MCM23: Book of Abstracts
Version: June 26, 2023

Contents

1 Monday 26th June

1.1 Keynote speakers

Speaker: Gareth O. Roberts

PDMPs: understanding limiting and transient behaviour

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

Piecewise deterministic Markov processes (PDMPs) have recently become exciting new alternatives for the MCMC user. Among their appealing properties are the fact that they are non-reversible, have huge flexibility, and they offer principled (ie approximation-free) rapid implementation through the use of subsampling in contexts where full target densities are expensive to compute. This talk will give an overview of some recent theoretical and methodological developments for PDMPs which have contributed to our understanding of these methods. For most of the presentation, I will concentrate on the two most widely adopted PDMP methods: the Zig-Zag (ZZ) and the Bouncy Particle Sampler (BPS). The first part of the talk will focus on limit results in stylised cases which allow us to understand the scaling behaviour of the mixing of ZZ and BPS as dimension grows (where ZZ performs better), and as anisotropy grows (where BPS seems more robust). These results use weak convergence techniques to infinite-dimensional limit processes, and (in the latter case) multi-scale methods for understanding how fast and slow components converge at different rates. In the second half the talk, I will focus more on how the algorithms behave in their transient phase. This is particularly relevant for PDMPs as simulation costs are often disproportionately high for PDMPs in their transient phase, so that understanding/reducing the computational effort involved in the transient phase is particularly important. Geometric and polynomial ergodicity results will be presented as well as methodological improvements which regain geometric ergodicity for heavy-tailed targets: the Speed-Up Zig-Zag and the Stereographic Projection Sampler.
Different parts of this work are joint work with Joris Bierkens, Kengo Kamatani, Giorgos Vasdekis, Jun Yang and Krys Latuszynski.

Speaker: Madalina Deaconu

Walk on moving spheres methods - new perspectives

Madalina Deaconu
Inria Nancy-Grand Est & IECL, madalina.deaconu@inria.fr

The aim of this talk is to introduce new numerical methods for the pathwise approximation and the hitting times of some boundaries approximation for a class of stochastic processes. We will start by introducing the walk on moving spheres method for Bessel processes and Brownian motions. This method exhibits a procedure which efficiently approximates the hitting time and the hitting position of a class of stochastic differential equations and can be used in questions arising from finance, geophysics or neuroscience. Some applications to the initial boundary value problem and to the ε-strong approximation technique will be also discussed.
Furthermore, we give convergence theorems and present numerical results that permit to illustrate the efficiency and accuracy of these methods.

1.2 Parallel sessions - Morning

1.2.1 Amphitheater 25 - MS Stochastic Computation and Complexity

Speaker: Andreas Neuenkirch

The robustness of the Euler scheme for scalar SDEs with non-Lipschitz diffusion coefficients

Annalena Mickel
Mathematical Institute and DFG Research Training Group 1953, University of Mannheim, Germany, amickel@mail.uni-mannheim.de

Andreas Neuenkirch
Mathematical Institute, University of Mannheim, Germany, aneuenki@mail.uni-mannheim.de

We consider stochastic differential equations (SDEs) that are given by

$$dV_t = a(V_t)dt + (b(V_t))^{1-\gamma}\, dW_t, \qquad t \in [0,T], \tag{1}$$

where $V_0 = v_0 \in \mathbf{R}$ is deterministic, $W = (W_t)_{t\in[0,T]}$ is a Brownian motion and $\gamma \in \left(0, \frac{1}{2}\right]$. We assume that $a : \mathbf{R} \to \mathbf{R}$ and $b : \mathbf{R} \to [0,\infty)$ are globally Lipschitz continuous. Well-known examples that fall into the class of SDE (1) are the CIR process, the CEV process or the Wright-Fisher diffusion.
We analyze the equidistant Euler scheme for SDE (1). Among other results, we show L^1-convergence order $1/2 - \varepsilon$ in the discretization points (for $\varepsilon > 0$ arbitrarily small) if

$$\int_0^T \mathbf{E}\left[\frac{1}{b(V_t)^{2\gamma}}\right] dt < \infty. \tag{2}$$

Thus, the loss of Lipschitzness, i.e. $\gamma > 0$, for the diffusion coefficient can be compensated by an appropriate inverse moment condition, that is the dt-integrability of the 2γ-inverse moment of $b(V_t)$, $t \in [0,T]$. This result yields in particular a unifying framework for the above mentioned SDEs. For the CIR or Wright-Fisher process, condition (2) corresponds to the non-attainability of the boundaries of their support, while for the CEV process condition (2) is always fulfilled.

Speaker: Konstantinos Dareiotis

Asymptotic error distribution of the Euler method for stochastic differential equations with irregular drifts

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

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

Khoa Lê
University of Leeds, UK, k.le@leeds.ac.uk

In this talk we will establish a central limit theorem for the Euler-Maruyama scheme approximating multidimensional SDEs with elliptic Brownian diffusion, under very mild regularity requirements on the drift coefficients. When the drift is Hölder continuous, we show that the limiting law of the rescaled fluctuations around the true solution is characterised by the solution of a hybrid Young-Itô differential equation. When the drift has positive Sobolev regularity, this limit is characterised by the solution of a transformed SDE. Our result is an extension of the result of Kurtz–Protter (1991) in which SDEs with bounded differentiable coefficients were considered. To compensate for the lack of regularity of the drifts, we utilize the regularization effect from the non-degenerate noise.

Speaker: Verena Schwarz

Randomized Milstein algorithm for approximation of solutions of jump-diffusion SDEs

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

Verena Schwarz
University of Klagenfurt, Austria, verena.schwarz@aau.at

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

We investigate the error of the randomized Milstein algorithm for solving scalar jump-diffusion stochastic differential equations. We provide a complete error analysis under substantially weaker assumptions than those known in the literature. If the jump-commutativity condition is satisfied, we prove the optimality of the randomized Milstein algorithm by proving a matching lower bound. Moreover, we give some insight into the multidimensional case by investigating the optimal convergence rate for the approximation of jump-diffusion type Lévys' areas. Finally, we report numerical experiments that support our theoretical findings.

Speaker: André Herzwurm

On Upper and Lower Error Bounds for Strong Approximation of Scalar SDEs with Reflecting Boundary

Mario Hefter
RPTU Kaiserslautern, Germany, hefter@mathematik.uni-kl.de

André Herzwurm
TH Rosenheim, Germany, andre.herzwurm@th-rosenheim.de

Klaus Ritter
RPTU Kaiserslautern, Germany, ritter@mathematik.uni-kl.de

We study pathwise approximation of scalar stochastic differential equations with reflecting boundary at a single point in time or globally in time by means of methods that are based on finitely many observations of the driving Brownian motion. We establish upper and lower error bounds in terms of the average number of evaluations of the driving Brownian motion.

1.2.2 Room 101, corridor 24-25 - MS On the power of iid information for (non-linear) approximation

Speaker: Mario Ullrich

On recent advances in approximation based on iid data

Albert Cohen
LJLL, Sorbonne Université, Paris, albert.cohen@sorbonne-universite.fr

Matthieu Dolbeault
LJLL, Sorbonne Université, Paris, matthieu.dolbeault@sorbonne-universite.fr

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

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

We survey on recent developments in the area of approximation of functions based on function evaluations, with a special emphasis on random sampling and high-dimensional problems. In particular, we discuss how the theory of random matrices and the recent solution of the Kadison-Singer conjecture interact to provide a final solution to the problem of *power of function values* for L_2-approximation in the worst-case setting.

