom 4, Wednesday, 09:00

Backward and truncated Euler–Maruyama schemes for radial Dunkl processes

Hoang-Long Ngo
Hanoi National University of Education, ngolong@hnue.edu.vn

Dai Taguchi
Okayama University, dai.taguchi.dai@gmail.com

In this talk, we consider the numerical approximation for a class of radial Dunkl processes corresponding to arbitrary (reduced) root systems. This class contains some well-known processes such as Bessel processes and Dyson's Brownian motions (see [1]). We introduce a backward and truncated Euler–Maruyama scheme, which can be implemented on a computer, and study its rate of convergece in L^p-norm.

[1] N. Demni. *Radial Dunkl processes: existence, uniqueness and hitting time.* C. R. Math. Acad. Sci. Paris, Ser. I **347** 1125–1128, (2009).

Back to overview

zoom 4, Wednesday, 09:30

The random periodic solution of a SDE with a monotone drift and its numerical approximation

Yue Wu
University of Oxford, yue.wu@maths.ox.ac.uk

Periodicity is widely exhibited in a large number of natural phenomena like oscillations, waves, or even lying behind many complicated ensembles such as biological and economic systems. However, periodic behaviours are often found to be subject to random perturbation or under the influence of noise. It was only until recently that the random periodic solution was endowed with a proper definition (c.f. Zhao and Zheng [4], Feng, Zhao and Zhou [1]), which is compatible with definitions of both the stationary solution (also termed as random fixed points) and the deterministic periodic solution.

In general, random periodic solutions cannot be solved explicitly. One may treat the numerical approximation that stay sufficient close to the true solution as a good substitute to study stochastic dynamics. It is worth mentioning here that this is a numerical approximation of an infinite time horizon problem. The classical numerical approaches including the Euler-Marymaya method and a modified Milstein method to simulate random period solutions of a dissipative system with global Lipchitz condition have been investigated in [2], the first paper that numerical schemes were used to approximate the random period trajectory.

We study the existence and uniqueness of the random periodic solution for a stochastic differential equation with a one-sided Lipschitz condition (also known as monotonicity condition) and the convergence of its numerical approximation via the backward Euler-Maruyama method [3]. The existence of the random periodic solutions are shown as the limits of the pull-back flows of the SDE and discretized

SDE respectively. We establish a convergence rate of the strong error for the backward Euler-Maruyama method and obtain the weak convergence result for the approximation of the periodic measure.

[1] C. R. Feng, H. Z. Zhao and B. Zhou, Pathwise random periodic solutions of stochastic differential equations, *Journal of Differential Equations,* Vol.251(2011), 119-149.

[2] C. R. Feng, Y. Liu and H. Z. Zhao, Numerical approximation of random periodic solutions of stochastic differential equations, *Zeitschrift fÃ¼r angewandte Mathematik und Physik,* 68.5(2017), 1-32.

[3] Y.Wu, The random periodic solution of a stochastic differential equation with a monotone drift and its numerical approximation *arXiv Preprint.* arXiv:2105.13477.

[4] H. Z. Zhao and Z. H. Zheng, Random periodic solutions of random dynamical systems, *J. Differential Equations,* Vol.248(2009), 2020-2038.

Back to overview

zoom 4, Wednesday, 10:00

Irregular drift and stochastic sewing in numerical analysis of SDEs

Oleg Butkovsky
WIAS Berlin, oleg.butkovskiy@gmail.com

Konstantinos Dareiotis
University of Leeds, k.dareiotis@leeds.ac.uk

Máté Gerencsér
TU Wien, mate.gerencser@tuwien.ac.at

Khoa Lê
TU Berlin, le@math.tu-berlin.de

We give an overview of the stochastic sewing approach in the numerical analytic context. We discuss how novel developments in the regularisation by noise area can be leveraged to derive sharp convergence rates of very simple numerical algorithms for stochastic equations with irregular drift, and survey a number of application for SDEs with Hölder and Sobolev drift coefficients, non-Markovian SDEs, Lévy-driven equations, SPDEs, and central limit theorems.

[1] O. Butkovsky, K. Dareiotis, M. Gerencsér *Approximation of SDEs – a stochastic sewing approach.* arXiv:1909.07961.

[2] K. Dareiotis, M. Gerencsér, K. Lê *Quantifying a convergence theorem of Gyöngy and Krylov* arXiv:2101.12185

Back to overview

zoom 4, Wednesday, 10:30

A deep neural network algorithm for semilinear elliptic PDEs with applications in insurance mathematics

Stefan Kremsner
University of Graz, stefan.kremsner@uni-graz.at

Alexander Steinicke
Montanuniversitaet Leoben, alexander.steinicke@unileoben.ac.at

Michaela Szölgyenyi
Universität Klagenfurt, Michaela.Szoelgyenyi@aau.at

In insurance mathematics, optimal control problems over an infinite time horizon arise when computing risk measures. An example of such a risk measure is expected discounted future dividend payments. In models which take multiple economic factors into account, this problem is high-dimensional. The solutions to such control problems correspond to solutions of deterministic semilinear (degenerate) elliptic PDEs. In the first part of the talk we give an introductory example and show this associated PDE.

Then we discuss the correspondence of elliptic PDEs to BSDEs with unbounded random terminal time. Finally we propose a deep neural network algorithm based on the BSDE-PDE correspondence for solving control problems in complex high-dimensional economic environments.

Special Session Stochastic Computation and Complexity (part E)

Back to overview

zoom 4, Wednesday, 13:30

Numerical methods for mean-field control problems

Wolfgang Stockinger
University of Oxford, wolfgang.stockinger@maths.ox.ac.uk

Christoph Reisinger
University of Oxford, christoph.reisinger@maths.ox.ac.uk

Yufei Zhang
University of Oxford, yufei.zhang@maths.ox.ac.uk

In this talk, we discuss numerical methods for mean field control (MFC) problems, which seek optimal control of McKean–Vlasov dynamics whose coefficients involve mean-field interactions both on the state and actions, and where objectives are optimized over open-loop strategies. Due to the non-Markovian nature of McKean-Vlasov SDEs, it is generally preferred to approach the MFC problem by employing a version of the Pontryagin stochastic maximum principle, rather than the Hamilton–Jacobi–Bellman (HJB) equations, which require to consider the enlarged (infinite-dimensional) state space $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. Many existing numerical methods assume that the first order optimality condition, arising from the stochastic maximum principle, can be solved explicitly; i.e., it is possible to find the optimal feedback map in terms of the adjoint equations, which then allows to rewrite the control problem as coupled mean-field forward backward SDE. However, in many cases this is not possible, in particular if the coefficients of the controlled dynamics or the running cost depend on the law of the control or if the cost is nonsmooth (e.g., ℓ_1-regularization).

At the core of our algorithm, we iteratively refine the optimal feedback control by applying an accelerated proximal gradient method to the MFC problem, which does not require to solve the first order optimality explicitly. The fully implementable procedure is a Markovian hybrid, i.e., Monte-Carlo and PDE based, algorithm, which outputs the optimal control in feedback form and solves the control problem along a flow of measures. Thereby, we improve the robustness of the numerically obtained feedback control in terms of model perturbation compared to pure sampling based algorithms. We also propose a version of our method based on a deep neural network approach which is applicable for high-dimensional control problems and will be illustrated for the well-known Cucker-Smale model.

[1] C. Reisinger, W. Stockinger, and Y. Zhang. *Optimal regularity of extended mean field controls and their piecewise constant approximation.* arXiv:2009.08175v2, 2020.

Back to overview

zoom 4, Wednesday, 14:00

Milstein MLMC for delay McKean–Vlasov equations

Jianhai Bao
Tianjin University, jianhaibao13@gmail.com

Christoph Reisinger
University of Oxford, christoph.reisinger@maths.ox.ac.uk

Panpan Ren
City University of Hong Kong, rppzoe@gmail.com

Wolfgang Stockinger
University of Oxford, wolfgang.stockinger@maths.ox.ac.uk

In this talk, we first derive Milstein schemes for an interacting particle system associated with point delay McKean–Vlasov stochastic differential equations, possibly with a drift term exhibiting super-linear growth in the state component. We prove strong convergence of order one and moment stability, making use of techniques from variational calculus on the space of probability measures with finite second order moments. Then, we introduce an antithetic multi-level Milstein scheme, which leads to optimal complexity estimators for expected functionals of solutions to certain delay McKean-Vlasov equations without the need to simulate Lévy areas.

[1] J. Bao, C. Reisinger, P. Ren, and W. Stockinger. *First order convergence of Milstein schemes for McKean-Vlasov equations and interacting particle systems.* Proceedings of the Royal Society A, 477(2245), 2021.

[2] J. Bao, C. Reisinger, P. Ren, and W. Stockinger. *Milstein schemes for delay McKean equations and interacting particle systems.* arXiv:2005.01165.

Back to overview

zoom 4, Wednesday, 14:30

The p-variation of Stochastic Evolution Equations and Applications in Numerical Analysis

Raphael Kruse
Martin-Luther-Universität Halle-Wittenberg, raphael.kruse@mathematik.uni-halle.de

Johanna Weinberger
Martin-Luther-Universität Halle-Wittenberg, johanna.weinberger@mathematik.uni-halle.de

Rico Weiske
Martin-Luther-Universität Halle-Wittenberg, rico.weiske@mathematik.uni-halle.de

In this talk we investigate the p-variation of the solution to a stochastic evolution equation. As it is known, if the solution is Hölder continuous with exponent $\gamma \in (0,1)$ then it is also of finite p-variation for any $p \geq \frac{1}{\gamma}$. However, if the initial condition is not smooth, then the exact solution is, in general, only locally Hölder continuous. In such a situation the p-variation semi-norm might serve as a better measure of the temporal regularity of the exact solution. In the second part of the talk, we indicate how to make use of the p-variation semi-norm in the numerical analysis of stochastic evolution equations.

Special Session Quantitative Aspects of SDEs and Applications

Back to overview

zoom 4, Thursday, 13:30

Efficient simulation of generalized Whittle-Matérn fields

Sonja Cox
Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Netherlands, s.g.cox@uva.nl

Kristin Kirchner
Delft Institute of Applied Mathematics, Delft University of Technology, Netherlands, k.kirchner@tudelft.nl

Non-stationary Gaussian random fields can be modelled by assuming that covariance operator to be given by a negative fractional power of a second-order elliptic differential operator of the form $L := -\nabla \cdot (A\nabla) + \kappa^2$, where A and κ are space-dependent. Such fields form a generalization of the (stationary) Whittle-Matérn fields, where A and κ are constants. In our work we established optimal strong convergence rates in Hölder and Sobolev norms for Galerkin approximations of these generalized Whittle-Matérn random fields. More specifically, we considered both spectral Galerkin methods and finite element methods. The latter, although significantly more tedious to analyse, are more suitable for non-stationary fields on non-standard domains.

Back to overview

zoom 4, Thursday, 14:00

Continuous in time tug-of-war with noise in irregular domains

Stefan Geiss
Department of Mathematics and Statistics, University of Jyväskylä, Finland, stefan.geiss@jyu.fi

Christel Geiss
Department of Mathematics and Statistics, University of Jyväskylä, Finland, christel.geiss@jyu.fi

Mikko Parviainen
Department of Mathematics and Statistics, University of Jyväskylä, Finland,
mikko.parviainen@jyu.fi

In this paper, we define a continuous time version of the tug-of-war with noise in a bounded domain. We show that for a suitably defined game, the value coincides with the Perron solution (in the sense of PDE theory) to the p-Laplace equation without regularity assumptions on the domain.

Back to overview

zoom 4, Thursday, 14:30

Federated mean-field stochastic control with common noise of numerous heterogeneous energy storage systems

Emmanuel Gobet
CMAP - Ecole Polytechnique, 91128 Palaiseau cedex, France, emmanuel.gobet@polytechnique.edu

Maxime Grangereau
,

We propose a stochastic control problem to control cooperatively Thermostatically Controlled Loads (TCLs) to promote power balance in electricity networks. We develop a method to solve this stochastic control problem with a decentralized architecture, in order to respect privacy of individual users and to reduce both the telecommunications and the computational burden compared to the setting of an omniscient central planner. This paradigm is called federated learning in the machine learning community, therefore we refer to this problem as a federated stochastic control problem. The optimality conditions are expressed in the form of a high-dimensional Forward-Backward Stochastic Differential Equation (FBSDE), which is decomposed into smaller FBSDEs modeling the optimal behaviors of the aggregate population of TCLs of individual agents. In particular, we show that these FBSDEs fully characterize the Nash equilibrium of a stochastic Stackelberg differential game. In this game, a coordinator (the leader) aims at controlling the aggregate behavior of the population, by sending appropriate signals, and agents (the followers) respond to this signal by optimizing their storage system locally. A mean-field-type approximation is proposed to circumvent telecommunication constraints and privacy issues. Convergence results and error bounds are obtained for this approximation depending on the size of the population of TCLs. A numerical illustration is provided to show the interest of the control scheme and to exhibit the convergence of the approximation. An implementation which answers practical industrial challenges to deploy such a scheme is presented and discussed.

Back to overview

zoom 4, Thursday, 15:30

Uniform in time estimates for the weak error of the Euler method for SDEs and a Pathwise Approach to Derivative Estimates for Diffusion Semigroups

Dan Crisan
Department of Mathematics, Imperial College London, London, UK, d.crisan@imperial.ac.uk

Paul Dobson
Delft Institute of Applied Mathematics, Delft University of Technology, The Netherlands. ,
p.dobson@tudelft.nl

Michela Ottobre
Department of Mathematics, Heriot-Watt University, Edinburgh EH14 4AS, United Kingdom,
m.ottobre@hw.ac.uk

I will present a criterion for uniform in time convergence of the weak error of the Euler scheme for Stochastic Differential equations (SDEs). The criterion requires i) exponential decay in time of the space-derivatives of the semigroup associated with the SDE and ii) bounds on (some) moments of the Euler approximation. I will give examples (and counterexamples) how both i) and ii) are needed to obtain the desired result. If the weak error converges to zero uniformly in time, then convergence of ergodic averages follows as well. I will also show that Lyapunov-type conditions are neither sufficient nor necessary in order for the weak error of the Euler approximation to converge uniformly in time and clarify relations between the validity of Lyapunov conditions, i) and ii).

The study of derivative estimates has attracted a lot of attention, however not many results are known in order to guarantee exponentially fast decay of the derivatives. Exponential decay of derivatives typically follows from coercive-type conditions involving the vector fields appearing in the equation and their commutators; here we focus on the case in which such coercive-type conditions are non-uniform in space. To obtain results under such space-inhomogeneous conditions we initiate a pathwise approach to the study of derivative estimates for diffusion semigroups.

The talk is based on joint work with Paul Dobson and Michela Ottobre and is baset on the paper [1].

[1] Crisan, D.; Dobson, P.; Ottobre, M.; Uniform in time estimates for the weak error of the Euler method for SDEs and a pathwise approach to derivative estimates for diffusion semigroups. Trans. Amer. Math. Soc. 374 (2021), no. 5, 3289-3330.