[1] M. Dolbeault, D. Krieg and M. Ullrich. *A sharp upper bound for sampling numbers in L_2.* accepted in Applied and Computational Harmonic Analysis, 2022.

[2] D. Krieg, P. Siedlecki, M. Ullrich and H. Wozniakowski. *Exponential tractability of L_2-approximation with function values.* accepted in Applied and Computational Harmonic Analysis, 2022.

[3] D. Krieg and M. Ullrich. *Function values are enough for L_2 approximation.* Foundations of Computational Mathematics 21, 2021.

[4] D. Krieg and M. Ullrich. *Function values are enough for L_2 approximation: Part II.* J. Complexity 66, 2021.

Speaker: Matthieu Dolbeault

Approximation with iid, reduced, or greedy sampling strategies

Albert Cohen
LJLL, Sorbonne Université, Paris, albert.cohen@sorbonne-universite.fr

Matthieu Dolbeault
LJLL, Sorbonne Université, Paris, matthieu.dolbeault@sorbonne-universite.fr

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

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

In this talk, we investigate sampling strategies for approximation of functions by weighted least-squares. Although a quasilinear sampling budget can already be achieved by iid random draws according to the adapted density [2, 4], further reductions of the needed number of points are possible, using the solution to the Kadison-Singer problem [1, 3]. However this involves a subsampling step, which is not algorithmically tractable. We show how greedy sampling methods can circumvent this defect, while attaining optimal sample sizes.

[1] A. Cohen and M. Dolbeault. *Optimal point-wise sampling for L^2 approximation.* Journal of Complexity, 68, 2022

[2] A. Cohen and G. Migliorati. *Optimal weighted least squares methods.* SMAI Journal of Computational Mathematics, 3, 2017

[3] M. Dolbeault, D. Krieg and M. Ullrich. *A sharp upper bound for sampling numbers in L_2.* accepted in Applied and Computational Harmonic Analysis, 2022

[4] D. Krieg and M. Ullrich. *Function values are enough for L_2 approximation.* Foundations of Computational Mathematics, 21, 2021

Speaker: Albert Cohen

Stable nonlinear inversion application and application to interface reconstruction from cell-averages

Albert Cohen
Sorbonne Université, Paris, France, albert.cohen@sorbonne-universite.fr

In this lecture, we present a general framework for solving inverse problems using nonlinear approximation spaces. The main principles build up on the so called Parametrized Background Data Weak method (PBDW) which can be thought as a linear counterpart. As an application we study the reconstruction of sharp interfaces from cell average at coarse resolutions for which linear methods are known to be uneffective. We discuss the convergence rates of these reconstructions and their optimality.

Speaker: Sebastian Moraga

Optimal learning of infinite-dimensional holomorphic functions from i.i.d. samples

Ben Adcock
Department of Mathematics, Simon Fraser University, Burnaby BC, Canada, ben_adcock@sfu.ca

Nick Dexter
Department of Scientific Computing, Florida State University, Tallahassee, USA, nick.dexter@fsu.edu

Sebastian Moraga
Department of Mathematics, Simon Fraser University, Burnaby BC, Canada, smoragas@sfu.ca

Over the last decade, learning functions in infinite dimensions have received increasing attention in Computer Science and Engineering (CSE). This is primarily due to the relevance of such functions in applications where specific quantities of interest involve solving parametric differential equations (DEs) in biology, engineering and physics, e.g., in computational uncertainty quantification (UQ). While acquiring accurate and reliable approximations of such functions is inherently difficult, current benchmark methods exploit the properties of classes of holomorphic functions to get certain algebraic convergence rates in infinite dimensions with respect to the number of samples m. Our work focuses on providing theoretical guarantees for holomorphic function approximation, indicating that these algebraic rates are the best possible for Banach-valued functions in infinite dimensions. First, we establish lower bounds using a reduction to a discrete problem along with m-widths, Gelfand widths and Kolmogorov widths theory for both known and unknown parametric dependence. In particular, we use information regarding the anisotropic regularity of the function to obtain optimal rates, up to log factors, in the so-called *known* and *unknown anisotropy* cases. In addition, our results show how to attain (close to) the lower bounds for these class functions using independent and identically distributed pointwise samples, polynomial approximation, and machine learning with deep neural networks.

1.2.3 Room 102, corridor 15-25 - MS Numerics for stochastic partial differential equations

Speaker: Guillaume Dujardin

Numerical methods for the nonlinear stochastic Manakov system

Guillaume Dujardin
Inria Lille Nord-Europe, France, guillaume.dujardin@inria.fr

The stochastic Manakov equation is an SPDE system arising in the study of pulse propagation in randomly birefringent optical fibers. We propose and analyse a Lie–Trotter splitting scheme for the time integration of this system. First, we prove that the strong order of the numerical approximation is 1/2 if the nonlinear term in the system is globally Lipschitz. Then, we show that the splitting scheme has convergence order 1/2 in probability and almost sure order 1/2− in the case of a cubic nonlinearity. We provide several numerical experiments illustrating the aforementioned results and the efficiency of the Lie-Trotter splitting scheme. Finally, we numerically investigate the possible blowup of solutions for some power-law nonlinearities. This is a joint work with A. Berg (University of Umeå) and D. Cohen (Chalmers).

Speaker: Ludovic Goudenège

Tamed Euler scheme for SPDE with distributional drift

Ludovic Goudenège
CNRS, France, goudenege@math.cnrs.fr

Alexandre Richard
CentraleSupélec, MICS, France, alexandre.richard@centralesupelec.fr

El Mehdi Haress
CNRS, France, el-mehdi.haress@centralesupelec.fr

Using regularization by noise technics and stochastic sewing lemma from [Le20], we will recall how we can define solutions to SDEs with distributional drift driven by fractional Brownian motion. The aim of this talk is to extend these technics to SPDE with distributional drift driven by white noise in the spirit of [BDG22].
We build the solutions as the limit of numerical solutions of tamed Euler schemes by combining the time-step convergence and the taming. An optimization in the taming permits obtaining a speed of convergence that recovers the known rates in the bounded regular cases [BDG21, DGL21, DAGI19] but extends the speed of convergence to drifts with negative regularities given constraints on the Hölder regularity of the noise [GHR22].
We will also present numerical simulations based on Monte-Carlo sampling and illustrate the difficulty to recover the expected theoretical speed of convergence.

[BDG21] O. Butkovsky, K. Dareiotis, and M. Gerencsér Approximation of SDEs: a stochastic sewing approach. Probab. Theory Related Fields, 181(4), 975–1034, 2021.

[BDG22] O. Butkovsky, K. Dareiotis, and M. Gerencsér Optimal rate of convergence for approximations of spdes with non-regular drift. arXiv preprint arXiv:2110.06148

[BW20] O. Butkovsky and F. Wunderlich Asymptotic strong Feller property and local weak irreducibility via generalized couplings *arXiv preprint arXiv:1912.06121*, 2020.