Back to overview

zoom 4, Thursday, 16:00

Convergence rate of random walk approximation of Backward SDEs

Philippe Briand
LAMA, Université Savoie Mont Blanc, France, philippe.briand@univ-smb.fr

Christel Geiss
Department of Mathematics and Statistics, University of Jyväskylä, Finland, christel.geiss@jyu.fi

Stefan Geiss
Department of Mathematics and Statistics, University of Jyväskylä, Finland, stefan.geiss@jyu.fi

Céline Labart
LAMA, Université Savoie Mont Blanc, France, celine.labart@univ-smb.fr

Antti Luoto
Elo, `Finland`

Briand, Delyon and Mémin have shown in [1] a Donsker-type theorem for forward-BSDE. If one approximates the Brownian motion B by a random walk B^n, the according solutions (X^n, Y^n, Z^n) converges weakly to (X, Y, Z). We investigate under which conditions (Y_t^n, Z_t^n) converges to (Y_t, Z_t) in L_2 and compute the rate of convergence in dependence of smoothness properties of the coefficients b and sigma, the terminal condition and the generator. The talk is based on joint results with Philippe Briand, Christel Geiss, Stefan Geiss and Antti Luoto.

[1] Philippe Briand, Bernard Delyon, Jean Mémin. Donsker-type theorem for BSDEs. Electron. Comun. in Probab., 6:1-14, 2001.

Back to overview

zoom 4, Thursday, 16:30

Approximation and error analysis of forward-backward SDEs driven by pure jump Lévy processes using shot noise series representations

Till Massing
Universität Duisburg-Essen, Germany, `till.massing@uni-due.de`

We consider the simulation of a system of decoupled forward-backward stochastic differential equations (FBSDEs) driven by a pure jump Lévy process $\{L_t\}$ and an independent Brownian motion $\{B_t\}$. We allow the Lévy process $\{L_t\}$ to have an infinite jump activity. Therefore, it is necessary for the simulation to employ a finite approximation of its Lévy measure. We use the general shot noise series representation method by [2] to approximate the driving Lévy process $\{L_t\}$. Going beyond [1], we compute the L^p error between the true and the approximated FBSDEs which arises by a finite truncation of the shot noise series (given sufficient conditions for existence and uniqueness of the FBSDE). In the next step, we derive the error between the true solution and the discretization of the approximated FBSDE using an appropriate backward Euler scheme.

[1] Soufiane Aazizi. Discrete-time approximation of decoupled forward-backward stochastic differential equations driven by pure jump Lévy Processes. Advances in Applied Probability, 45(3):791-821, 2013.

[2] Jan Rosinski. Series Representations of Lévy processes from the perspective of point processes, pages 401-415. Birkhäuser Boston, Boston, MA, 2001.

Back to overview

zoom 4, Thursday, 09:30

A derivative-free scheme of order up to 1.5 for SPDEs of Nemytskii-tpye

Claudine von Hallern
Universität Hamburg, claudine.von.hallern@uni-hamburg.de

Ricarda Mißfeldt
Universität zu Lübeck, missfeldt@math.uni-luebeck.de

Andreas Rößler
Universität zu Lübeck, roessler@math.uni-luebeck.de

We propose a numerical scheme that is free of derivatives and achieves an order of temporal convergence of up to $1.5 - \varepsilon$ for some $\varepsilon > 0$. The proposed scheme is tailored to stochastic partial differential equations of Nemytskii-type and is given as a family of exponential stochastic Runge-Kutta type schemes. It is based on the exponential Wagner-Platen type scheme introduced in [1] but involves lower computational cost as no derivatives have to be computed. In this talk, we introduce and discuss the scheme and its convergence properties. Moreover, we look at some examples and illustrate our findings with simulations.

[1] S. Becker, A. Jentzen, A. and P. E. Kloeden. *An exponential Wagner-Platen type scheme for SPDEs.* SIAM J. Numer. Anal., 2016

Back to overview

zoom 4, Thursday, 10:00

Strong Convergence Rates for Space-Time Discrete Numerical Approximation Schemes for Stochastic Burgers Equations

Martin Hutzenthaler
University of Duisburg-Essen, martin.hutzenthaler@uni-due.de

Arnulf Jentzen
University of Münster, ajentzen@uni-muenster.de

Felix Lindner
University of Kassel, lindner@mathematik.uni-kassel.de

Primož Pušnik
ETH Zürich, primoz.pusnik@sam.math.ethz.ch

The main result presented in this talk establishes strong convergence rates on the whole probability space for explicit space-time discrete numerical approximations for a class of stochastic evolution equations with possibly non-globally monotone coefficients such as stochastic Burgers equations with additive trace-class noise. The key idea in the proof of our main result is (i) to bring the classical Alekseev-Gröbner formula from deterministic analysis into play and (ii) to employ uniform exponential moment estimates for the numerical approximations.

Back to overview

zoom 4, Thursday, 10:30

Numerical approximation of solutions of stochastic Allen-Cahn equations

Ludovic Goudenège
CNRS-CentraleSupélec-Université Paris-Saclay, goudenege@math.cnrs.fr

Charles-Édouard Bréhier
CNRS-ICJ-Université Lyon 1, brehier@math.univ-lyon1.fr

We consider a class of various Euler schemes in time (tamed, explicit-implicit, exponential) applied to stochastic partial differential equations of Allen-Cahn type driven by additive noise. This type of semi-linear parabolic SPDEs is characterized by technical difficulties due to drift nonlinearities with polynomial growth. Still, they possess a one-sided Lipschitz continuity property that permits control of the solutions.

As a classical approach for these semi-discretized schemes, the first step is to obtain the moment bounds for the approximated solutions obtained by numerical schemes in finite time (see [1, 2] for results using splitting schemes). In a second step, we obtain the error estimates in terms of the time-step size.

We compare the numerical approximated solutions to a continuous solution at a fixed time and in a weak sense. It is natural to consider this type of error in law, which is compatible with statistical estimation. With at most polynomial growing with respect to the time horizon of the moment bounds, we can hope to study the long-time behavior, and in particular, the approximation of the invariant distribution, in the spirit of [3]. Thus, when the long-time behavior is reachable, we quantify the error to approximate averages with respect to the invariant distribution.

Finally,we will illustrate the effectiveness of each scheme, from a theoretical point of view, concerning the realized weak order, and from a practical point of view, concerning the computational cost.

[1] C.-E. Bréhier, and L. Goudenège. *Weak convergence rates of splitting schemes for the stochastic Allen-Cahn equation.* Bit Numer Math, Volume 60, 543–582, 2020.

[2] C.-E. Bréhier, and L. Goudenège. *Analysis of some splitting schemes for the stochastic Allen-Cahn equation.* DCDS Series B, Volume 24(8), 2019.

[3] C.-E. Bréhier. *Approximation of the invariant distribution for a class of ergodic SPDEs using an explicit tamed exponential Euler scheme.* ArXiv:2010.00508.

Special Session SPDEs Simulation and Modelling (part B)

Back to overview

zoom 4, Friday, 09:00

Sparse grid MLMC for Zakai equations

Christoph Reisinger
University of Oxford, christoph.reisinger@maths.ox.ac.uk

Zhenru Wang
University of Oxford, zhenru.wang@maths.ox.ac.uk

In this talk, we analyse the accuracy and computational complexity of estimators for expected functionals of the solution to multi-dimensional parabolic stochastic partial differential equations (SPDE) of Zakai-type. Here, we use the Milstein scheme for time integration and an alternating direction implicit (ADI) splitting of the spatial finite difference discretisation, coupled with the sparse grid combination technique and multilevel Monte Carlo sampling (MLMC). In the two-dimensional case, we find by detailed Fourier analysis that for a root-mean-square error (RMSE) ε, MLMC on sparse grids has the optimal complexity $O(\varepsilon^{-2})$, whereas MLMC on regular grids has $O(\varepsilon^{-2}(\log\varepsilon)^2)$, standard MC on sparse grids $O(\varepsilon^{-7/2}(|\log\varepsilon|)^{5/2})$, and MC on regular grids $O(\varepsilon^{-4})$. Numerical tests confirm these findings empirically. We give a discussion of the higher-dimensional setting without detailed proofs, which suggests that MLMC on sparse grids always leads to the optimal complexity, standard MC on sparse grids has a fixed complexity order independent of the dimension (up to a logarithmic term), whereas the cost of MLMC and MC on regular grids increases exponentially with the dimension.

[1] C. Reisinger and Z. Wang. *Stability and error analysis of an implicit Milstein finite difference scheme for a two-dimensional Zakai SPDE.* BIT Numerical Mathematics, 59, 987–1029, 2019.

[2] C. Reisinger and Z. Wang. *Analysis of sparse grid multilevel estimators for multi-dimensional Zakai equations.* To appear in Sparse Grids and Applications, H.-J. Bungartz et al. (Eds), Springer Lectures Notes in Computational Science and Engineering, 2021.

Back to overview

zoom 4, Friday, 09:30

Connecting random fields on manifolds and stochastic partial differential equations in simulations

Annika Lang
Chalmers & University of Gothenburg, annika.lang@chalmers.se

Random fields on manifolds can be used as building blocks for solutions to stochastic partial differential equations or they can be described by stochastic partial differential equations. In this talk I present recent developments in numerical approximations of random fields and solutions to stochastic partial differential equations on manifolds and connect the two. More specifically, we look at the stochastic wave equation on the sphere and approximations of Gaussian random fields on manifolds using suitable finite element methods. Throughout the talk, theory and convergence analysis are combined with numerical examples and simulations.

This talk is based on joint work with David Cohen, Erik Jansson, Mihály Kovács, and Mike Pereira.

Back to overview

zoom 4, Friday, 10:00

Gaussian random fields defined through SPDEs: Regularity, simulation and applications in statistics

David Bolin
King Abdullah University of Science and Technology, david.bolin@kaust.edu.sa

Sonja G. Cox
University of Amsterdam, s.g.cox@uva.nl

Kristin Kirchner
Delft University of Technology, k.kirchner@tudelft.nl

Mihály Kovács
Pázmány Péter Catholic University, Budapest, kovacs.mihaly@itk.ppke.hu

Many models in spatial statistics are based on Gaussian random fields (GRFs). Motivated by the relation between the Matérn class of GRFs and stochastic partial differential equations (SPDEs), we consider GRFs as solutions of fractional-order elliptic SPDEs driven by additive spatial white noise on a bounded Euclidean domain. We discuss the regularity of these GRFs in Sobolev and Hölder spaces and propose a numerical approximation which we show to converge in these spaces at explicit and sharp rates. We furthermore discuss the computational benefits of the proposed method compared to kernel-based approaches, illustrated by several numerical experiments and explicit applications in statistics.

This talk will be based on results of the articles [1, 2, 3, 4].

[1] D. Bolin and K. Kirchner. The rational SPDE approach for Gaussian random fields with general smoothness. *J. Comp. Graph. Stat.*, 29(2):274–285, 2020.

[2] D. Bolin, K. Kirchner, and M. Kovács. Weak convergence of Galerkin approximations for fractional elliptic stochastic PDEs with spatial white noise. *BIT*, 58(4):881–906, 2018.

[3] D. Bolin, K. Kirchner, and M. Kovács. Numerical solution of fractional elliptic stochastic PDEs with spatial white noise. *IMA J. Numer. Anal.*, 40(2):1051–1073, 2020.

[4] S. G. Cox and K. Kirchner. Regularity and convergence analysis in Sobolev and Hölder spaces for generalized Whittle–Matérn fields. *Numer. Math.*, 146(4):819–873, 2020.

Back to overview

zoom 4, Friday, 10:30

The long time approximation for SDEs driven by a Levy process

Erika Hausenblas
Montanuniversität Leoben, erika.hausenblas@unileoben.ac.at

In the talk we will speak about the numerical approximation of sdes with jumps. Here, the main focus is the long term approximation. We will compare different strategies and present some error estimates for the weak convergence.

6.5 Zoom Channel 5

back to table of contents

Clicking on the name of the speaker will lead you to the abstract. NOTE THAT ALL TIMES ARE UTC+2.

	Monday	Tuesday	Wednesday	Thursday	Friday
				Special Session **Uncertainty Quantification for hyperbolic PDEs**	Contributed talks on **Physics and Engineering**
09:00				Alexandra Würth	Bert Mortier
09:30				Elisa Iacomini	Chi-Ok Hwang
10:00				Thomas Schillinger	Emil Løvbak
10:30				Michael Schuster	Philippe Blondeel
		Contributed talks on **Sampling** (part A)	Contributed talks on **PDMPs and ODEs**		
13:30					
14:00		Hozumi Morohosi	Andrea Bertazzi		
14:30		Fan Yang	Sebastiano Grazzi		
15:30		Piergiacomo Sabino	Natalia Czyzewska		
16:00		Amparo Gil	Flavius Guiaş		
16:30					
		Contributed talks on **Sampling** (part B)	Special Session **Wasserstein gradient flows and their applications**	Contributed talks on **Optimization**	
19:00		Yiran Chen	Pierre Monmarché	Kamélia Daudel	
19:30		Mark Huber	Wuchen Li	Mathieu Gerber	
20:00		Chaofan Huang	Katy Craig	Sifan Liu	
20:30		Alexander Shkolnik	Nicolás García Trillos		

Back to overview

zoom 5, Tuesday, 14:00

Empirical Bernstein copula random number generation via mixture model

Hozumi Morohosi
National Graduate Institute for Policy Studies, morohosi@grips.ac.jp

This work is a computational study of generating random numbers which follow an empirical Bernstein copula constructed from a sample. Empirical Bernstein copula [2] is devised as a nonparametric method to describe the dependent relation between random variables. Its simple structure helps us to generate random numbers from it by comparing it to a mixture model. We will propose a simple algorithm to generate random numbers following an empirical Bernstein copula based on some consideration of the relation between empirical Bernstein copula and kernel density estimation. We also report some numerical experiment results, which include a goodness-of-fit test, parametric bootstrap, time series modeling [1].

[1] M. Hofert, I. Kojadinovic, M. Mächler, and J. Yan. *Elements of Copula Modeling with R*, Springer, 2018.

[2] A. Sancetta and S. Satchell. The Bernstein copula and its applications to modeling and approximations of multivariate distributions. *Econometric Theory*, 20(2004), 535–562.

Back to overview

zoom 5, Tuesday, 14:30

Monte Carlo simulation of rough volatility models

Alexandre Richard
Université Paris-Saclay, CentraleSupélec, MICS and CNRS FR-3487,
alexandre.richard@centralesupelec.fr

Xiaolu Tan
The Chinese University of Hong Kong, xiaolu.tan@cuhk.edu.hk

Fan Yang
The Chinese University of Hong Kong, fyang@math.cuhk.edu.hk

We study the Monte Carlo simulation problem of some rough volatility models in mathematical finance, which are given by the stochastic Volterra equation. We first consider the Volterra equation with Lipschitz coefficient functions. By using and adapting some results from Zhang (2008), together with the Garsia-Rodemich-Rumsey lemma, we obtain the convergence rates of the Euler scheme and Milstein scheme under the supremum norm. We then apply these schemes to approximate the expectation of functionals of such Volterra equations by the (Multilevel) Monte-Carlo method, and analyse their complexity. Next, we consider the rough Heston model, whose volatility coefficient function has a non-Lipschitz term $\sqrt{x}$. Using weak convergence techniques in Abi Jaber, Cuchiero, Larsson and Pulido (2019), we obtain a convergence result of the discrete time scheme. Finally, we provide some numerical examples of different simulation methods.