[GHR22] Ludovic Goudenège, El Mehdi Haress et Alexandre Richard. Numerical approximation of SDEs with fractional noise and distributional drift. *hal-03715427v1*, 2022.

[Le20] K. Lê A stochastic sewing lemma and applications. Electronic Journal of Probability, 25, 1–55, 2020.

[DGL21] K. Dareiotis, M. Gerencsér and K. Lê. Quantifying a convergence theorem of Gyongy and Krylov. arXiv preprint arXiv:2101.12185, 2021

[DAGI19] T. De Angelis, M. Germain and E. Issoglio. A numerical scheme for stochastic differential equations with distributional drift. arXiv preprint arXiv:1906.11026, 2019

Speaker: Gabriel Lord

Adaptive/Tamed methods for SPDEs with additive noise

Gabriel Lord
Radboud University, Nijmegen, NL, gabriel.lord@ru.nl

Stuart Campbell
Heriot-Watt University, Edinburgh, UK, sc58@hw.ac.uk

We consider strong convergence of tamed exponential and adaptive time-step methods for SPDEs with non-globally Lipschitz diffusion terms. A typical example would the the stochastic Allen-Cahn or Swift-Hohenberg equations. Although not designed to control local error directly we observe in numerical simulations an improved error constant and that the adaptivity is good at capturing the local behaviour. The talk will introduce the issue with standard Euler methods for non-globally Lipschitz coefficients and introduce two different approaches for the adaptivity as well as the tamed exponential method. We will briefly examine the proof of strong convergence and compare numerically methods.

Speaker: Kerstin Schmitz

Convergence of a finite-volume scheme for a stochastic heat equation with a multiplicative Lipschitz noise

Kerstin Schmitz
Universität Duisburg-Essen, Germany, kerstin.schmitz@uni-due.de

We study an approximation by a finite-volume scheme in space and a semi-implicit discretization in time for a stochastic heat equation driven by a Lipschitz continuous multiplicative noise. Here the nonlinearity in the stochastic integral leads to a lack of compactness. More precisely, the weak convergence of our finite-volume approximations that we obtain by a priori estimates is not enough to identify the weak limit in the nonlinear term in the stochastic integral. Therefore we use the stochastic compactness method based on Skorokhod's representation theorem to get the convergence of our finite-volume approximations to a martingale solution. By a pathwise uniqueness argument we then get stochastically strong convergence to the unique variational solution of our parabolic problem.

1.2.4 Room 104, corridor 15-25 - MS Piecewise Deterministic Markov Processes and Related Topics

Speaker: George Deligiannidis

Non-Reversible Parallel Tempering: a Scalable Highly Parallel MCMC Scheme

Saifuddin Syed
Department of Statistics, University of Oxford, UK, saifuddin.syed@stats.ox.ac.uk

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

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

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

Parallel tempering (PT) methods are a popular class of Markov chain Monte Carlo schemes used to sample complex high-dimensional probability distributions. They rely on a collection of N interacting auxiliary chains targeting tempered versions of the target distribution to improve the exploration of the state-space. We provide here a new perspective on these highly parallel algorithms and their tuning by identifying and formalizing a sharp divide in the behaviour and performance of reversible versus non-reversible PT schemes. We show theoretically and empirically that a class of non-reversible PT methods dominates its reversible counterparts and identify distinct scaling limits for the non-reversible and reversible schemes, the former being a piecewise-deterministic Markov process and the latter a diffusion. These results are exploited to identify the optimal annealing schedule for non-reversible PT and to develop an iterative scheme approximating this schedule. We provide a wide range of numerical examples supporting our theoretical and methodological contributions. The proposed methodology is applicable to sample from a distribution π with a density L with respect to a reference distribution π_0 and compute the normalizing constant. A typical use case is when π_0 is a prior distribution, L a likelihood function and π the corresponding posterior.

[1] S. Syed, A. Bouchard-Côté, G. Deligiannidis, A. Doucet (2022). *Non-reversible parallel tempering: a scalable highly parallel MCMC scheme.* Journal of the Royal Statistical Society (Series B), 84(2), 321-350

Speaker: Kengo Kamatani

Scaling of Piecewise Deterministic Monte Carlo for Anisotropic Targets

Joris Bierkens
Delft Institute of Applied Mathematics, The Netherlands, Joris.Bierkens@tudelft.nl

Kengo Kamatani
The Institute of Statistical Mathematics, Japan, kamatani@ism.ac.jp

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

Piecewise deterministic Markov processes are continuous-time Markov processes driven by deterministic flows and jumps. Recently, the piecewise deterministic Markov process has attracted the attention of the Monte Carlo community as a possible alternative to traditional Markov chain Monte Carlo methods. The Zig-Zag sampler and the Bouncy particle sampler are widely used as examples of the piecewise deterministic Monte Carlo methodology.
In this paper, the performance of these processes is analysed when the target distribution is an anisotropic probability distribution. Continuous time processes naturally fit this problem because there is no rejection scheme unlike the traditional Markov chain Monte Carlo method, which suffers from a scaling issue. However, this does not mean that the processes are always effective for this problem. Even when the process mixes well, it may require a large number of jumps, resulting in a high computational cost. To understand this in a quantitative sense it is important to determine the influence of anisotropies in the target distribution.
In this paper, we use a multi-scale analysis framework for this analysis. We conclude that when the Gaussian target distribution has two scales, of order 1 and ϵ, then the computational cost of the Bouncy particle sampler is $\mathcal{O}(\epsilon^{-1})$, and the computational cost of the Zig-Zag sampler is $\mathcal{O}(\epsilon^{-1})$ or $\mathcal{O}(\epsilon^{-2})$ depending on the target distribution. Since the cost of the traditional Markov chain Monte Carlo method, the random walk Metropolis algorithm, is of the order of $\mathcal{O}(\epsilon^{-2})$, there is an advantage to using piecewise deterministic Markov processes.

Speaker: Augustin Chevallier

Adaptive Metropolized PDMP sampling using the No-U-Turn criterion

Augustin Chevallier
Department of Mathematics and Statistics, Lancaster University, UK, a.chevallier@lancaster.ac.uk

PDMP samplers are promising new PDMP samplers. However, they are at this moment mostly theoretical, due to the difficulty of simulating these processes. Here, a Metropolised PDMP version of a numerical approximation of the PDMP will be presented, along with the adaptation to PDMP samplers of the acclaimed No-U-Turn sampler. In this case, both the numerical step size and the path length can be chosen adaptively, which might lead to a more robust algorithm than NUTS.

Speaker: Sebastiano Grazzi

PDMP samplers with boundary conditions

Joris Bierkens
Delft Institute of Applied Mathematics, The Netherlands, Joris.Bierkens@tudelft.nl

Sebastiano Grazzi
Department of Statistics, University of Warwick, UK, Sebastiano.Grazzi@warwick.ac.uk

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

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

In this talk, I will formally introduce piecewise deterministic Markov processes (PDMPs) endowed with "sticky floors", "soft/hard walls" and "teleportation portals" which can be used for Monte Carlo simulation and allow to target efficiently a rich class of measures arising in Bayesian inference.
I will motivate and illustrate the framework with three challenging applications: Bayesian variable selection, for sampling the latent space of infection times with unknown infected population size in the SIR model with notifications and for sampling efficiently the invariant measure in hard-sphere models.
The class of processes presented here extends [1] and is joint work with J. Bierkens, G. Roberts, and M. Schauer.

[1] Bierkens, J., Grazzi, S., Meulen, F. V. D., Schauer, M. *Sticky PDMP samplers for sparse and local inference problems.* Statistics and Computing, 33(1), 1-31.

1.3 Parallel sessions - Afternoon

1.3.1 Amphitheater 25 - MS Numerical methods in statistical physics

Speaker: Pierre Monmarché

Quantitative convergence bounds for kinetic Langevin and HMC

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

We will present non-asymptotic efficiency bounds obtained for a family of MCMC samplers based on splitting schemes of the Hamiltonian or (underdamped) Langevin dynamics, under the assumption that the target measure satisfies a log-Sobolev inequality. The estimates are explicit and have a sharp dependency in the parameters of the problem (step-size, log-Sobolev constant, dimension...). The proof is based on a discrete-time adaptation of Villani's modified entropy method.

Speaker: Régis Santet

Unbiasing HMC Algorithms For General Hamiltonian Functions

Tony Lelièvre
Ecole des Ponts ParisTech et Inria Paris, France, tony.lelievre@enpc.fr

Régis Santet
Ecole des Ponts ParisTech et Inria Paris, France, regis.santet@enpc.fr

Gabriel Stoltz
Ecole des Ponts ParisTech et Inria Paris, France, gabriel.stoltz@enpc.fr

Hamiltonian Monte Carlo (HMC) [1] is a Markov Chain Monte Carlo method that allows to sample high dimensional probability measures. It relies on the integration of the Hamiltonian dynamics to propose a move which is then accepted or rejected thanks to a Metropolis procedure. Unbiased sampling is guaranteed by the preservation of key properties of the Hamiltonian dynamics (symplecticity and time-reversibility in particular) by the numerical integrators. For separable Hamiltonian functions, explicit numerical schemes, such as the Störmer–Verlet integrator, satisfy these properties [2]. For numerical reasons [3], one may consider a Hamiltonian function which is nonseparable. To integrate numerically a nonseparable Hamiltonian function while preserving the two key properties, an implicit integrator has to be used. Unfortunately, actual implicit numerical schemes cannot be reversible, as already noted in the context of constrained stochastic differential equations [4, 5]. We show here how to enforce the numerical reversibility in the HMC setting to guaranteed that the sampling is unbiased for general Hamiltonian which are even functions of the momenta. Our numerical results demonstrate that this correction is indeed relevant in practice.

[1] Duane, S. and Kennedy, A. D. and Pendleton, B. J. and Roweth, D. *Hybrid Monte Carlo.* Physics Letters B, 195(2):216–222, 1987.

[2] E. Hairer, C. Lubich, and G. Wanner. *Geometric Numerical Integration.* Colume 31 of Springer Series in Computational Mathematics. Springer-Verlag, Berlin, second edition, 2006.

[3] Girolami, M. and Calderhead, B. *Riemann Manifold Langevin and Hamiltonian Monte Carlo methods.* Journal of the Royal Statistical Society: Series B (Statistical Method- ology), 73(2):123–214, 2011.

[4] Zappa, E. and Holmes-Cerfon, M. and Goodman, J. *Monte Carlo on manifolds: sampling densities and integrating functions.* Communications on Pure and Applied Mathematics, 71(12):2609–2647, 2018.

[5] Lelièvre, T. and Rousset, M. and Stoltz, G. *Hybrid Monte Carlo methods for sampling probability measures on submanifolds.* Numerische Mathematik, 143(2):379–421, October 2019.

Speaker: Pierre Illien

Brownian dynamics simulations of colloids propelled by mesoscale phase separations

Jeanne Decayeux
PHENIX, CNRS/Sorbonne Université, Paris, France, jeanne.decayeux@sorbonne-universite.fr

Vincent Dahirel
PHENIX, CNRS/Sorbonne Université, Paris, France, vincent.dahirel@sorbonne-universite.fr

Marie Jardat
PHENIX, CNRS/Sorbonne Université, Paris, France, marie.jardat@sorbonne-universite.fr

Pierre Illien
PHENIX, CNRS/Sorbonne Université, Paris, France, pierre.illien@sorbonne-universite.fr

The motion of active colloids is generally achieved through their anisotropy, as exemplified by Janus colloids. Recently, there was a growing interest in the propulsion of isotropic colloids, which requires some local symmetry breaking. Although several mechanisms for such propulsion were proposed, little is known about the role played by the interactions within the environment of the colloid, which can have a dramatic effect on its propulsion. Here, we propose a minimal model of an isotropic colloid in a bath of solute particles that interact with each other and simulated using Brownian dynamics. These interactions lead to a spontaneous phase transition close to the colloid, to directed motion of the colloid over very long timescales and to significantly enhanced diffusion, in spite of the crowding induced by solute particles. We determine the range of parameters where this effect is observable in the model, and we propose an effective Langevin equation that accounts for it and allows one to determine the different contributions at stake in self-propulsion and enhanced diffusion.

[1] J. Decayeux, V. Dahirel, M. Jardat, P. Illien. *Spontaneous propulsion of an isotropic colloid in a phase-separating environment.* Phys. Rev. E, **104**, 034602 (2021).

[2] J. Decayeux, M. Jardat, P. Illien, V. Dahirel. *Conditions for the propulsion of a colloid surrounded by a mesoscale phase separation.* Eur. Phys. J. E, **45**, 96 (2022).

Speaker: Pierfrancesco Urbani

Dynamical mean-field theory for stochastic gradient descent in high dimensions

Pierfrancesco Urbani
Université Paris-Saclay, CNRS, CEA, Institut de physique théorique, 91191, Gif-sur-Yvette, France, pierfrancesco.urbani@ipht.fr

Artificial neural networks (ANNs) trained with the stochastic gradient-descent (SGD) algorithm have achieved impressive performances in a variety of applications. However, the theory behind this practical success remains largely unexplained. A general consensus has arisen that the answer requires a detailed description of the trajectory traversed during training. This task is highly nontrivial for at least two reasons. First, the high dimension of the parameter space where ANNs typically operate defies standard mathematical techniques. Second, SGD navigates a non-convex loss landscape following out-of-equilibrium dynamics with a complicated state-dependent noise. In this talk, I will consider prototypical learning problems that are amenable to an exact characterisation. I will show how dynamical mean-field theory from statistical physics can be used to derive an effective low-dimensional description of the network performance and the learning dynamics of multi-pass SGD [1]. Finally, I will discuss how different sources of algorithmic noise affect the performance of the network [2, 3].

[1] Mignacco, F., Krzakala, F., Urbani, P., and ZdeborovÃ¡, L. (2020) *Dynamical mean-field theory for stochastic gradient descent in Gaussian mixture classification.* Advances in Neural Information Processing Systems, 33, 9540-9550.