[1] X. Zhang *Euler schemes and large deviations for stochastic Volterra equations with singular kernels.* Journal of Differential Equations, 244(9):2226–2250, 2008.

[2] E. Abi Jaber, C. Cuchiero, M. Larsson, and S. Pulido *A weak solution theory for stochastic Volterra equations of convolution type.* ArXiv preprint arXiv:1909.01166, 2019.

Back to overview

zoom 5, Tuesday, 15:30

Exact Simulation of Tempered Stable Distributions and Finite Variation Ornstein-Uhlenbeck processes

Piergiacomo Sabino
Quantitative Risk Management, E.ON SE[1], piergiacomo.sabino@eon.com

The energy and commodity markets exhibit a strong mean-reversion and sudden spikes which makes the use of the Lévy-driven Ornstein-Uhlenbeck processes more advisable than the standard Gaussian framework. Constructing Lévy processes of Ornstein-Uhlenbeck type is a task closely related to the notion of self-decomposability (see Barndorff-Nielsen and Shephard [1] and Sato [5]). In particular, their transition laws are linked to the properties of what we call the *a-remainder* of their self-decomposable stationary laws. In the present study we fully characterize the Lévy triplet of these *a-remainder*'s and we provide a general framework to deduce the transition laws of the finite variation Ornstein-Uhlenbeck processes associated with tempered stable distributions (see Grabchak [2]). We focus finally on the subclass of the exponentially-modulated tempered stable laws (see Rosiński [4]) and we derive the algorithms for an exact generation of the skeleton of Ornstein-Uhlenbeck processes related to such distributions, with the further advantage of adopting a procedure computationally more efficient than those already available in the existing literature (see Qu et al [3] and Zhang [6]).

[1] O.E. Barndorff-Nielsen and N. Shephard. Non-Gaussian Ornstein-Uhlenbeck-based Models and some of their Uses in Financial Economics. *Journal of the Royal Statistical Society: Series B*, 63(2):167–241, 2001.

[2] M. Grabchak. On the Simulation of General Tempered Stable Ornstein-Uhlenbeck Processes. *Journal of Statistical Computation and Simulation*, 90(6):1057–1081, 2020.

[3] Y. Qu, A. Dassios, and H. Zhao. Exact Simulation of Ornstein-Uhlenbeck Tempered Stable Processes. *Journal of Applied Probability*, 0(0), 2021. Forthcoming.

[4] J. Rosinski. Tempering Stable Proceses. *Stochastic Processes and their Applications*, 117(6):677 – 707, 2007.

[5] K. Sato. *Lévy Processes and Infinitely Divisible Distributions.* Cambridge U.P., Cambridge, 1999.

[6] S. Zhang. Exact Simulation of Tempered Stable Ornstein-Uhlenbeck Proceses. *Journal of Statistical Computation and Simulation*, 81(11):1533–1544, 2011.

Back to overview

[1]The views, opinions, positions or strategies expressed in this article are those of the authors and do not necessarily represent the views, opinions, positions or strategies of, and should not be attributed to E.ON SE.

zoom 5, Tuesday, 16:00

Computation and inversion of the Student's-t distribution

Amparo Gil
Departmento de Matemática Aplicada y CC de la Comput, Universidad de Cantabria, Spain,
amparo.gil@unican.es

Javier Segura
Departmento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Spain,
javier.segura@unican.es

Nico M. Temme
CWI, The Netherlands, nico.temme@cwi.nl

There is a very close relationship between some special functions and some of the most popular distribution functions in statistics. For example, the incomplete beta function is the central beta distribution. Particular cases include other well-known distributions such as the geometric, binomial, negative binomial or the central Student's-t distribution. Therefore, standard methods for the computation and inversion of special functions [1] can also be applied to evaluate and invert distribution functions. The problem of inversion appears, for example, when computing percentage points of the distribution functions; also, it is closely related to the generation of random variates from a continuous probability density function needed, for example, in Monte Carlo or quasi-Monte Carlo methods. In [2] and [3] we have considered the inversion of the central beta distribution and the binomial/negative binomial distributions. In this talk, we describe recent work on the computation and inversion of the ubiquitous Student's-t distribution. We illustrate the performance of all the obtained approximations with numerical examples.

[1] A. Gil, J. Segura, N.M. Temme. Numerical methods for special functions. *SIAM*, 2007.

[2] A. Gil, J. Segura, N.M. Temme. Efficient algorithms for the inversion of the cumulative central beta distribution. *Numer. Algorithms*, 74(1): 77–91, 2017.

[3] A. Gil, J. Segura, N.M. Temme. Asymptotic inversion of the binomial and negative binomial cumulative distribution functions. *Electron. Trans. Numer. Anal.*, 52: 270-280, 2020.

Contributed talks on Sampling (part B)

Back to overview

zoom 5, Tuesday, 19:00

Monte Carlo and Quasi-Monte Carlo Simulation of Copulas

Yiran Chen
Florida State University, ychen3@math.fsu.edu

Giray Ökten
Florida State University, okten@math.fsu.edu

Multivariate models with dependent variables are popular in financial industry. Simulation of copulas can be done by Monte Carlo (MC) methods or quasi-Monte Carlo (QMC) methods. Goodness-of-fit tests can be used to find the best simulation algorithms for copulas. We introduce a new goodness-of-fit test based on the collision test, Voronoi diagrams and low-discrepancy sequences, and present numerical results on option pricing and VaR estimation via copula models. The numerical results show the improvements of QMC simulations of copulas over MC simulations, and the advantages of the new goodness-of-fit test for copulas for large samples.

Back to overview

zoom 5, Tuesday, 19:30

Secant bounds on light tailed discrete distributions

Mark Huber
Claremont McKenna College, Mark.Huber@claremontmckenna.edu

Siki Wang
Claremont McKenna College, SWang21@students.claremontmckenna.edu

Bounding the tails of discrete distributions such as Poisson, Binomial, and Negative Binomial arises in many contexts. For instance, Poisson tails are used in the general TPA algorithm for automatic continuously self-reducible problems, and estimation of bounded nonnegative random variables with unknown variance. Typical approaches to bounding these tails are *geometric* and *normal* approaches. This work introduces a third *secant* approach. This new approach is asymptotically closest to the geometric approach, but always results in a strictly better bound. The normal approach turns out to be better in low relative error regimes, with a crossover point where the secant dominates. Evaluation of these bounds from theorertical and experimental perspectives will be presented.

Back to overview

zoom 5, Tuesday, 20:00

Population Quasi-Monte Carlo

Chaofan Huang
Gerogia Institue of Technology, chuang397@gatech.edu

V. Roshan Joseph
Gerogia Institue of Technology, roshan@gatech.edu

Simon Mak
Duke University, sm769@duke.edu

Monte Carlo methods are widely used for approximating complicated, multidimensional integrals for Bayesian inference. Population Monte Carlo (PMC) is an important class of Monte Carlo methods, which utilizes a population of proposals to generate weighted samples that approximate the target distribution. The generic PMC framework iterates over three steps: samples are simulated from a set of proposals, weights are assigned to such samples to correct for mismatch between the proposal and target distributions, and the proposals are then adapted via resampling from the weighted samples. When the target distribution is expensive to evaluate, the PMC has its computational limitation since the convergence rate is $\mathcal{O}(N^{-1/2})$. To address this, we propose in this paper a new *Population Quasi-Monte Carlo* (PQMC) framework, which integrates Quasi-Monte Carlo ideas within the sampling and adaptation steps of PMC. A key novelty in PQMC is the idea of importance support points resampling, a deterministic method for

finding an "optimal" subsample from the weighted proposal samples. Moreover, within the PQMC framework, we develop an efficient covariance adaptation strategy for multivariate normal proposals. Lastly, a new set of correction weights is introduced for the weighted PMC estimator to improve the efficiency from the standard PMC estimator. We demonstrate the improved empirical convergence of PQMC over PMC in extensive numerical simulations and a friction drilling application.

[1] Mak, Simon and V. Roshan Joseph. *Support points.* The Annals of Statistics 46.6A, pp. 2562–2592, 2018.

[2] O. Cappé and A. Guillin and J. M. Marin and C. P. Robert. *Population Monte Carlo.* Journal of Computational and Graphical Statistics, 13.4, pp. 907–929, 2004.

Back to overview

zoom 5, Tuesday, 20:30

Simulation of Maximum Drawdown for Diffusions

Liang Feng
University of California, Santa Barbara, `liangfeng@pstat.ucsb.edu`

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

In financial industry, drawdown-based rules have been highly effective in improving the performance of money managers. These rules are important for both risk management as well as portoflio return. One of the primary quantities of interest in this setting is the maximum drawdown $\mathcal{M}_T(X)$ of a stochastic process X by time T, defined as

$$\mathcal{M}_T(X) = \sup_{t \leq T} \left(\sup_{s \leq t} (X_s - X_t) \right).$$

In words, $\mathcal{M}_T(X)$ is the largest drop in X from peak to bottom on $[0, T]$. In academia, maximum drawdown statistics are rarely considered however, likely owing to challenges with tractability. Moreover, surprisingly, little to no literature on the simulation of maximum drawdown exists. We address this gap between theory and practice in this work.

We begin by developing a basic acceptance-rejection (AR) sampler for the case of a Brownian motion W, for which the density of $\mathcal{M}_T(W)$ is well-known as an alternating infinite series that makes inversion sampling prohibitive. Our AR algorithm generates $4/\pi$ candidates on average before a sample of $\mathcal{M}_T(W)$ is drawn. We use this scheme as a basis to systematically treat other Brownian motion based processes (Vasicek, OU, CIR, etc). Finally, leveraging exact methods for diffusion simulation, we develop an AR scheme to sample $\mathcal{M}_T(X)$ for a diffusion X with a general drift and volatility under some conditions on these coefficients. Numerical results illustrate the performance of the method with the Euler scheme used as the benchmark. These simulations also shed insight into the order of covergence of Euler schemes for various function of the maximum drawdown.

Back to overview

zoom 5, Wednesday, 14:00

Euler schemes for piecewise deterministic Markov processes

Andrea Bertazzi
Delft University of Technology, a.bertazzi@tudelft.nl

Joris Bierkens
Delft University of Technology, joris.bierkens@tudelft.nl

Paul Dobson
Delft University of Technology, p.dobson@tudelft.nl

Piecewise deterministic Markov processes (PDMP) received substantial interest in recent years as an alternative to classical Markov chain Monte Carlo algorithms. The applicability of PDMP to real world scenarios is currently limited by the fact that these processes can be simulated only when specific properties of the target distribution are known beforehand. In order to overcome this problem, we introduce an Euler-type discretisation scheme for PDMP which does not need such pre-requisite knowledge. For the resulting schemes we study both pathwise convergence to the continuous process as the step size converges to zero and convergence in law to the target measure in the long time limit. Finally, we study the empirical performance of the Euler discretised zig-zag and bouncy particle samplers and compare it to the corresponding continuous time processes.

[1] M. H. A. Davis. *Piecewise-Deterministic Markov Processes: A General Class of Non-Diffusion Stochastic Models.* Journal of the Royal Statistical Society. Series B (Methodological), vol. 46, no. 3, 1984.

Back to overview

zoom 5, Wednesday, 14:30

Sticky PDMP samplers for sparse and local inference problems

Joris Bierkens
Delft University of Techlology, joris.bierkens@tudelft.nl

Sebastiano Grazzi
Delft University of Techlology, s.grazzi@tudelft.nl

Frank van der Meulen
Delft University of Techlology, f.h.vandermeulen@tudelft.nl

Moritz Schauer
Chalmers University of Technology, University of Gothenburg, smoritz@chalmers.se

During the talk, I will present the *sticky PDMP samplers [1]*. This is a new class of efficient Monte Carlo methods based on continuous-time piecewise deterministic Markov processes (PDMPs) suitable for inference in high dimensional sparse models, i.e. models for which there is prior knowledge that many coordinates are likely to be exactly 0. This is achieved with the fairly simple idea of endowing existing PDMP samplers with sticky coordinate axes, coordinate planes etc. Upon hitting those subspaces, an event is triggered, during which the process *sticks* to the subspace, this way spending some time in a sub-model. That introduces *non-reversible* jumps between different (sub-)models. During the talk, I will touch upon computational aspects of the algorithm and illustrate the method for a number of statistical models where both the sample size N and the dimensionality d of the parameter space are large.

[1] J. Bierkens, S. Grazzi, F. van der Meulen, and M. Schauer. *Sticky PDMP samplers for sparse and local inference problems.* arXiv: 2103.08478, 2021.

Back to overview

zoom 5, Wednesday, 15:30

Numerical approximation of nonlinear delay differential equations under nonstandard assumptions

Natalia Czyżewska
AGH Doctoral School, Faculty of Applied Mathematics,
AGH University of Science and Technology in Krakow, Poland, nczyzew@agh.edu.pl

Delay differential equations (DDEs) have been studied for at least 200 years. However, before Bellen and Zennaro [1] there is a lack of books describing and unifying various approaches and techniques in the literature concerning solving DDEs. Nevertheless, classical literature for DDEs assumes some regularity of the right-hand side function of the initial value problem, commonly Lipschitz condition, alike to classical literature for ordinary differential equations. Meanwhile, it turns out that the real world applications need nonstandard assumptions. Namely, a phase change of metallic materials can be described by DDE with a non-Lipschitz right-hand side function, see [5, chapter 3.3] and [4, 3].

Firstly, we will consider a case when a right-hand side function is one-dimensional, locally Hölder continuous and monotone [3]. Then we will sketch a case with a multidimensional right-hand side function which is also locally Hölder continuous and fulfills linear growth together with one-side Lipschitz conditions [2]. Finally, the tests of the above results in simulations will be presented.

Joint work with Paweł Morkisz and Paweł Przybyłowicz, both from Faculty of Applied Mathematics, AGH University of Science and Technology in Krakow, Poland.

[1] A. Bellen and M. Zennaro. *Numerical methods for delay differential equations.* Oxford, New York, 2003.

[2] N. Czyżewska, P. Morkisz, P. Przybyłowicz. *Optimal solution of delay differential equations under nonstandard assumptions.* In preparation.

[3] N. Czyżewska, J. Kusiak. P. Morkisz, P. Oprocha, M. Pietrzyk, P. Przybyłowicz, Ł. Rauch and D. Szeliga. *On mathematical aspects of evolution of dislocation density in metallic materials.* 2020, arXiv:2011.08504.