[2] Mignacco, F., Urbani, P., and ZdeborovÃ¡, L. (2021). *Stochasticity helps to navigate rough landscapes: comparing gradient-descent-based algorithms in the phase retrieval problem.* Machine Learning: Science and Technology, 2(3), 035029.

[3] Mignacco, F., and Urbani, P. (2022). *The effective noise of stochastic gradient descent.* Journal of Statistical Mechanics: Theory and Experiment, 2022(8), 083405.

1.3.2 Room 101, corridor 24-25 - MS Variance reduction techniques for rare events

Speaker: Nadhir Ben Rached

Importance sampling via stochastic optimal control for McKean-Vlasov stochastic differential equation

Nadhir Ben Rached
University of Leeds, UK, n.benrached@leeds.ac.uk

Abdul-Lateef Haji-Ali
Heriot-Watt University, UK, A.HajiAli@hw.ac.uk

Shyam Mohan Subbiah Pillai
RWTH Aachen University, Germany, subbiah@uq.rwth-aachen.de

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

This work [1] investigates Monte Carlo methods to estimate probabilities of rare events associated with solutions to the $d-$dimensional McKean-Vlasov stochastic differential equations (SDEs). The equation is usually approximated using a stochastic interacting $P-$particle system, a set of $P-$coupled $d-$dimensional SDEs. We use importance sampling (IS) to reduce high relative variance of Monte Carlo estimators of rare event probabilities. In the SDE context, optimal measure change is derived using stochastic optimal control to minimize estimator variance, which when applied to stochastic particle systems yields a $P \times d-$ dimensional partial differential control equation, which is cumbersome to solve. The work in [2] circumvented this problem by a decoupling approach, producing a $d-$dimensional control partial differential equation. Based on the decoupling approach, we develop a computationally efficient double loop Monte Carlo (DLMC) estimator. We systematically provide work-error analysis and formulate optimal complexity of the DLMC estimator. Our adaptive DLMC algorithm estimates rare event probabilities with significantly reduced runtimes required to achieve relative tolerance TOL_r compared with standard Monte Carlo estimators without IS. Our numerical experiments on the Kuramoto model from statistical physics show $O(TOL_r^{-4})$ complexity for our DLMC estimator with significantly reduced constant, leading to substantial computational gains.

[1] Nadhir Ben Rached, Abdul-Lateef Haji-Ali, Shyam Mohan Subbiah Pillai, Raúl Tempone, *Single Level Importance Sampling for McKean-Vlasov Stochastic Differential Equation.* arXiv preprint arXiv:2207.06926, 2022.

[2] Goncalo dos Reis, Greig Smith, Peter Tankov. *Importance sampling for McKean-Vlasov SDEs.* arXiv preprint arXiv:1803.09320, 2018.

Speaker: Shyam Mohan Subbiah Pillai

Multilevel and Multi-index Monte Carlo methods for rare events associated with McKean-Vlasov equation

Nadhir Ben Rached
University of Leeds, UK, n.benrached@leeds.ac.uk

Abdul-Lateef Haji-Ali
Heriot-Watt University, UK, A.HajiAli@hw.ac.uk

Shyam Mohan Subbiah Pillai
RWTH Aachen University, Germany, subbiah@uq.rwth-aachen.de

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

This work combines multilevel and multi-index Monte Carlo methods with importance sampling (IS) to estimate rare event quantities that can be expressed as the expectation of sufficiently regular observables of the solution to the McKean-Vlasov stochastic differential equation. We first extend the double loop Monte Carlo (DLMC) estimator, introduced in this context in our previous work [1], to the multilevel / multi-index setting. We formulate novel multilevel DLMC (MLDLMC) [2] and multi-index DLMC (MIDLMC) estimators, and perform a comprehensive work-error analysis yielding new and improved complexity results. Crucially, we also devise an antithetic sampler to estimate level differences that guarantees reduced work complexity for both MLDLMC and MIDLMC estimators compared to single level DLMC. To tackle rare events, we apply the same single level IS scheme, obtained via stochastic optimal control in [1] over all levels/indices of the MLDLMC/MIDLMC estimators respectively. Combining IS and efficient

hierarchical sampling methods not only reduces computational complexity by multiple orders, but also drastically reduces the associated constant, when compared to the naive single level DLMC estimator. We illustrate effectiveness of proposed estimators on the Kuramoto model from statistical physics with sufficiently regular observables, confirming reduced complexity from $O(TOL_r^{-4})$ for the naive DLMC estimator to $O(TOL_r^{-2})$ (up to logarithmic terms) using MIDLMC, while providing feasible estimates of rare event quantities up to prescribed relative error tolerance TOL_r.

[1] Nadhir Ben Rached, Abdul-Lateef Haji-Ali, Shyam Mohan Subbiah Pillai, Raúl Tempone, *Single Level Importance Sampling for McKean-Vlasov Stochastic Differential Equation.* arXiv preprint arXiv:2207.06926, 2022.

[2] Nadhir Ben Rached, Abdul-Lateef Haji-Ali, Shyam Mohan Subbiah Pillai, Ra'ul Tempone, *Multilevel Importance Sampling for McKean-Vlasov Stochastic Differential Equation.* arXiv preprint arXiv:2208.03225, 2022.

Speaker: Bruno Tuffin

Bounds, Assessment and Confidence Intervals for Exponential Approximations

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

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

Bruno Tuffin
Inria, France, bruno.tuffin@inria.fr

Geometric sums can often be approximated by an exponential random variable when the number of summands is large [3], simplifying estimations in many rare contexts. It is typically the case when estimating the distribution of the waiting time in an M/G/1 queue or the hitting time to a rare set for a regenerative process [1, 2]. One of the main advantages of the approximation is to reduce the estimation of the whole distribution to the estimation of its mean, for which efficient (rare event) simulation are usually available in the literature.
Though, the approximation error introduces a bias to the estimation which needs to be assessed and ensured to be negligible with respect to the statistical error, or included in the overall estimation error. This presentation contributes to solve this issue: it provides deterministic approximation bounds when parameters are known, which can be used to control the validity of the estimation and to propose thresholds on the mean of the geometric number of summands for which the approximation is advised to be used.

[1] P.W. Glynn, M.K. Nakayama, B. Tuffin (2018) Using simulation to calibrate exponential approximations to tail-distribution measures of hitting times to rarely visited sets. *Proceedings of the 2018 Winter Simulation Conference* (Piscataway, NJ: Institute of Electrical and Electronics Engineers).

[2] P.W. Glynn, M.K. Nakayama, B. Tuffin (2020) Comparing regenerative-simulation-based estimators of the distribution of the hitting time to a rarely visited set. Bae KH, Feng B, Kim S, Lazarova-Molnar S, Zheng Z, Roeder T, Thiesing R, eds., *Proceedings of the 2020 Winter Simulation Conference* (Piscataway, New Jersey: IEEE).

[3] V. Kalashnikov (1997) *Geometric Sums: Bounds for Rare Events with Applications* (Dordrecht, The Netherlands: Kluwer Academic Publishers).