[4] N. Czyżewska, J. Kusiak. P. Morkisz, P. Oprocha, M. Pietrzyk, P. Przybyłowicz, Ł. Rauch, D. Szeliga. *Prediction of Distribution of Microstructural Parameters in Metallic Materials Described by Differential Equations with Recrystallization Term.* International Journal for Multiscale Computational Engineering, **17**(3) (2019), 361-371.

[5] M. Pietrzyk, Ł. Madej, Ł. Rauch, D. Szeliga. *Computational Materials Engineering: Achieving high accuracy and efficiency in metals processing simulations.* Butterworth-Heinemann, Elsevier, Amsterdam, 2015.

Back to overview

zoom 5, Wednesday, 16:00

Coefficient Estimation in Nonautonomous Ordinary Differential Equations Using the Principle of a Solver Family Based on Markov Jump Processes

Flavius Guiaş
Dortmund University of Applied Sciences and Arts, flavius.guias@fh-dortmund.de

The values of the parameter vector driving a system of ordinary differential equations are estimated by an optimization algorithm which minimizes an error function constructed in order to match given data obtained from statistics or measurements. We consider systems of ordinary differential equations of the form $\vec{y}'(t) = \vec{F}(\vec{R}(t), \vec{p}; \vec{y}(t))$ which depend on a set $\vec{R}(t)$ of time-dependent parameters and on another set $\vec{p}$ of time independent parameters. The goal is to find a set of parameters such that the error in mean-square sense between the solution $\vec{y}(t_i)$ and given data points $\vec{Y}_i$, with $i = 1, \dots n$, is minimal, considering that the initial condition is $\vec{y}(0) = \vec{Y}_0$.

Each function call in searching for a minimizer needs to calculate numerical approximations $\vec{y}_i^*$ for $\vec{y}(i)$. Our choice for the solver is based on the requirements that it should be cheap and accurate, since the optimization algorithm involves a large number of function evaluations.

A modified version of the stochastic algorithm presented in [1] and [2] fulfills these criteria. The original method is based on a predictor computed as a path of a Markov jump process whose values are improved by using steps of Picard or Runge-Kutta type, which have a strong effect in reducing the stochastic fluctuations. The result of this scheme is a smooth numerical approximation of high order of the solution of the differential equation.

In the method presented here, instead of using the data produced by the Markov jump process, we use the variations of the statistical data for computing the predictor and obtain next an improved approximation by the steps of the stochastic solver on a short time interval, where the parameters are assumed to be constant.

We use an initial guess for the parameter vector and apply several runs of a splitting algorithm, by optimizing alternatively with respect to $\vec{p}$ by keeping $\vec{R}$ fixed, and then with respect to $\vec{R}$ by keeping $\vec{p}$ fixed. The minimum of the error functional is searched by a combination of deterministic and Monte-Carlo methods in order to obtain an optimal set of parameters which delivers a solution as close as possible to the real data points.

[1] F. Guiaş and P. Eremeev. Improving the stochastic direct simulation method with applications to evolution partial differential equations. *Appl. Math. Comput.*, **289** : 353–370, (2016)

[2] F. Guiaş. High precision stochastic solvers for large autonomous systems of differential equations. *Int. J. Math. Models and Meth. in Appl. Sci.* **13** : 60–63, (2019)

Back to overview

zoom 5, Wednesday, 19:00

Wasserstein contraction for non-elliptic diffusions

Pierre Monmarché
Sorbonne Université, pierre.monmarche@sorbonne-universite.fr

In some recent works concerned with the use of the (underdamped) Langevin diffusion (or of the randomized HMC process) for MCMC purpose, it has been used that, for a convex and smooth target log density, assuming the damping is large enough, these processes induce a contraction of the Wasserstein distances. This is well understood in the elliptic reversible case, since the famous results of Sturm and Von Renesse in 2005 who gave several characterizations of this Wasserstein contraction. We will see what remains of these characterizations for diffusion processes on R^d with constant diffusion matrices in the non-elliptic (possibly non-hypoelliptic) case, which covers for instance the underdamped Langevin process.

[1] P. Monmarché. *Almost sure contraction for diffusions on R^d. Application to generalised Langevin diffusions.* arXiv e-prints, arXiv:2009.10828, 2020.

Back to overview

zoom 5, Wednesday, 19:30

Entropy dissipation via information Gamma calculus

Qi Feng
University of Michigan, qif@umich.edu

Wuchen Li
University of South Carolina, wuchen@mailbox.sc.edu

In this talk, we present some convergence behaviors for some non-gradient degenerate stochastic differential equations towards their invariant distributions. Our method extends the connection between Gamma calculus and Hessian operators in the Wasserstein space. In detail, we apply Lyapunov methods in the space of probabilities, where the Lyapunov functional is chosen as the relative Fisher information. We derive the Fisher information induced Gamma calculus to handle non-gradient drift vector fields and degenerate diffusion matrix. Several examples are provided for non-reversible Langevin dynamics, sub-Riemannian diffusion process, and variable-dependent underdamped Langevin dynamics.

[1] Q. Feng, and W. Li *Hypoelliptic entropy dissipation for stochastic differential equations.* arXiv:2102.00544, 2021

Back to overview

zoom 5, Wednesday, 20:00

A blob method for diffusion and applications to sampling and two layer neural networks

Katy Craig
University of California, Santa Barbara, kcraig@math.ucsb.edu

Karthik Elamvazhuthi
University of California, Los Angeles, karthikevaz@math.ucla.edu

Matt Haberland
Cal Poly, San Luis Obispo, mhaberla@calpoly.edu

Olga Turanova
Michigan State University, turanova@msu.edu

Given a desired target distribution and an initial guess of that distribution, composed of finitely many samples, what is the best way to evolve the locations of the samples so that they more accurately represent the desired distribution? A classical solution to this problem is to allow the samples to evolve according to Langevin dynamics, the stochastic particle method corresponding to the Fokker-Planck equation. In today's talk, I will contrast this classical approach with a deterministic particle method corresponding to the porous medium equation. This method corresponds exactly to the mean-field dynamics of training a two layer neural network for a radial basis function activation function. We prove that, as the number of samples increases and the variance of the radial basis function goes to zero, the particle method converges to a bounded entropy solution of the porous medium equation. As a consequence, we obtain both a novel method for sampling probability distributions as well as insight into the training dynamics of two layer neural networks in the mean field regime.

Back to overview

zoom 5, Wednesday, 20:30

On Wasserstein gradient flows and the search for neural network architectures

Nicolás García Trillos
University of Wisconsin Madison, garciatrillo@wisc.edu

Neural networks have revolutionized machine learning and artificial intelligence in unprecedented ways, establishing new benchmarks in performance in applications such as image recognition and language processing. Such success has motivated researchers and practitioners in multiple fields to develop further applications. This environment has driven several novel research directions. In particular, one crucial question that has received increased recent attention is related to the design of good neural architectures using data-driven approaches and minimal human intervention. In this talk I will discuss a framework in which ideas from optimal transport can be used to motivate algorithms for the exploration of the architecture space. In the first part of the talk, I will abstract the problem of neural architecture search slightly and discuss how optimal transport can motivate first order and second order gradient descent schemes for the optimization of a semi-discrete objective function. I will then return to the original neural architecture search problem, and using the ideas discussed during the first part of the talk, I will motivate two algorithms for neural architecture search called NASGD and NASAGD. I will wrap up by discussing the performance of our algorithms when searching an architecture for a classification problem with the CIFAR-10 data set, and providing some perspective on future research directions. This talk is based on joint work with Félix Morales and Javier Morales.

Back to overview

zoom 5, Thursday, 09:00

Data driven uncertainty quantification in macroscopic traffic flow models

Alexandra Würth
Inria Sophia Antipolis, alexandra.wuerth@inria.fr

Mickaël Binois
Inria Sophia Antipolis, mickael.binois@inria.fr

Paola Goatin
Inria Sophia Antipolis, paola.goatin@inria.fr

Simone Göttlich
Universität Mannheim, goettlich@uni-mannheim.de

We propose a Bayesian calibration technique for parameter identification and uncertainty quantification in macroscopic traffic flow models, exploiting different loop detector data sets. We validate the results comparing the error metrics of both first and second order models. While needing more parameter calibration, second order models generally perform better in reconstructing traffic quantities of interest.

Macroscopic traffic flow models have been employed for decades to describe the spatio-temporal evolution of aggregate traffic quantities such as density and mean velocity. Classically, macroscopic traffic models are calibrated either by fitting the so-called fundamental diagram (i.e., the density-flow or density-speed mapping described by the model flux function) or by minimizing some error measure of the simulation output. The calibration can be done against either data provided by loop detectors at fixed locations or trajectory data. To our knowledge however, few works have been devoted evaluating the uncertainty of both models and data. Thus, we propose to follow a *Bayesian approach*, which allows us to evaluate the parameter probability distribution given the observed data. Moreover, following Kennedy-O'Hagan [1], we introduce a bias term to better account for possible discrepancies between the mathematical models and reality; this bias term is modeled by a Gaussian process.

[1] M. C. Kenney and A. O'Hagan. *Bayesian calibration of computer models.* Journal of the Royal Statistical Society: Series B (Statistical Methodology), 63(3): 425-464, 2001.

Back to overview

zoom 5, Thursday, 09:30

Uncertainty quantification for traffic flow models via a stochastic Galerkin approach

Stephan Gerster
RWTH Aachen University, gerster@igpm.rwth-aachen.de

Michael Herty
RWTH Aachen University, herty@igpm.rwth-aachen.de

Elisa Iacomini
RWTH Aachen University, iacomini@igpm.rwth-aachen.de

Although traffic models have been extensively studied, obtaining trustful forecast from these models is still challenging, since the evolution of traffic is also exposed to the presence of uncertainties.

In this talk, we will investigate the propagation of uncertainties in traffic flow models, especially in macroscopic second order models applying the stochastic Galerkin approach. Hyperbolic preserving stochastic Galerkin formulations are presented in conservative form, and for smooth solutions also in the corresponding non-conservative form. This allows one to obtain stabilization results, when the system is relaxed to a first-order model [2]. We will illustrate the theoretical results with numerical simulations.

[1] S. Gerster, M. Herty, E. Iacomini. *Stability analysis of a hyperbolic stochastic Galerkin formulation for the Aw-Rascle-Zhang model with relaxation.* Mathematical Biosciences and Engineering, 2021.

[2] M. Herty, G. Puppo, S. Roncoroni, G. Visconti. *The BGK approximation of kinetic models for traffic.* Kinetic & Related Models, 2020.

[3] S. Jin and R. Shu. *A study of hyperbolicity of kinetic stochastic Galerkin system for the isentropic Euler equations with uncertainty.* Chinese Annals of Mathematics, Series B, 2019.

Back to overview

zoom 5, Thursday, 10:00

Microscopic and macroscopic traffic flow models including random accidents: theory and numerical simulations

Simone Göttlich
Universität Mannheim, goettlich@uni-mannheim.de

Thomas Schillinger
Universität Mannheim, schillinger@math.uni-mannheim.de

We introduce microscopic and macroscopic stochastic traffic models including traffic accidents. The microscopic model is based on a Follow-the-Leader approach with ordinary differential equations whereas the macroscopic model is described by a scalar conservation law with space dependent flux function.

Accidents are introduced as a stochastic disturbance of the deterministic evolution of the traffic models and are directly linked to the traffic situation. Specifically, we focus on two types of accidents: First, we consider accidents due to high traffic flux which correspond to the idea that accidents are more likely having both, a high density and high velocity of the vehicle. Second, we investigate rear-end collisions which for example can be observed at tailbacks.

In contrast to other accident models, we consider a framework in which the traffic situation and accidents are connected in a bi-directional relation. On the one hand, accidents influence the traffic flow by adjusting the velocity function in the microscopic model and the flux function in the macroscopic model. On the other hand, the probability measures modeling the accident times and accident positions depend on the current traffic situation.

Concentrating on a microscopic submodel and based on a Lax-Friedrichs discretization, convergence of the microscopic model to the macroscopic model is shown.

Numerical investigations using Monte Carlo simulations are presented to compare the above models and show their convergence behaviour. Additionaly, data from a German highway is used to validate the accident models.

[1] Göttlich, S. and Knapp, S. *Modeling random traffic accidents by conservation laws.* Mathematical Biosciences and Engineering 17 (2020), pp. 1677-1701.

[2] Karlsen, K. and Towers J. D. *Convergence of the Lax-Friedrichs Scheme and Stability for Conservation Laws with a Discontinuous Space-Time Dependent Flux.* Chinese Annals of Mathematics 25.03 (2004), pp. 287-318.

[3] Wagner, D. H. *Equivalence of the Euler and Lagrangian equations of gas dynamics for weak solutions.* Journal of Differential Equations 68.1 (1987), pp. 118-136.

Back to overview

zoom 5, Thursday, 10:30

Kernel density estimation on gas networks: Probabilistic robustness

Michael Schuster
FAU Erlangen-Nürnberg, michi.schuster@fau.de

Elisa Strauch
TU Darmstadt, strauch@mathematik.tu-darmstadt.de

Martin Gugat
FAU Erlangen-Nürnberg, gugat@math.fau.de

Jens Lang
TU Darmstadt, lang@mathematik.tu-darmstadt.de

Motivated by optimization problems with probabilistic constraints we compute the probability that certain variables are feasible under uncertain gas demand. Feasibility in this context means, that the solution of the gas network model satisfies bounds for the pressure at the network junctions. The probabilities can be seen as a measure for the robustness of these variables.

The key technique for computing these probabilities is *kernel density estimation*, which allows us to estimate unknown probability density functions using a given sample of gas demand scenarios. We provide an efficient way to compute these probabilities and show how the kernel density estimator approach can be applied to optimization problems with probabilistic constraints.

[1] M. Schuster, E. Strauch, M. Gugat and J. Lang. *Probabilistic Constrained Optimization on Flow Networks.* Optim. Eng. 2021.

Back to overview

zoom 5, Thursday, 19:00

Monotonic Alpha-divergence Minimisation

Kamélia Daudel
Télécom Paris, Institut Polytechnique de Paris,
kamelia.daudel@gmail.com

Randal Douc
Télécom SudParis, Institut Polytechnique de Paris,
randal.douc@telecom-sudparis.eu

François Roueff
Télécom Paris, Institut Polytechnique de Paris,
francois.roueff@telecom-paris.fr

In this work [1], we introduce a novel iterative algorithm which carries out α-divergence minimisation by ensuring a systematic decrease in the α-divergence at each step. In its most general form, our framework allows us to simultaneously optimise the weights and components parameters of a given mixture model.

Notably, our approach permits us to build on various methods previously proposed for α-divergence minimisation such as gradient (e.g [3]) or power [2] descent schemes. Furthermore, we shed a new light on an integrated Expectation Maximization algorithm [4]. As our methodology involves computing integrals which might be intractable, we resort to Monte Carlo approximations in practice. We provide empirical evidence that our approach yields improved results, all the while illustrating the numerical benefits of having introduced some flexibility through the parameter α of the α-divergence.