Speaker: Charly Andral

The Importance Markov Chain

Charly Andral
CEREMADE, Université Paris-Dauphine, andral@ceremade.dauphine.fr

Randal Douc
SAMOVAR, Télécom SudParis, randal.douc@telecom-sudparis.eu

Hugo Marival
SAMOVAR, Télécom SudParis, hugo.marival@telecom-sudparis.eu

Christian P. Robert
CEREMADE, Université Paris-Dauphine & University of Warwick, xian@ceremade.dauphine.fr

The Importance Markov chain is a new algorithm bridging the gap between rejection sampling and importance sampling, moving from one to the other through a tuning parameter. Based on a modified sample of an instrumental Markov chain targeting an instrumental distribution (typically with a MCMC kernel), the Importance Markov chain aims to construct an extended Markov chain where the marginal distribution of the first component converges to the target distribution. We obtain geometric ergodicity for this extended kernel, under mild assumptions on the instrumental kernel. For example, when targeting a multimodal distribution, the instrumental distribution can be chosen as a tempered version of the target which allows the algorithm to explore its modes more efficiently. A law of large numbers and a central limit theorem are also obtained. Computationally, the algorithm is easy to implement and preexisting librairies can be used to sample from the instrumental distribution. In the particular case where the instrumental chain is an i.i.d. sample, the independant Importance Markov chain outperforms in our simulations the independant Metropolis-Hastings algorithm.

[1] Charly Andral, Randal Douc, Hugo Marival, and Christian P Robert. The Importance Markov Chain. *arXiv preprint arXiv:2207.08271*, 2022.

1.3.3 Room 102, corridor 15-25 - MS Slice sampling and adaptive MCMC

Speaker: Mareike Hasenpflug

Slice Sampling on the Sphere

Mareike Hasenpflug
University of Passau, Germany, mareike.hasenpflug@uni-passau.de

We introduce a geodesic slice sampler on the Euclidean sphere (in arbitrary but fixed dimension) that can be used for approximate sampling from distributions that have a density with respect to the corresponding surface measure. Such distributions occur e.g. in the modelling of directional data or shapes. Under some mild conditions we show that the corresponding transition kernel is well-defined, in particular, that it is reversible with respect to the distribution of interest. Moreover, if the density is bounded, then we obtain a uniform ergodicity convergence result. Finally, we illustrate the performance of the geodesic slice sampler on the sphere with numerical experiments.

Speaker: Philip Schaer

Making Polar Slice Sampling Efficiently Implementable

Michael Habeck
Friedrich Schiller University Jena, michael.habeck@uni-jena.de

Daniel Rudolf
University of Passau, daniel.rudolf@uni-passau.de

Philip Schaer
Friedrich Schiller University Jena, philip.schaer@uni-jena.de

Despite promising convergence guarantees and a simple implementation proposal that is applicable under mild conditions, polar slice sampling [1] has not found its way into broad practical application in the twenty years since its inception. In our view, the reason for this lies in the fact that the implementation proposed in [1] is based on an acceptance-rejection scheme that quickly becomes extremely computationally costly when the target density starts to deviate from rotational invariance. With the resulting astronomical costs per iteration, the remarkable convergence speed in terms of the number of iterations becomes irrelevant. To remedy this unfortunate state of affairs, we develop a sampler that closely approximates polar slice sampling while being about as efficient as other practically applied slice sampling methods. Numerical experiments show our method to outperform the related approaches elliptical slice sampling [2] and hit-and-run uniform slice sampling [3, Section 29.7] in various settings. A particularly noteworthy aspect of our method is its suitability for target distributions that exhibit funneling, i.e. those that feature both very narrow and very wide regions of high probability mass. An experiment with the most well-known such distribution, Neal's funnel [4], which is generally considered very challenging, shows our method to substantially outperform the two related approaches in terms of the speed at which various target quantities are computed.

[1] G. O. Roberts, J. S. Rosenthal *The Polar Slice Sampler* Stochastic Models, 18(2), 257–280, 2002.

[2] I. Murray, R. Adams, D. MacKay *Elliptical Slice Sampling* Journal of Machine Learning Research, 9, 541–548, 2010.

[3] David MacKay *Information Theory, Inference and Learning Algorithms* Cambridge University Press, 2003.

[4] Radford M. Neal *Slice Sampling* The Annals of Statistics, 31(3), 706–767, 2003.

Speaker: Julian Hofstadler

Adaptive MCMC for doubly intractable distributions

Julian Hofstadler
University of Passau, Germany, julian.hofstadler@uni-passau.de

We consider posterior densities with two unknown quantities, the normalizing constant and an intractable multiplicative factor in the likelihood function. Densities of this form appear, e.g., if one uses Bayesian inference in the context of biophysical problems. Not being able to evaluate the likelihood function leads to computational issues in classical (adaptive) MCMC algorithms, however, in the past years different methods have been suggested to overcome this problem. We discuss an adaptive MCMC scheme that relies on approximating the likelihood function and, moreover, we present convergence results for bounded measurable functions.

Speaker: Andi Wang

Comparison theorems for Hybrid Slice Sampling

Daniel Rudolf
Passau University, Germany, daniel.rudolf@uni-passau.de

Björn Sprungk
TU Freiberg, Germany, bjoern.sprungk@math.tu-freiberg.de

Sam Power
University of Bristol, UK, sam.power@bristol.ac.uk

Andi Wang
University of Warwick, UK, Andi.Wang@warwick.ac.uk

Hybrid Slice Sampling is a variant of Simple Slice Sampling, whereby exact sampling from the uniform distribution on a level set is replaced with sampling from a Markov kernel which is merely invariant with respect to this distribution. It is known that the spectral gap of the Hybrid Slice Sampling algorithm, if it exists, is smaller than that of the Simple Slice Sampling algorithm [2]. Following the framework first presented in [1], we shall derive new comparison results, giving convergence guarantees for Hybrid Slice Sampling, relative to convergence rates of Simple Slice Sampling. These comparisons are valid even in subgeometric scenarios when the spectral gap is 0, and significantly extend and simplify previous results in the literature [2].

[1] Andrieu, C., Lee, A., Power, S., Wang, A.Q. *Comparison of Markov chains via weak Poincaré inequalities with application to pseudo-marginal MCMC.* The Annals of Statistics, 50(6): 3592-3618, 2022.

[2] K. Latuszynski and D. Rudolf. *Convergence of hybrid slice sampling via spectral gap.* arXiv preprint.

1.3.4 Room 104, corridor 15-25 - Contributed Talks

Speaker: Noufel Frikha

On the convergence of the Euler-Maruyama scheme for McKean-Vlasov SDEs

Noufel Frikha
Université Paris 1 Panthéon Sorbonne, Centre d'Économie de la Sorbonne, 106 Boulevard de l'Hôpital, 75642 Paris Cedex 13, France, noufel.frikha@univ-paris1.fr

Xuanye Song
Université Paris Cité, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, F-75013 Paris, France, xsong@lpsm.paris

Clément Rey
Ecole Polytechnique, CMAP, France, clement.rey@polytechnique.edu

Relying on the backward Kolmogorov partial differential equation investigated in [1, 2] stated on the strip $[0,T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$, $\mathcal{P}_2(\mathbb{R}^d)$ being the Wasserstein space, that is, the space of probability measures on $\mathbb{R}^d$ with a finite second-order moment, we obtain several new results concerning the approximation error of some non-linear stochastic differential equations in the sense of McKean-Vlasov by the corresponding Euler-Maruyama time discretization scheme of its system of interacting particles. We notably present explicit error estimates, at the level of the trajectories, at the level of the semigroup and at the level of the densities. Some Gaussian density estimates for the Euler-Maruyama scheme are also established by relying on the fundamental solution for the related parabolic linear operator here stated on $[0,T] \times \mathcal{P}_2(\mathbb{R}^d)$.

[1] Paul-Eric Chaudru de Raynal, Noufel Frikha. *Well-posedness for some non-linear SDEs and related PDE on the Wasserstein space.* Journal de Mathématiques Pures et Appliquées, volume 159, March 2022, pages 1-167. https://doi.org/10.1016/j.matpur.2021.12.001 .

[2] Paul-Eric Chaudru de Raynal, Noufel Frikha. *From the backward Kolmogorov PDE on the Wasserstein space to propagation of chaos for McKean-Vlasov SDEs.* Journal de Mathématiques Pures et Appliquées, volume 156, December 2021, pages 1-124 https://doi.org/10.1016/j.matpur.2021.10.010

Speaker: Gonçalo dos Reis

High order splitting methods for stochastic differential equations

Gonçalo dos Reis
University of Edinburgh, UK, G.dosReis@ed.ac.uk

Calum Strange
University of Edinburgh, UK, c.strange-1@sms.ed.ac.uk

James Foster
University of Bath, UK, jmf68@bath.ac.uk

In this talk, we will discuss how ideas from rough path theory can be leveraged to develop high order numerical methods for SDEs. To motivate our approach, we consider what happens when the Brownian motion driving an SDE is replaced by a piecewise linear path. We show that this procedure transforms the SDE into a sequence of ODEs – which can then be discretized using an appropriate ODE solver. Moreover, to achieve a high accuracy, we construct these piecewise linear paths to match certain "iterated" integrals of the Brownian motion. At the same time, the ODE sequences obtained from this path-based approach can be interpreted as a splitting method, which neatly connects our work to the existing literature. For example, we show that the well-known Strang splitting falls under this framework and can be modified to give an improved convergence rate. We will conclude the talk with a couple of examples, demonstrating the flexibility and convergence properties of our methodology.
(joint work with James Foster and Calum Strange, https://arxiv.org/abs/2210.17543)

Speaker: Wei Cai

An Iterative Probabilistic Method for Mixed Problems of Laplace Equations with the Feynman–Kac Formula of Killed Brownian Motions

Wei Cai
Department of Mathematics, Southern Methodist University, Dallas, TX75275, USA, cai@smu.edu

In this talk, a probabilistic method [1] using the Feynman–Kac formula of killed Brownian motions will be presented to solve the mixed boundary value problems (BVPs) of 3D Laplace equations. To avoid using reflecting Brownian motions and the calculation of their local time L(t) in the Feynman–Kac representation of solutions for Neumann and Robin BVPs, the proposed method uses an iterative approach to approximate the solutions where each iteration will use the Feynman–Kac formula to solve a pure Dirichlet problem, thus only involving killed Brownian motions. First, the boundary of the domain is decomposed with overlapping local patches formed by the intersection of hemispheres superimposed on the domain boundary. The iteration starts with an arbitrary initial guess for the Dirichlet data on the Neumann and Robin boundaries; then, using the Feynman-Kac formula for a pure Dirichlet problem

with the current available Dirichlet data on the whole boundary, the solution over the hemispheres can be obtained by the Feynman—Kac formula for the killed Brownian motions, sampled by a walk-on-spheres (WOS) algorithm. Second, by solving a local boundary integral equation (BIE) over each hemisphere and a local patch on the domain boundary, the Dirichlet data on the Neumann and Robin boundaries can be updated. By continuing this process, the proposed hybrid probabilistic and deterministic BIE-WOS method gives a highly parallel algorithm for the global solution of any mixed-type BVPs of the Laplace equations. Numerical results of various mixed interior and exterior BVPs demonstrate the parallel efficiency and accuracy of the proposed method.

[1] Cuiyang Ding, Changhao Yan, Xuan Zeng, And Wei Cai *A Parallel Iterative Probabilistic Method for Mixed Problems of Laplace Equations with the Feynman–Kac Formula of Killed Brownian Motions.* SIAM J. SCI. COMPUT, Vol. 44, No. 5, pp. A3413–A3435, 2022.

Speaker: Stefano Pagliarani

Numerical approximation of McKean-Vlasov SDEs via Stochastic Gradient Descent

Ankush Agarwal
University of Glasgow, UK, Ankush.Agarwal@glasgow.ac.uk

Stefano Pagliarani
University of Bologna, Italy, stefano.pagliarani9@unibo.it

Gonçalo dos Reis
University of Edinburgh, UK, G.dosReis@ed.ac.uk

We explore a novel method for the numerical resolution of McKean-Vlasov (MKV) SDEs that does not rely on the simulation of interacting particle systems. Specifically, we investigate the use of Stochastic Gradient Descent (SGD), which has recently emerged as a powerful tool for high-dimensional optimization problems in machine learning. We consider a class of MKV SDEs with separable coefficients, whose solutions can be identified with a family of functions of time that solve a functional fixed-point equation. We propose an algorithm that combines a projection method with SGD, to efficiently and accurately approximate these functions in terms of linear combinations of elementary functions. We provide a convergence result under simplified (Lipschitz) assumptions, and present numerical experiments to demonstrate the effectiveness and efficiency of the approach, for different regimes of learning rate and batch size, and under different regularity and growth assumptions on the coefficients.

2 Tuesday 27th June

2.1 Keynote speakers

Speaker: Sophie Donnet

Using a Sequential Monte Carlo algorithm to find the mesoscale structure of an ecological network

Sophie Donnet
INRAE, MIA Paris Saclay, France, sophie.donnet@inrae.fr

This work is motivated by the analysis of ecological interaction networks. Stochastic block models are widely used in this field to decipher the structure that underlies a network or that is shared by a collection of networks. Efficient algorithms based on variational approximations exist for frequentist inference and sometimes for Bayesian inference, but without statistical guaranties as for the resulting estimates. We propose to combine the variational estimation with a sequential Monte-Carlo algorithm to efficiently sample the posterior distribution and to perform model selection. We use this methodology to find common meso-structures in a collection of bipartite ecological networks.

Speaker: Wenzel Jakob

Differential Monte Carlo Methods for Inverse Problems (not only) in Computer Graphics

Wenzel Jakob
Realistic Graphics Lab at EPFL's School of Computer and Communication Sciences, Switzerland, wenzel.jakob@epfl.ch

Rendering algorithms rely on Monte Carlo integration to convert a three-dimensional model of a virtual world into a photorealistic image. We are surrounded by their output in essentially any visual medium. This talk examines the associated *inverse problem*, which infers a 3D world observed through a given set of 2D images (e.g., photos from different viewpoints). One "obvious" way to approach this problem combines the derivative of the Monte Carlo estimator with some form of stochastic gradient descent to evolve a tentative world until it becomes consistent with the images. However, this naïve approach runs into various obstacles and dead ends. I will present my group's work on solving such problems at scale, focusing on the following aspects:

- Automatic differentiation is highly inefficient when applied to this problem. I will show how the derivative calculation can be cast into an efficient physical analog that can be solved using linear time and constant space complexity.
- The underlying integrals are riddled with discontinuities, causing severe gradient bias that breaks the optimization. I will review methods to address this problem.
- Monte Carlo sampling strategies designed for the original computation are no longer appropriate for its derivative; I will review tailored derivative sampling strategies.
- The combination of Monte Carlo integration and gradient descent involves a high degree of redundancy that can be mitigated using control variates.