[1] K. Daudel, R. Douc and F. Roueff. *Monotonic Alpha-divergence Minimisation.* Available on arxiv at https://arxiv.org/pdf/2103.05684.pdf, 2021.

[2] K. Daudel, R. Douc and F. Portier. *Infinite-Dimensional Alpha-Divergence Minimisation for Variational Inference.* To appear in the Annals of Statistics and available on arxiv at https://arxiv.org/pdf/2005.10618.pdf, 2020.

[3] Y. Li and R. E Turner. *Rényi divergence variational inference.* Advances in Neural Information Processing Systems 29, pages 1073–1081. Curran Associates, Inc., 2016. http://papers.nips.cc/paper/6208-renyi-divergence-variational-inference.pdf.

[4] O. Cappé, R. Douc, A. Guillin, J. Marin, and C. P Robert. *Adaptive importance sampling in general mixture classes.* Statistics and Computing, 18(4):447–459, 2008.

Back to overview

zoom 5, Thursday, 19:30

Global Stochastic Optimization using Particle Filter Algorithms

Mathieu Gerber
University of Bristol, mathieu.gerber@bristol.ac.uk

Randal Douc
SAMOVAR, Télécom SudParis, Institut Polytechnique de Paris, randal.douc@it-sudparis.eu

We introduce a new algorithm to learn on the fly the parameter value $\theta_\star = \mathbb{E}[\log f_\theta(Y_0)]$ from a sequence $(Y_t)_{t\geq 1}$ of independent copies of Y_0, with $\{f_\theta, \theta \in \Theta \subseteq \mathbb{R}^d\}$ a parametric model for the distribution of Y_0. The main idea of the proposed approach is to define a sequence $(\tilde{\pi}_t)_{t\geq 1}$ of probability distributions on Θ which (i) is shown to concentrate on $\theta_\star$ as $t \to \infty$ and (ii) can be estimated in an online fashion by means of a standard particle filter (PF) algorithm. Denoting by $\tilde{\pi}_t^N$ the PF approximation of $\tilde{\pi}_t$ with support size $N \in \mathbb{N}$, the parameter value $\theta_\star$ is then estimated by $\tilde{\theta}_t^N := \int_\Theta \theta \tilde{\pi}_t^N(\mathrm{d}\theta)$. Our convergence result for $(\tilde{\pi}_t)_{t\geq 1}$ holds under minimal assumptions on the statistical model and, in particular, does not assume that the objective function $\theta \mapsto \mathbb{E}[\log f_\theta(Y_0)]$ is uni-modal. This makes the PF stochastic optimization (PFSO) algorithm introduced in this work a derivative-free global stochastic optimization method. The sequence $(\tilde{\pi}_t)_{t\geq 1}$ depends on a learning rate $h_t \to 0$, and we observe that the slower $h_t \to 0$ the greater is the ability of $\tilde{\theta}_t^N$ to escape from a local optimum of the objective function but the slower is its convergence rate towards $\theta_\star$. Interestingly, our numerical experiments suggest that, for a finite N, the Polyak-Ruppert arranging $\bar{\theta}_t^N = t^{-1}\sum_{s=1}^t \tilde{\theta}_s^N$ of $\{\tilde{\theta}_s^N\}_{s=1}^T$ converges to $\theta_\star$ at the optimal $t^{-1/2}$ rate in situations where $\tilde{\theta}_t^N$ converges to this parameter value at a sub-optimal rate. We illustrate the practical usefulness of the proposed PFSO algorithm for online parameter learning and for computing the maximum likelihood estimator in challenging statistical models.

Back to overview

zoom 5, Thursday, 20:00

Quasi-Newton Quasi-Monte Carlo in Numerical Optimization

Sifan Liu
Stanford University, sfliu@stanford.edu

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

Numerical optimization solves problems of the form $\min_\omega E_\xi f(\omega;\xi)$. Common practice is to run stochastic gradient descent (SGD) with the gradient being estimated by Monte Carlo methods. Two issues remain for this approach: (1) It often requires large sample size to reduce Monte Carlo error of the gradient estimators; (2) Even if the gradient estimator is accurate, SGD can preform poorly when the Hessian of the objective function is ill-conditioned. We propose to use randomized quasi-Monte Carlo (RQMC) and stochastic quasi-Newton methods to address these problems. We prove that RQMC achieves a smaller stationary variance than plain Monte Carlo when using a stochastic quasi-Newton algorithm. The effectiveness of the proposed algorithm is illustrated by simulations in variational inference.

Back to overview

zoom 5, Friday, 09:00

Source term estimation for random walk approximations of kinetic equations

Bert Mortier
KU Leuven, Department of Computer Science, bert.mortier@kuleuven.be

Martine Baelmans
KU Leuven, Department of Mechanical Engineering, martine.baelmans@kuleuven.be

Giovanni Samaey
KU Leuven, Department of Computer Science, giovanni.samaey@kuleuven.be

Recently, new algorithms have been proposed [3, 2] to simulate the Boltzmann-Bhatnager-Gross-Krook kinetic equation when collisionality is high under a diffusive scaling. In the limit of infinite collision rates, the standard Monte Carlo method, which executes each collision individually, becomes too computationally expensive. Simultaneously, the kinetic model converges to an advection-diffusion equation in this limit, for which relatively cheap simulation methods exist. The new algorithms that were developed, combine the standard Monte Carlo method with a random walk Monte Carlo method. The proposed hybridizations maintain the correct behaviour of the kinetic model in both the kinetic and the diffusive limits. These methods are attractive since they are fully Monte Carlo, admitting relatively simple code. Furthermore, these methods do not resolve the, then-irrelevant, velocity dimensions in the advection-diffusion regime.

The discussed simulation methods provide samples according to the kinetic model that are asymptotically correct. The quantities of interest from Boltzmann-BGK simulations often are distributions over the domain, such as source terms in the context of nuclear fusion simulations. The work presented here focuses on extracting quantities of interest in a low-variance-way from samples that follow the advection-diffusion model. We do so, by adopting a track-length-like strategy [1].

[1] B. Mortier *Advanced Monte Carlo simulation and estimation for kinetic neutral particles in the plasma edge of fusion reactors.* PhD thesis, 2020.

[2] B. Mortier, M. Baelmans and G. Samaey *Kinetic-diffusion asymptotic-preserving Monte Carlo algorithm for Boltzmann-BGK in the diffusive scaling.* arXiv preprint arXiv:2012.08985, 2020

[3] G. Dimarco, L. Pareschi and G. Samaey *Asymptotic-Preserving Monte Carlo methods for transport equations in the diffusive limit.* SIAM Journal on Scientific Computing 40(1): A504–A528, 2018

Back to overview

zoom 5, Friday, 09:30

Last-passage algorithm for charge distribution over a finite region

Unjong Yu
Gwangju Institute of Science and Technology, uyu@gist.ac.kr

Hoseung Jang
Gwangju Institute of Science and Technology, jhs1751@gist.ac.kr

Chi-Ok Hwang
Gwangju Institute of Science and Technology, chwang@gist.ac.kr

First-passage [1, 2, 3] and last-passage [4, 5, 6] Monte Carlo algorithms have been used for computing charge distribution on a conducting object. First-passage algorithms are used for an overall charge distribution on a conducting object and Given-Hwang's last-passage algorithms are used for a charge density at a specific point on a flat or spherical surface of a conductor. In this talk, I will present a last-passage algorithm for computing charge distribution over a finite region of a conducting object.

[1] H.-X. Zhou, A. Szabo, J. F. Douglas, and J. B. Hubbard. J. Chem. Phys. 100, 3821 (1994).

[2] J. A. Given, J. B. Hubbard, and J. F. Douglas. J. Chem. Phys. 106, 3721 (1997).

[3] C.-O. Hwang and M. Mascagni. J. Appl. Phys. 95, 3798 (2004).

[4] C.-O. Hwang and M. Mascagni. J. Korean Phys. Soc. 42, L1 (2003).

[5] C. Yan, W. Cai, and X. Zeng. SIAM J. Sci. Comput. 35-4, B868 (2013).

[6] H. Jang, U. Yu, Y. Chung, and C.-O. Hwang. Adv. Theory Simul. 3, 2000075 (2020).

Back to overview

zoom 5, Friday, 10:00

Accelerating plasma-edge neutral particle simulations with particle-based multilevel Monte Carlo

Emil Løvbak
KU Leuven, emil.loevbak@kuleuven.be

Bert Mortier
KU Leuven, bert.mortier@kuleuven.be

Giovanni Samaey
KU Leuven, giovanni.samaey@kuleuven.be

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

A major computational bottleneck in the in-silico design of tokamak fusion reactor components is the simulation of kinetic equations, modeling neutral particles near the plasma-edge. Simulating these diffusively-scaled kinetic equations is costly due to their high dimensionality, as they model the particle distribution in a position-velocity phase space. In addition, high collision-rates introduce constraints on the simulation time step size, further increasing the computational cost of naive simulation techniques.

In recent years, a new class of multilevel Monte Carlo particle algorithms has been developed at KU Leuven, which are capable of significantly accelerating computations in this area. These algorithms make use of asymptotic-preserving schemes, which allow approximate simulations to be performed with large time steps while still converging to the considered kinetic equation for small time steps. In this talk we present an overview of the current state of the art concerning these asymptotic-preserving multilevel Monte Carlo algorithms (APMLMC) as well as their applicability to problems in plasma-edge simulations.

[1] B. Mortier, P. Robbe, M. Baelmans, G. Samaey (2020) Multilevel Asymptotic-Preserving Monte Carlo for Particle Simulations arXiv:2004.04071

[2] E. Løvbak, B. Mortier, G. Samaey and S. Vandewalle (2020) Multilevel Monte Carlo with improved correlation for kinetic equations in the diffusive scaling V. Krzhizhanovskaya et al. (eds.) *LNCS – ICCS 2020*, pp. 374–388

[3] E. Løvbak, G. Samaey and S. Vandewalle (2020) A Multilevel Monte Carlo Asymptotic-Preserving Particle Method for Kinetic Equations in the Diffusion Limit B. Tuffin and P. L'Ecuyer (eds.) *MCQMC 2018*, pp. 383–402

[4] E. Løvbak, G. Samaey and S. Vandewalle (2019) A multilevel Monte Carlo method for asymptotic-preserving particle schemes in the diffusive limit In Review, arXiv:1907.04610

Back to overview

zoom 5, Friday, 10:30

p-refined Multilevel Quasi-Monte Carlo with Locally Nested Random Field Evaluation Points Applied to a Geotechnical Engineering Problem

Philippe Blondeel
Department of Computer Science, KU Leuven, philippe.blondeel@kuleuven.be

Pieterjan Robbe
Department of Computer Science, KU Leuven, pieterjan.robbe@kuleuven.be

Stefan Vandewalle
Department of Computer Science, KU Leuven, stefan.vandewalle@kuleuven.be

François Stijn
Department of Civil Engineering, KU Leuven, stijn.francois@kuleuven.be

Geert Lombaert
Department of Civil Engineering, KU Leuven, geert.lombaert@kuleuven.be

Engineering problems are often characterized by significant uncertainty in their material parameters. Multilevel sampling methods are a straightforward manner to account for this uncertainty. A popular multilevel method consists of the classical Multilevel Monte Carlo method (MLMC) [1]. This method relies on a hierarchy of successive refined Finite Element meshes of the considered engineering problem, where most of the samples are taken on coarse and computationally cheap meshes, while a decreasing number of samples are taken on finer and computationally expensive meshes. Classically, the mesh hierarchy is constructed by selecting a coarse mesh discretization of the problem, and recursively applying an h-refinement approach to it. However, the h-refined mesh hierarchy increases the number of degrees of freedom almost geometrical with increasing level, leading to a large computational cost during runtime of MLMC. An efficient manner to reduce this computational cost, is by means of the improved sampling method, called p-refined Multilevel Quasi-Monte Carlo (p-MLQMC) [2]. The p-MLQMC method combines a hierarchy of p-refined Finite Element meshes, with a deterministic Quasi-Monte Carlo sampling rule, yielding computational savings with respect to classical MLMC. However, the challenge in the p-MLQMC method consists in adequately incorporating the uncertainty, represented as a random field, in the Finite Element model. In order to do so, point evaluations of the random field are computed at certain carefully chosen spatial locations, and assigned to the quadrature points used for the numerical integration of the element stiffness matrices. We investigated how these points, at which the random field is evaluated by means of the Karhunen-Loève expansion, are to be selected in order to obtain a lower computational cost. We have shown that using sets which consist of points which are nested across the

mesh hierarchy, i.e., the Local Nested Approach (LNA), yields a speedup up to a factor 5 with respect to sets which consist of points which are not nested, i.e., the Non-Nested Approach (NNA). All the aforementioned implementations are benchmarked on a slope stability problem, with spatially varying uncertainty in the ground.

[1] Giles, Michael B. (2008). *Multilevel Monte Carlo Path Simulation. Oper. Res*, 56(3):607–617.

[2] Blondeel, Philippe and Robbe, Pieterjan and van Hoorickx, Cédric and François, Stijn and Lombaert, Geert and Vandewalle, Stefan (2020). *p-Refined Multilevel Quasi-Monte Carlo for Galerkin Finite Element Methods with Applications in Civil Engineering. Algorithms*, 13(5):1–30.

6.6 Zoom Channel 6

back to table of contents

Clicking on the name of the speaker will lead you to the abstract. NOTE THAT ALL TIMES ARE UTC+2.

	Monday	Tuesday	Wednesday	Thursday	Friday
		Contributed talks on **Numerics for SPDEs**			Contributed talks on **Discrepancy and Dispersion**
09:00		Andreas Stein			
09:30		Goncalo dos Reis			Kosuke Suzuki
10:00		Hamza Ruzayqat			Shin Harase
10:30		Robin Merkle			Benjamin Sommer
			Contributed Talks on **MCQMC in Statistics**	Special Session **Recent Advances on the Laplace Approx. (part A)**	
13:30					
14:00					
14:30					
15:30				Philipp Wacker	
16:00			Natalya Tracheva	Remo Kretschmann	
16:30			Nathan Yang	Mareike Hasenpflug	
	Special Session **Quasi Monte Carlo Software (part A)**	Special Session **Quasi Monte Carlo Software (part B)**	Contributed talks on **MCQMC Integration**	Special Session **Recent Advances on the Laplace Approx. (part B)**	
19:00	Sergei Kucherenko	Aleksei Sorokin	Yuya Suzuki	Pauching Yap	
19:30	Onyekachi Osisiogu	Pieterjan Robbe	Art Owen	Umberto Villa	
20:00	Cyril Chimisov	Adrian Ebert	Raaz Dwivedi	Youssef Marzouk	
20:30	Pamphile T. Roy	Fred J. Hickernell	Adam Knapp		

zoom 6, Monday, 19:00

Construction and comparison of high-dimensional Sobol' sequence generators

Sergei Kucherenko
Imperial College London, UK, s.kucherenko@imperial.ac.uk

Danil Asotsky
Esti Map, Russia, danil@asotsky.ru

Emanouil Atanassov
IICT, Bulgarian Academy of Sciences, Bulgaria, emanouil@parallel.bas.bg

Pamphile T. Roy
iTranslate, France, roy.pamphile@gmail.com

Sobol′ sequences are the most widely used in practice low-discrepancy sequences (LDS), e.g. they are the only LDS used in Finance. The definition of the Sobol′ sequences offers substantial freedom in choosing the so-called direction numbers, so that the basic properties of the sequences are ensured, while additional equidistribution properties are achieved through some optimization procedure. The effect of the uniformity properties A and A′ on the generator performance in high dimensional problems was explored in Sobol' et al.(2012). It was shown that these properties provide an additional guarantee of uniformity for high-dimensional problems even at a small number of sampled points. By imposing additional uniformity properties on low dimensional projections of the sequence in addition to the uniformity properties of the d-dimensional sequence itself the efficiency of the Sobol′ sequence can be increased. These criteria were used for the construction of the Sobol′ sequence generator produced by BRODA Ltd. (2021).

Sobol′ sequences are known to have poor two-dimensional projections at low number of sampled points. Joe and Kuo (2008, 2021) provided a set of direction numbers alleviating this problem. These are obtained by treating Sobol′ sequences in (d) dimensions as (t, d)-sequences and then optimizing the t-values of the two-dimensional projections. In this work we compare these two well known sets of direction numbers and based on them Sobol′ sequence generators on a set of test problems ranging from integral evaluations to option pricing.

[1] BRODA Ltd. http://broda.co.uk, 2021.

[2] S. Joe and F. Kuo. *Constructing Sobol Sequences with Better Two-Dimensional Projections. SIAM Journal on Scientific Computing.* 30: 2635 – 2654, 2008.

[3] S. Joe and F. Kuo. http://web.maths.unsw.edu.au/ fkuo/sobol/, 2021.

[4] I.M. Sobol', D. Asotsky, A. Kreinin and S. Kucherenko. *Construction and comparison of high-dimensional Sobol' generators, Wilmott Magazine.* 56, 64–79, 2012.

zoom 6, Monday, 19:30

Analysis and implementation of a component-by-component digit-by-digit construction algorithm for lattice rules

Adrian Ebert
RICAM, Austrian Academy of Sciences, Austria, adrian.ebert@ricam.oeaw.ac.at

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

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

Onyekachi Osisiogu
RICAM, Austrian Academy of Sciences, Austria, onyekachi.osisiogu@oeaw.ac.at

Lattice rules are quasi-Monte Carlo rules for approximating integrals over the s-dimensional unit cube. For dimensions $s > 2$, no explicit constructions are known such that one usually has to resort to computer search algorithms. In this talk, we consider a search algorithm that combines the component-by-component search with a digit-by-digit construction scheme. This algorithm constructs the generating vector in a component-by-component (CBC) way in which each component of the vector is constructed digit-by-digit (DBD) and the resulting construction is called the component-by-component digit-by-digit (CBC-DBD) algorithm. The quality function of the algorithm is independent of the smoothness parameter α of the underlying function space such that the proven error convergence rates hold simultaneously for multiple α. Furthermore, we show that the constructed lattice rules achieve the almost optimal order for the worst-case error convergence rate in the considered weighted function spaces. We present a fast implementation of the construction that exploits a smaller search space and analyze its complexity.

[1] A. Ebert, P. Kritzer, D. Nuyens, O. Osisiogu. *Digit-by-digit and component-by-component constructions of lattice rules for periodic functions with unknown smoothness.* Journal of Complexity, 101555, 2021.

Back to overview

zoom 6, Monday, 20:00

RQMC Tools for Tensorflow

Cyril Chimisov
Google, London, UK, cyrilchimisov@google.com

Yocheved Darmon
Université de Montréal, Canada, yocheved.darmon@umontreal.ca

Yves Gaeremynck
Google, London, UK, ygaeremynck@google.com

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

We present a set of Tensorflow functions designed to quickly compute randomized quasi-Monte Carlo (RQMC) point sets of arbitrary sizes and dimensions in parallel, and return them in tensors. The supported point sets are ordinary lattice rules, polynomial lattice rules, and digital nets in base 2, with the standard randomization choices such as a random shift modulo 1, a random digital shift, a linear scramble, and the tent transformation [1, 2]. Parameters for the point sets (the modulus, generating vectors, and generating matrices) can be read from files in a standard format produced by Latnet Builder [3]. Files with pre-scrambled generating matrices are also offered. These matrices were scrambled in a component by component fashion to minimize a measure of discrepancy. Numerical examples will be given to illustrate the speed gains and compare the variances of different point sets and randomization methods.

[1] J. Dick and F. Pillichshammer. *Digital Nets and Sequences: Discrepancy Theory and Quasi-Monte Carlo Integration*. Cambridge University Press, Cambridge, U.K., 2010.

[2] P. L'Ecuyer. Randomized quasi-Monte Carlo: An introduction for practitioners. In P. W. Glynn and A. B. Owen, editors, *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016*, pages 29–52, Berlin, 2018. Springer.

[3] P. L'Ecuyer, P. Marion, M. Godin, and F. Fuchhammer. A tool for custom construction of QMC and RQMC point sets. In *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2020*, 2020. https://arxiv.org/abs/2012.10263.

Back to overview

zoom 6, Monday, 20:30

Quasi-Monte Carlo for everyone: `scipy.stats.qmc` and `torch.quasirandom`

Pamphile T. Roy
iTranslate, `roy.pamphile@gmail.com`

Maximilian Balandat
Facebook, `balandat@gmail.com`

The Python programming language is becoming more and more popular, not just as a general purpose language, but as a scripting language for science and engineering. Besides its accessibility and flexibility, a key reason for this is its rich library ecosystem for scientific computing.

On one end of the spectrum, there are fundamental libraries such as SciPy [5], which provide basic and well established scientific tools for thousands of other packages. Anyone using Python for science is most certainly using SciPy either directly or as a dependency from another tool. On the other end, there are modern deep learning libraries such as PyTorch [4], which provide a modular and flexible high-level python interface to high-performance numerical backends with support for autodifferentiation and modern hardware such as GPUs, TPUs and custom accelerators. These libraries are evolving fast and used by millions of end-users; in fact, the overwhelming majority of modern machine learning development is done in such libraries. Initially focused on deep learning, they have been developing more and more towards supporting general purpose scientific computing workloads.

Our work introduced QMC capabitlities to both SciPy and PyTorch, making QMC available to millions of users through the libraries themselves and the many others that build on them . Bringing QMC to everyone will help advance the state-of-the-art in QMC methods in two ways:

1. With QMC methods now much easier to use and available "out-of-the-box", researchers and engineers will be able to utilize their benefits across a much broader range of problems, expanding their reach and impact far beyond their current usage.

2. As QMC methods are applied to more and a more diverse set of applications, this will raise a whole set of new practical (i.e., implementation-related) as well as fundamental methodological questions. For instance, recent applications of QMC in the context of sample average approximation in Bayesian optimization [1] led to the development of a strong law of large numbers for QMC [3], and our implementation work inspired additional work on highlighting the importance of including the first point in the *Sobol'* sequence [2].

Since SciPy 1.7, the following set of features is available:

- *Sobol'* and *Halton* sequences (scrambled and unscrambled),
- Multinominal and Multivariate sampler,
- LHS (optimized on C^2 coming soon),
- Discrepancy measures (C^2, wrap around, star-L_2, mixed),
- Scaling utility,
- Fast numerical inverse using Hermite spline (sample any distribution with QMC).

As for PyTorch, the same version of *Sobol'* as SciPy is provided, but implemented in PyTorch's native ATen Tensor library.

In this presentation, we propose to walk through the new features using examples. We hope to generate discussions, gather feedbacks to improve the libraries and call for new contributors.

Acknowledgements. The authors acknowledge Professors Art B. Owen, Fred Hickernell and Sergei Kucherenko for helpful discussions. The SciPy maintainer team also provided support and help regarding the design and integration. We thank especially Ralf Gommers and Matt Haberland for their thorough reviews.

[1] Maximilian Balandat et al. Botorch: A framework for efficient monte-carlo bayesian optimization. In H. Larochelle, M. Ranzato, R. Hadsell, M. F. Balcan, and H. Lin, editors, *Advances in Neural Information Processing Systems*, volume 33, pages 21524–21538. Curran Associates, Inc., 2020.

[2] Art B. Owen. On dropping the first Sobol' point. *arXiv e-prints*, page arXiv:2008.08051, August 2020.

[3] Art B. Owen and Daniel Rudolf. A strong law of large numbers for scrambled net integration. *SIAM Review*, 63(2):360–372, 2021.

[4] Adam Paszke et al. Pytorch: An imperative style, high-performance deep learning library. In H. Wallach, H. Larochelle, A. Beygelzimer, F. Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[5] Pauli Virtanen et al. Scipy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods*, 17(3):261–272, 2020.

Back to overview

zoom 6, Tuesday, 19:00

Building QMCPy's Quasi-Monte Carlo Framework

Aleksei G. Sorokin
Illinois Institute of Technology, asorokin@hawk.iit.edu

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

Sou-Cheng T. Choi
Kamakura Corporation, schoi32@iit.edu

Jagadeeswaran Rathinavel
Wi-Tronix, jrathin1@hawk.iit.edu

Michael J. McCourt
SigOpt, mccourt@sigopt.com

Quasi-Monte Carlo methods are used in many scientific applications to perform efficient, high dimensional numerical integration. This talk will describe the internal structure of QMCPy, an open source, Python library that implements research from across the QMC community into a cohesive, extensible framework. We overview the architecture of good QMC software and provide examples illustrating how our package categorizes these principles into an extensible object oriented framework. Specifically, we will overview the point generators, variable transforms (also known as importance sampling), and guaranteed approximation algorithms available in QMCPy.

[1] S.-C. T. Choi, F. J. Hickernell, R. Jagadeeswaran, M. J. McCourt, and A. G. Sorokin. Quasi-Monte Carlo software, 2021. arXiv:2102.07833 [cs.MS].

[2] S.-C. T. Choi, F. J. Hickernell, R. Jagadeeswaran, M. J. McCourt, and A. G. Sorokin. QMCPy Homepage, 2020+. qmcpy.org.

Back to overview

zoom 6, Tuesday, 19:30

Multilevel Quasi-Monte Carlo for System Reliability Assessment with QMCPy

Pieterjan Robbe
KU Leuven, pieterjan.robbe@kuleuven.be

Michiel Vromans
KU Leuven, michiel.vromans@kuleuven.be

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

We extend the Multilevel Monte Carlo framework for reliability theory from [1] to Multilevel Quasi-Monte Carlo. We show that the level selection for the Multilevel Monte Carlo method and the construction of good Quasi-Monte Carlo rules can be combined in a single screening phase that uses the same set of warm-up samples. We illustrate the efficiency of our approach using reliability block diagrams of a fire alarm system and a simplified hydraulic power plant. All simulations are performed with QMCPy [2].

[1] L. J. M. Aslett, T. Nagapetyan, and S. J. Vollmer. *Multilevel Monte Carlo for Reliability Theory.* Reliability Engineering & System Safety, Volume 165, 2017.

[2] S.-C. T. Choi, F. J. Hickernell, M. McCourt, J. Rathinavel and A. Sorokin. *QMCPy: A quasi-Monte Carlo Python Library.* Working. 2020. https://qmcsoftware.github.io/QMCSoftware/

Back to overview

zoom 6, Tuesday, 20:00

Computation of gradients of multivariate normal probabilities using automatic differentiation techniques and QMC methods

Adrian Ebert
Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austria,
adrian.ebert@ricam.oeaw.ac.at

Hernan Leövey
Structured Energy Management Team, Axpo, Baden, Switzerland, hernaneugenio.leoevey@axpo.com

Dirk Nuyens
Department of Computer Science, KU Leuven, Belgium, dirk.nuyens@cs.kuleuven.be

In this talk, we discuss the derivation and implementation of a fast algorithm for the simultaneous estimation of truncated multivariate normal probabilities as well as their gradients with respect to the boundaries of the domain. In particular, we combined the well-known algorithm for the calculation of multivariate normal probabilities by Alan Genz (see [1]) and used automatic differentiation (back-propagation) to calculate the gradients of the multivariate normal probability

$$P(\boldsymbol{a},\boldsymbol{b}) := \frac{1}{\sqrt{(2\pi)^s|\Sigma|}} \int_{a_1}^{b_1} \cdots \int_{a_s}^{b_s} e^{-\frac{1}{2}\boldsymbol{x}^\top \Sigma^{-1}\boldsymbol{x}}\, d\boldsymbol{x}$$

w.r.t. the borders $\boldsymbol{a},\boldsymbol{b}$ of the box domain $[a_1,b_1]\times\cdots\times[a_s,b_s]$. The numerical evaluation of gradients of multivariate normal probabilities is often required in stochastic optimization problems with probabilistic constraints, see, e.g., [2]. We demonstrate the effectiveness of our algorithm by numerical experiments.

[1] A. Genz. *Numerical Computation of Multivariate Normal probabilities.* Journal of Computational and Graphical Statistics, Vol. 1, No. 2, 141–149, 1992.

[2] I. Bremer, R. Henrion, A. Möller. *Probabilistic constraints via SQP solver: application to a renewable energy management problem.* Computational Management Science, Vol. 12, 435–459, 2015.

Back to overview

zoom 6, Tuesday, 20:30

Advances and Challenges in Quasi-Monte Carlo Software

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

Quasi-Monte Carlo (qMC) theory has developed over than six decades. Computer implementations of quasi-Monte Carlo algorithms have followed behind. With practitioners seeking ways to speed up their Monte Carlo calculations, and theorists eager to see qMC expand into new application areas, we need qMC software that is reliable, efficient, accessible, comprehensive, and easy-to-use. This talk describes the progress that has been made so far by various groups and what challenges lie ahead.

Contributed talks on Numerical Methods for SPDEs

Back to overview

zoom 6, Tuesday, 09:00

Numerical approximation of a stochastic transport problem with Lévy noise

Andreas Stein
ETH Zürich, andreas.stein@sam.math.ethz.ch

Semilinear hyperbolic stochastic partial differential equations (SPDEs) have various applications in the natural and engineering sciences. From a modelling point of view the Gaussian setting can be too restrictive, since applications in mathematical finance or phenomena as porous media and pollution models indicate an influence of noise of a different nature.

In order to capture temporal discontinuities and allow for heavy-tailed distributions, Hilbert space-valued Lévy processes (or Lévy fields) as driving noise terms are considered. The numerical discretization of the corresponding SPDE involves several difficulties: low spatial and temporal regularity of the solution to the problem entails slow convergence rates and instabilities for space-time discretization schemes. We resolve this problem by combining a backward Euler time stepping scheme with a discontinuous Galerkin discretization in space based on the pathwise weak formulation of the SPDE. A numerical analysis of the strong L^2-error reveals that this approach results in a stable method with optimal spatial convergence.

To handle the infinite-dimensional driving noise, we couple the space-time discretization scheme with truncated Karhunen-Loéve expansions for the Lévy process. This yields a fully discrete approximation scheme for the stochastic transport problem, and the mean-squared error bounds of the space-time discretization are extended to encompass the truncated series expansion of the noise.

This is joint work with Andrea Barth (SimTech/University of Stuttgart).

[1] A. Barth and A. Stein *A stochastic transport problem with Lévy noise: fully discrete numerical approximation.* arXiv preprint, arXiv:1910.14657, 2019.

Back to overview

zoom 6, Tuesday, 09:30

Split time-step schemes for McKean-Vlasov SDEs

Xingyuan Chen
University of Edinburgh, X.Chen-176@sms.ed.ac.uk

Gonçalo dos Reis
University of Edinburgh, G.dosReis@ed.ac.uk

We present two fully probabilistic Euler schemes, an explicit and a implicit split-step explicit Euler scheme, for the simulation of McKean-Vlasov Stochastic Differential Equations (MV-SDEs) with drifts of super–linear growth (in space) and random initial condition.

The general split-step scheme attains the standard 1/2 rate in stepsize and closes the gap left in [2] regarding efficient implicit methods and their convergence rate for this class of McKean-Vlasov SDEs.

The explicit Euler scheme, under certain structural conditions (but with non-constant diffusion matrix), draws on ideas from splitting operators to produce a order 1 Euler convergent method.

Several numerical examples are presented including a comparative analysis of other known algorithms for this class (taming [2] and adaptive time-stepping [3]) across parallel and non-parallel implementations.

This is joint work with X. Chen (U. of Edinburgh) [1].

[1] X. Chen and G. dos Reis *A flexible split-step scheme for MV-SDEs.* arXiv:2105.09688, 2021 May 21,
http://arxiv.org/abs/2105.09688

[2] G. dos Reis, S. Engelhardt and G. Smith *Simulation of McKean-Vlasov SDEs with super-linear growth.* IMA Journal of Numerical Analysis, draa099, 2021 Jan,
https://doi.org/10.1093/imanum/draa099

[3] C. Reisinger and W. Stockinger *An adaptive Euler-Maruyama scheme for McKean SDEs with super-linear growth and application to the mean-field FitzHugh-Nagumo model.* 2020 May, arXiv preprint arXiv:2005.06034
https://arxiv.org/abs/2005.06034

Back to overview

zoom 6, Tuesday, 10:00

Unbiased Estimation of the Solution to Zakai's Equation

Hamza Ruzayqat
King Abdullah University of Science and Technology, hamza.ruzayqat@kaust.edu.sa

Ajay Jasra
King Abdullah University of Science and Technology, ajay.jasra@kaust.edu.sa

The solution to Zakai's equation can be useful for model selection in statistics or as a solution of a particular stochastic partial differential equation in applied mathematics. In this work, we consider the non-linear filtering problem in continuous time and in particular the solution to Zakai's equation or the normalizing constant. We develop a methodology to produce finite variance, almost surely unbiased estimators of the solution to Zakai's equation. That is, given access to only a first order discretization of solution to the Zakai equation, we present a method which can remove this discretization bias. The

approach, under assumptions, is proved to have finite variance and is numerically compared to using a particular multilevel Monte Carlo method.

Back to overview

zoom 6, Tuesday, 10:30

Subordinated Gaussian random fields in elliptic PDEs

Robin Merkle
Universität Stuttgart, robin.merkle@mathematik.uni-stuttgart.de

Andrea Barth
Universität Stuttgart, andrea.barth@mathematik.uni-stuttgart.de

To model subsurface flow in uncertain heterogeneous fractured media an elliptic equation with a discontinuous stochastic diffusion coefficient - also called random field - may be used. We propose a new subordination approach to generate Lévy-type discontinuous random fields on a two-dimensional spatial parameter domain (see [1]). The resulting subordinated Gaussian random fields are used in the diffusion coefficient of an elliptic model equation. In this work, we use (multilevel) Monte Carlo - Finite Element methods to approximate the mean of the solution to the described stochastic elliptic PDE. We investigate the weak convergence of the problem, where we equilibrate all error contributions resulting from the (pathwise) Finite Element approximation and the approximation of the diffusion coefficient itself based on the results from [2].

[1] A. Barth and R. Merkle. *Subordinated Gaussian random fields.* submitted to Advances in Applied Probability, 2020

[2] A. Barth and R. Merkle. *Subordinated Gaussian random fields in elliptic PDEs.* submitted to Stochastics and Partial Differential Equations: Analysis and Computations, 2020

Contributed talks on MCQMC in Statistics

Back to overview

zoom 6, Wednesday, 16:00

On the use of the combined kernel-projection statistical estimator and randomized projection estimator

Natalya Tracheva
Laboratory of Monte Carlo methods, Institute of Computational Mathematics and Mathematical Geophysics SB RAS,
Mechanics and Mathematics Departmen, Novosibirsk State University, Russia, tnv@osmf.sscc.ru

In this talk, I would like to discuss the problems that arise in the practical applications of two particular statistical estimators. One of them is the randomized projection estimator which is based on the projection expansion on the orthonormal polynomial basis. Though this estimator seems to be the promising one and has been applied to solve a number of problems [1, 2, 3], the conversion rates vary sufficiently depending on the basis that has been chosen. Thus it is hard to give general

recommendations on the practical application of such an estimator, especially when we consider two-dimensional problems. Another estimation in consideration is a combined kernel-projection statistical estimator that was suggested in work [4] for the two-dimensional distribution density. It was constructed in the following way: for one of the variables the classical one-dimensional kernel estimator is formed and for other – the projection estimator. The optimal parameters for such an estimator were obtained in [5] within the assumptions made about the convergence rate of the orthogonal decomposition in use. I planning to illustrate the difference in two approaches on the two-dimensional problem of estimation of the bidirectional angle distribution of the polarized radiation flux transmitted through and backscattered by the scattering layered substance.

[1] N. V. Tracheva, S. A. Ukhinov. On the evaluation of spatial–angular distributions of polarization characteristics of scattered radiation. *Statistical Papers*, 59 (4): 1541–1557, 2018.

[2] N. V. Tracheva, S. A. Ukhinov. A new Monte Carlo method for estimation of time asymptotic parameters of polarized radiation. *Mathematics and Computers in Simulation*, 161: 84–92, 2019.

[3] S. V. Rogazinsky. Statistical modelling algorithm for solving the nonlinear Boltzmann equation based on the projection method. *Russian Journal of Numerical Analysis and Mathematical Modelling*, 32 (3): 197–202, 2017.

[4] G. A. Mikhailov, N. V. Tracheva, S. A. Ukhinov. New Statistical Kernel-Projection Estimator in the Monte Carlo Method. *DOKLADY MATHEMATICS*, 102 (1): 313–317, 2020.

[5] G. A. Mikhailov, N. V. Tracheva, S. A. Ukhinov. A new kernel-projective statistical estimator in the Monte Carlo method. *Russian Journal of Numerical Analysis and Mathematical Modelling*, 35 (6): 341–353, 2020.

Back to overview

Victim of Your (Customer's) Own Success

Nathan Yang
Cornell University, ncy6@cornell.edu

Yikun Jiang
UC Berkeley, jiangyk@berkeley.edu

Under the context of digital health, we explore the potential opportunity costs associated with giving customers a sense of success. Using large-scale data from a popular mobile fitness application, we show that many users who enjoy certain past successes need not necessarily upgrade to the premium version, or even continue using the app. These observed dynamics motivate a structural model of self-control, where users might exhibit present-bias (i.e., quasi-hyperbolic discounting). The model estimates reveal that improvements in the app's ability to help users lose weight do not necessarily increase their incentives to upgrade.

Back to overview

zoom 6, Wednesday, 19:00

Scaled lattice rules for integration over $\mathbb{R}^d$

Yuya Suzuki
NTNU Norway, yuya.suzuki@ntnu.no

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

In this talk, we show that by simply scaling lattice rules from the unit cube $[0,1]^d$ to properly sized boxes on $\mathbb{R}^d$, taking into account all errors, we can achieve higher-order convergence in approximating an integral on $\mathbb{R}^d$ where the order of convergence is close to the smoothness of the integrand function in a certain Sobolev space of dominating mixed smoothness. In particular, we show that the method of adding Bernoulli polynomials to a function to make it "periodic" on a box without changing its integral value over the box, is equivalent to an orthogonal projection from a well chosen Sobolev space of dominating mixed smoothness to an associated periodic Sobolev space of the same dominating mixed smoothness, which we call a Korobov space. We also conduct numerical experiments comaparing with other methods such as (i) direct product of Gauss–Hermite quadrature, (ii) Sparse grid based on Gauss–Hermite, (iii) and scaled interlaced Sobol' sequence.

Back to overview

zoom 6, Wednesday, 19:30

Importance sampling and randomized quasi-Monte Carlo

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

Sifan Liu
Stanford University, sfliu@stanford.edu

Scrambled net integration can attain a root mean squared error (RMSE) of $O(n^{-3/2}\log(n)^{d-1})$ for d-dimensional problems that are smooth enough [6]. Singular integrands f over $[0,1]^d$ such as those that arise when simulating Gaussian random variables pose a severe challenge. When $f \in L^2[0,1]^d$, the RMSE is $o(n^{-1/2})$ but not necessarily $o(n^{-3/2+\epsilon})$. Chelson [2] proposed to combine importance sampling with quasi-Monte Carlo and followup works include [1, 3, 4, 5]. We present some theoretical and empirical results on the difficulty of attaining an $o(n^{-3/2+\epsilon})$ RMSE via importance sampling and randomized quasi-Monte Carlo while also keeping the mean dimension of the resulting integrand from growing too quickly with the nominal dimension d.

[1] Aistleitner, C. and Dick, J. (2015). Functions of bounded variation, signed measures, and a general Koksma-Hlawka inequality. *Acta Arithmetica*, 167(2):143–71.

[2] Chelson, P. (1976). *Quasi-random techniques for Monte Carlo methods*. PhD thesis, The Claremont Graduate School.

[3] Chopin, N. and Ridgway, J. (2017). Leave Pima Indians alone: binary regression as a benchmark for Bayesian computation. *Statistical Science*, 32(1):64–87.

[4] Hörmann, W. and Leydold, J. (2005). Quasi importance sampling. Technical Report 57, WU Vienna.

[5] Hörmann, W. and Leydold, J. (2007). Importance sampling to accelerate the convergence of quasi-Monte Carlo. Technical Report 49, WU Vienna.

[6] Owen, A. B. (1997a). Monte Carlo variance of scrambled net quadrature. *SIAM Journal of Numerical Analysis*, 34(5):1884–1910.

Back to overview

zoom 6, Wednesday, 20:00

Kernel Thinning

Raaz Dwivedi
University of California, Berkeley, raaz.rsk@berkeley.edu

Lester Mackey
Microsoft Research New England, Stanford University, lmackey@stanford.edu

We introduce kernel thinning, a simple algorithm for generating better-than-Monte-Carlo approximations to target distributions $\mathbb{P}$ on $\mathbb{R}^d$ for estimating function integrals in a reproducing kernel Hilbert space. Kernel thinning is an online algorithm which given an input sequence of size n outputs a coreset of size $n^{\frac{1}{2}}$ in $\mathcal{O}(n^2)$ computation time and requires $\mathcal{O}(n \min(n, d))$ storage. When the input sequence admits an integration error of $\mathcal{O}(n^{-\frac{1}{2}})$, the integration error for the kernel thinning coreset is bounded as $\mathcal{O}_d(n^{-\frac{1}{2}}\sqrt{\log n \log\log n})$ for compactly supported $\mathbb{P}$ and $\mathcal{O}_d(n^{-\frac{1}{2}}\sqrt{(\log n)^{d+1}\log\log n})$ for sub-exponential $\mathbb{P}$. In comparison, $n^{\frac{1}{2}}$ i.i.d. Monte Carlo sample points from $\mathbb{P}$ yield only $\Omega(n^{-\frac{1}{4}})$ integration error. Our guarantees for sub-exponential distributions match the classical rates of Quasi-Monte Carlo point sets [2] on the unit cube but apply to general target distributions on $\mathbb{R}^d$ and a variety of popular kernels. We illustrate the benefits of kernel thinning for various infinite-dimensional kernels including Gaussian, Matérn, B-spline, Wendland, inverse multiquadric, and hyperbolic secant kernels. En route, we introduce efficient strategies to generate near-optimal L^∞ coresets of size $n^{\frac{1}{2}}$ with L^∞ error $\mathcal{O}(\sqrt{\frac{d}{n}\log n \log\log n})$ for a wide range of kernels and $\mathbb{P}$ using a Hilbert space generalization of the self-balancing walk of Alweiss et al [1].

[1] R. Alweiss, Y. P. Liu, M. Sawhney. *Discrepancy Minimization via a Self-Balancing Walk* . arXiv preprint *arXiv:2006.14009*, 2020.

[2] A. B. Owen and P. W. Glynn. *Monte Carlo and Quasi-Monte Carlo Methods: MCQMC 2016, Stanford, CA, August 14-19*. Springer, Volume 241, 2018.

Back to overview

zoom 6, Wednesday, 20:30

Monte Carlo Integration for Complex Exponential Integrals

Adam Knapp
U. S. Naval Research Laboratory, adam.knapp@nrl.navy.mil

Among the integral forms presenting severe challenges to the Monte Carlo integration methods are high-dimensional, highly oscillatory integrals of the form,

$$I = \int d\mathbf{X}\, f(\mathbf{X}) e^{i\omega g(\mathbf{X})},$$

Intuitively, this difficulty in numerical integration may be understood to originate from the value of the integral being the difference between the positive and negative parts of the integral which may be decrease as some function of the dimension of the integrand making statistical error estimation difficult. Despite these challenges, this class of integrals is commonly encountered in engineering and physics applications like electromagnetic scattering, time dependent quantum mechanical calculations and the multidimensional Fourier analysis-related integrals used in signal processing.

We show that integrals of this form may be exactly transformed into a related integral of the form,

$$I = \int d\mu\, d\mathbf{X}\, s(\mu) f(\mathbf{X}) p(\mu, \mathbf{X}),$$

such that $p(\mu, \mathbf{X}) \geq 0$ and may act as a sampling function and $s(\mu)$ is a single variable, mildly oscillatory integral of fewer than 10 sign changes on compact support. The resultant high-dimensional integral may be numerically integrated with Monte Carlo methods with the variance being particularly amenable to control variate methods optimized for this particular application.

Special Session Recent Advances on the Laplace Approximation and Related Sampling Methods (part A)

Back to overview

zoom 6, Thursday, 15:30

Laplace-based Importance Sampling

Claudia Schillings
Universität Mannheim, c.schillings@uni-mannheim.de

Björn Sprungk
Technische Universität Bergakademie Freiberg, bjoern.sprungk@math.tu-freiberg.de

Philipp Wacker
Friedrich-Alexander-Universität Erlangen-Nürnberg, wacker@math.fau.de

A frequent task in inference is the one of sampling from a measure, for example if we have constructed a posterior measure from data and we want to interpret it. There is a large number of sampling methods and in this talk we will take a closer look at importance sampling. Importance sampling's performance is strongly dependent on a suitable choice of reference measure. Broadly speaking, the reference measure should be as similar to the measure of interest as possible while still being elementary enough such that we can explicitly sample from it. One possible choice for the reference measure is the prior measure but

it is readily observed that this does not work for high-dimensional problems. For this reason we propose choosing the Laplace approximation to the measure of interest: It is both relatively close to the posterior measure and can be explicitly sampled as it is a Gaussian measure. In this talk, we will talk about the Laplace approximation, about importance sampling including scenarios where prior-based importance sampling does not work well and how Laplace-based importance sampling can improve computation.

[1] C. Schillings, B. Sprungk, and P. Wacker. *On the convergence of the Laplace approximation and noise-level-robustness of Laplace-based Monte Carlo methods for Bayesian inverse problems*. Numerische Mathematik (2020), Springer.

Back to overview

zoom 6, Thursday, 16:00

Non-asymptotic error estimates for the Laplace approximation in Bayesian inverse problems

Tapio Helin
LUT University, tapio.helin@lut.fi

Remo Kretschmann
LUT University, remo.kretschmann@lut.fi

We consider the Gaussian posterior approximation according to Laplace' method in Bayesian inverse problems.
While the asymptotic behaviour of the Laplace approximation in the small noise limit has been studied in previous works, one is, in practice, often interested in bounding the approximation error for a given, fixed noise level. For this purpose, asymptotic estimates are of limited use. On the one hand, they may not be sharp enough for realistic noise levels. On the other hand, the nonlinearity of the forward mapping may cause a large approximation error even for low noise levels without this being reflected in an asymptotic estimate. We provide sharp, non-asymptotic error estimates for the Laplace approximation that hold for common noise levels, are fast to compute, and quantify the influence of the nonlinearity of the problem. Other Gaussian approximations have been shown to converge to the posterior distribution in the small noise limit even if the number of parameters is allowed to tend to infinity with a certain rate. We prove a similar result for the Laplace approximation and quantify the influence of the problem dimension on the approximation error for a fixed noise level.

Back to overview

zoom 6, Thursday, 16:30

Wasserstein convergence of the Laplace approximation

Mareike Hasenpflug
Universität Göttingen, mareike.hasenpflug@uni-goettingen.de

For increasing n we study the difference of a probability measure μ_n with density determined by $\exp\big(-n\phi(x)\big)$ to its weak limit measure provided in [1]. If ϕ has a unique global minimum at x_0, we show that μ_n converges with $n^{-1/2}$ to the Dirac measure δ_{x_0} in terms of the Wasserstein distance. Estimates and discussions beyond the case of a unique global minimum are provided. Our investigation is motivated by a convergence result for the Laplace approximation w.r.t. the Hellinger distance stated in [2]. Related to that we suggest a generalization of the Laplace approximation to a Gaussian mixture and discuss its Wasserstein convergence if the setting is generalized to ϕ having finitely many global minima.

[1] Hwang, Chii-Ruey, *Laplace's method revisited: weak convergence of probability measures*, Ann. Probab., 8(6):1177-1182, 1980.

[2] Schillings, C., Sprungk, B., Wacker, P., *On the convergence of the Laplace approximation and noise-level-robustness of the Laplace-based Monte Carlo methods for Bayesian inverse problems*, Numer. Math., 145(4):915-971, 2020.

Special Session Recent Advances on the Laplace Approximation and Related Sampling Methods (part B)

Back to overview

zoom 6, Thursday, 19:00

Laplace Approximation for Addressing Catastrophic Forgetting in Few-Shot Problems

Pauching Yap
University College London, p.yap@cs.ucl.ac.uk

Hippolyt Ritter
University College London,

David Barber
University College London,

Laplace approximation along with Bayesian online learning (BOL) [1] is capable of overcoming catastrophic forgetting in large-scale supervised classification [2]. This benefit can be extended to address catastrophic forgetting in the few-shot classification setting by coupling with Bayesian online meta-learning (BOML) instead [3]. BOML combines BOL and meta-learning to consider the posterior of the meta-parameters rather than the posterior of the model parameters in BOL.

Laplace approximation suggests using a Gaussian approximate posterior in both BOL and BOML by Taylor expanding the log-posterior around a mode up to the second order. This results in the requirement for a Hessian approximation in order to update the precision of the Gaussian approximate posterior. For BOL in large-scale supervised classification, we can use a Kronecker-factored block-diagonal Hessian approximation. The meta-learning element in BOML, however, breaks the desirable Kronecker-factored structure in Hessian approximation.

We present the necessary adjustments to the Hessian approximation in BOML to retain the desirable block-diagonal Kronecker-factored structure, and illustrate that Laplace approximation with BOML is indeed able to overcome catastrophic forgetting in few-shot problems. This talk is based on the paper [3].

[1] M. Opper *A Bayesian Approach to Online Learning 2016, Online Learning in Neural Networks.* Cambridge University Press, 1998.

[2] H. Ritter, A. Botev and D. Barber *Online Structured Laplace Approximations for Overcoming Catastrophic Forgetting.* Advances in Neural Information Processing Systems 31, 2018.

[3] P. Yap, H. Ritter and D. Barber *Addressing Catastrophic Forgetting in Few-Shot Problems.* Proceedings of the 38th International Conference on Machine Learning, 2021.

Back to overview

zoom 6, Thursday, 19:30

Scalable Laplace Approximation for Bayesian Optimal Experimental Design

Umberto Villa
Washington University in St Louis, uvilla@wustl.edu

Omar Ghattas
University of Texas at Austin, omar@oden.utexas.edu

Optimal experimental design (OED) for Bayesian nonlinear inverse problems governed by partial differential equations (PDEs) is an extremely challenging problem. First, the parameter to be inferred is often a spatially correlated field and it leads—after discretization—to a high dimensional parameter space. Second, the forward model is often extremely complex and computationally expensive to solve. A common objective function for OED is the expected information gain (EIG). Naïve evaluation of EIG is unfeasible for large-scale problems due to the large number of samples required in the double-loop Monte Carlo.

To overcome these difficulties, we invoke an approximation of the EIG based on the Laplace approximation of the posterior. Each evaluation of the objective function then requires computing a sample average approximation (over possible realization of the data) of the information gain (IG) between the Laplace approximation and the Gaussian prior distribution. An analytical formula for the IG is available and it involves the log-determinant and trace of the posterior covariance operator. Randomized eigensolver algorithms allows us to efficiently estimate such invariants at a cost that is independent of the dimension of the parameter space, thus allowing for a scalable evaluation of the objective function. Variational adjoint methods, implemented in `hIPPYlib` [1], are then used to efficiently compute the gradient of the OED objective function. A numerical example related to single phase subsurface flow will demonstrate scalability of the proposed method with respect to both parameter and design dimension.

[1] U. Villa, N. Petra, O. Ghattas *hIPPYlib: An Extensible Software Framework for Large-Scale Inverse Problems Governed by PDEs; Part I: Deterministic Inversion and Linearized Bayesian Inference.* ACM Trans. Math. Softw., 47(2), April 2021.

Back to overview

zoom 6, Thursday, 20:00

Greedy variational inference and preconditioning of sampling schemes using transport

Youssef Marzouk
MIT, ymarz@mit.edu

Transformations of measure underlie many powerful tools for Bayesian inference and Monte Carlo simulation. A central idea is to deterministically couple a probability measure of interest with a tractable "reference" measure (e.g., a standard Gaussian). In principle, this approach enables perfect simulation; in practice, the approximation of suitable transport maps provides a flexible framework that encompasses both sampling and variational inference.

In this setting, we will discuss a framework for solving high-dimensional Bayesian inference problems using structured low-dimensional transport maps or flows. The action of these maps is confined to a low-dimensional subspace, and the subspace is identified by minimizing an upper bound on the Kullback–Leibler divergence. Our framework provides a principled way of identifying and exploiting low-dimensional structure in an inference problem. It focuses the expressiveness of a transport map along the directions of most significant discrepancy from the posterior, and can be used to build deep compositions of lazy

maps, where low-dimensional projections of the parameters are iteratively transformed to match the posterior. We prove weak convergence of the generated sequence of distributions to the posterior, and demonstrate the benefits of the framework on challenging inference problems in machine learning and differential equations.

We will also discuss the use of such maps in "preconditioning" asymptotically exact sampling schemes such as Markov chain Monte Carlo (MCMC). In particular, we will elucidate new links between transport maps and Riemannian manifold MCMC.

Contributed talks on Discrepancy and Dispersion

Back to overview

zoom 6, Friday, 09:30

Koksma-Hlawka inequality on finite posets

Kosuke Suzuki
Hiroshima University, kosuke-suzuki@hiroshima-u.ac.jp

The Koksma-Hlawka inequality, which plays an important role in quasi-Monte Carlo method, bounds the integration error of an integrand f using a set of sample points P by the product of the Hardy-Krause variation of f and the star-discrepancy of P. This stems from Abel summation formula or what we call "Koksma-Hlawka inequality on grids". As a generalization of this, we develop the Koksma-Hlawka inequality on finite posets. To this aim we introduce counterparts of the star-discrepancy and the Hardy-Krause variation on finite posets in a natural way via Möbius transform.

Back to overview

zoom 6, Friday, 10:00

A Construction of Short-Period Tausworthe Generators with Low Discrepancies over $\mathbf{F}_b$

Shin Harase
Ritsumeikan University, harase@fc.ritsumei.ac.jp

A one-dimensional sequence $u_0, u_1, u_2, \ldots \in [0,1)$ is said to be completely uniformly distributed (CUD) if overlapping s-blocks $(u_i, u_{i+1}, \ldots, u_{i+s-1})$, $i = 0, 1, 2, \ldots$, are uniformly distributed for every dimension $s \geq 1$. This concept naturally arises in Markov chain quasi-Monte Carlo.

Recently, Harase [1] searched for short-period Tausworthe generators over the two-element field $\mathbf{F}_2 := \{0, 1\}$ that approximate CUD sequences in terms of the t-value, which is a central criterion in the theory of (t, m, s)-nets and (t, s)-sequences. The key technique was to use the continued fraction expansion. Namely, Tausworthe generators can be viewed as polynomial Korobov lattice point sets with a denominator polynomial $p(x)$ and a numerator polynomial $q(x)$, and hence, the t-value is zero (i.e., optimal) if and only if the partial quotients in the continued fraction of $q(x)/p(x)$ are all of degree one. As a result, Tausworthe generators with t-values zero for dimension $s = 2$ and small for $s \geq 3$ were obtained. However, in the case of $\mathbf{F}_2$, we have the following restrictions:

(a) There exists no maximal-period Tausworthe generator with t-value zero for dimension $s = 3$ [2].

(b) For each irreducible polynomial $p(x)$, there are exactly two polynomials $q(x)$ for which the partial quotients of the continued fraction expansion of $q(x)/p(x)$ all have degree one [3].

In fact, in the case of the finite fields $\mathbf{F}_b$ of order $b \geq 3$, we can construct maximal-period Tausworthe generators with t-value zero for $s = 3$ for some m. Additionally, the number of $q(x)$ is expected to be much more than two in general.

In this talk, we therefore explore a construction of Tausworthe generators over $\mathbf{F}_b$ with t-value zero for $s = 3$, and report some numerical results.

[1] S. Harase. A table of short-period Tausworthe generators for Markov chain quasi-Monte Carlo. J. Comput. Appl. Math. 384 (2021), 113136, 12 pp.

[2] H. Kajiura, M. Matsumoto, and K. Suzuki. Characterization of matrices B such that (I, B, B^2) generates a digital net with t-value zero. Finite Fields Appl. 52 (2018) 289-300.

[3] J.P. Mesirov and M.M. Sweet, Continued fraction expansions of rational expressions with irreducible denominators in characteristic 2. J. Number Theory 27 (2) (1987) 144-148.

Back to overview

zoom 6, Friday, 10:30

A dispersion toolkit

Benjamin Sommer
Max-Planck-Institut für Informatik, besommer@mpi-inf.mpg.de

Thomas Lachmann
Johannes Kepler Universität Linz, thomas.lachmann@jku.at

Jaspar Wiart
Johannes Kepler Universität Linz, jaspar.wiart@jku.at

Gurprit Singh
Max-Planck-Institut für Informatik, gsingh@mpi-inf.mpg.de

Dispersion measures the greatest empty region amidst an n element point set $\mathcal{P}$ in d-dimensions. Numerically measuring dispersion is a demanding task for large n and d, both regarding computational as well as algorithmic complexity. Currently however, researchers need to implement dispersion measures on their own from scratch, because no dedicated tool exists to date, to the best of our knowledge. We are developing a high-performance, multi-threaded, cross-platform framework to solve this problem, tailored to workstations and to computing clusters. It comprises implementations of previously suggested algorithms computing dispersion using modern C++, to establish a common ground regarding comparability as well as algorithmic research. Currently, we are supporting exact and $(1 - \epsilon)$-approximate algorithms for $d = 2$ and $d = 3$, while this list of implementations is going to be expanded in the future.

Another contribution is an experimental algorithm measuring dispersion based on axes-aligned hyperboxes in $d \geq 2$, bounded by the problem domain. The idea is to compute the locally largest hyperbox being bounded by $\mathbf{p}_i, \forall \mathbf{p}_i \in \mathcal{P}$, out of which the biggest one is taken. This local dispersion is computed along each axis independently, based on a grow & shrink algorithm. With the assumption that all point coordinates differ, this algorithm leads to computational complexity of $\mathcal{O}(dn^2)$ in time and of $\mathcal{O}(dn)$ in space, while its extent of approximation remains to be investigated in theory. Based on this algorithm, we additionally provide an implementation of an optimisation scheme using gradient descent to numerically find a point set with minimal dispersion, contributing towards the understanding of dispersion bounds.

Given a modified Fibonacci lattice $\mathcal{P}_\mathrm{f}$ in $d = 2$, for finite n, we succeeded in finding a point set with notably smaller dispersion than the best known point set to date. However, this improvement appears to vanish for $n \to \infty$, and our numerical experiments indicate $\lim_{n\to\infty} \min_{\mathcal{P}} n \operatorname{disp}(\mathcal{P}) \approx 1.8944.. \approx \frac{\varphi^3}{\sqrt{5}}$, $\varphi = \frac{\sqrt{5}+1}{2}$, using dispersion based on axes-aligned rectangles constraint to $[0, 1]^2$. In addition, this toolkit already contains several utilities, most notably a statistics module to estimate confidences needed for stochastic point sets, an analysis module helping to gain understanding concerning convergence of dispersion, a visualisation module, as well as a transduction module to ease interoperability with other software, for instance Mathematica's matrices. Finally, we provide a detailed manual of how functionalities of this toolkit are to be used to help everyone conducting research regarding point set dispersion.

Index