2.2 Parallel sessions - Morning

2.2.1 Amphitheater 25 - MS Numerical methods in statistical physics

Speaker: Gilles Vilmart

Accelerated convergence to equilibrium and reduced asymptotic variance for Langevin dynamics using Stratonovich perturbations

Assyr Abdulle
EPF Lausanne, Switzerland,

Grigorios A. Pavliotis
Imperial College London, UK, g.pavliotis@imperial.ac.uk

Gilles Vilmart
University of Geneva, Switzerland, gilles.vilmart@unige.ch

We propose a new approach for sampling from probability measures in, possibly, high dimensional spaces. By perturbing Langevin dynamics by a suitable Stratonovich perturbation that preserves the invariant measure of the original system, we show that accelerated convergence to equilibrium and reduced asymptotic variance can be achieved, leading, thus, to a computationally advantageous sampling algorithm. The new perturbed Langevin dynamics is reversible with respect to the target probability measure and, consequently, does not suffer from the drawbacks of nonreversible Langevin samplers, such as undesired oscillatory transient behaviours.

[1] A. Abdulle, G. A. Pavliotis, and G. Vilmart, Accelerated convergence to equilibrium and reduced asymptotic variance for Langevin dynamics using Stratonovich perturbations, arXiv:1903.03024, *C. R. Acad. Sci. Paris; Ser. I* 357 (2019), 349–354.

[2] A. Abdulle, G. A. Pavliotis, and G. Vilmart, Accelerated Langevin dynamics using Stratonovich perturbations, in preparation.

Speaker: Petr Plechac

Estimating linear response and sensitivity analysis of non-equilibrium steady states

Petr Plechac
University of Delaware, USA, plechac@udel.edu

Gabriel Stoltz
École des Ponts ParisTech et Inria Paris, France, gabriel.stoltz@enpc.fr

Ting Wang
DEVCOM Army Research Laboratory, Aberdeen Proving Ground, USA, tingw@udel.edu

In this talk we present numerical schemes for linear response computations of invariant measures from fluctuations at steady states. The schemes are based on Girsanov's change-of-measure theory and apply reweighting of trajectories by factors derived from a linearization of the Girsanov weights. This approach leads to a class of estimators we call martingale product estimators. We show that the derived estimators are of bounded variance with respect to the integration time, which is a desirable feature for long time simulations. We discuss how the discretization error can be improved to second order accuracy in the time step by modifying the weight process in an appropriate way.We provide numerical evidence demonstrating the efficiency and accuracy of the method.

Speaker: Dominic Phillips

Coordinate Transforms for Efficient Brownian Dynamics Simulations

Dominic Phillips
University of Edinburgh, UK, Dominic.Phillips@ed.ac.uk

Brownian dynamics with multiplicative noise is a powerful tool for modelling stochastic processes in physical and biological sciences. However, numerical integrators for these processes can lose accuracy or even fail to converge when the diffusion term is configuration-dependent. One remedy is to construct a transform to a constant-diffusion process, and sample the transformed process instead. In this work, we explain how coordinate-based and time-rescaling-based transforms can be used either individually or in combination to map a general class of variable-diffusion Brownian motion processes into constant-diffusion ones. The transforms are invertible, thus allowing recovery of the original dynamics. We motivate our methodology using examples in one dimension before then considering multivariate diffusion processes. We illustrate the benefits of the transforms through numerical simulations, demonstrating how the right combination of integrator and transform can improve computational efficiency and the order of convergence to the invariant distribution. Notably, the transforms that we derive are applicable to a class of multibody, anisotropic Stokes-Einstein diffusion that has applications in biophysical modelling.

Speaker: Shiva Darshan

Sticky Coupling as a Control Variate for Sensitivity Analysis

Nawaf Bou-Rabee
Rutgers University Camden, USA, nawaf.bourabee@rutgers.edu

Shiva Darshan
École des Ponts ParisTech et Inria Paris, France, shiva.darshan@enpc.fr

Andreas Eberle
Universität Bonn, Germany, eberle@uni-bonn.de

Gabriel Stoltz
École des Ponts ParisTech et Inria Paris, France, gabriel.stoltz@enpc.fr

A standard method to compute transport coefficients is to simulate Langevin dynamics perturbed by a small non-equilibrium forcing up to a time T and time-average over the trajectory a desired observable divided by the magnitude of the forcing, η. Unfortunately, this method suffers from large finite-time sampling bias and variance in the limit of small forcing—on the order of $(T\eta)^{-1}$ and $(T\eta^2)^{-1}$ respectively. For overdamped Langevin dynamics, we propose a method to reduce the bias and variance of this computation using a version of the reference (unperturbed) dynamics sticky coupled [1] to the perturbed dynamics as a control variate. We will show that when the potential of the dynamics is strongly convex at infinity, this sticky coupling based estimator's reduces the bias and variance by a factor of η^{-1} compared to the standard method. The case of strongly convex at infinity potentials includes commonly used systems such as Lennard-Jones particles confined to box by a quadratic potential.

[1] A. Eberle and R. Zimmer. *Sticky Couplings of Multidimensional Diffusions with Different Drifts.* Ann. Inst. H. Poincaré Probab. Statist., 55:4, 2370–2394, November 2019.

2.2.2 Room 101, corridor 24-25 - MS SDEs - Theory, Numerics and Applications

Speaker: Mireille Bossy

Stochastic approach for the simulation of non-spherical particles in turbulence

Mireille Bossy
INRIA, mireille.bossy@inria.fr

Investigating the dynamics of non-spherical particles suspended in turbulent flows is paramount to several industrial, biological and environmental applications (to name a few examples, fibres in papermaking industries, plankton in the ocean or even bacteria, complex-shaped ice-crystals in clouds, soot and ashes in the atmosphere).
In addition to its translational dynamics, a non-spherical particles displays rotational dynamics. And both depend on the particle properties (especially its shape and inertia). Besides, in many of these applications, the flow is highly turbulent and it has a profound effect on the rotational dynamics, alignment trends and correlations of anisotropic particles. Dealing with such level of description in engineering applications implies to introduce a dedicated model of the flow velocity gradient seen by the particles.
In this talk, we will introduce a first stochastic model for the orientation dynamics of the spheroids (parametrized by their aspect ratio) and briefly review several aspects of the proposed SDE, in particular its physical validation, the time transition between Levy noise and Brownian noise regimes due to the presence of high vorticity structures in turbulence and the related numerical simulation issues.
This is based on joint works with J. Bec [1], L. Campana [1, 2], C. Henry [2] and work in progress Paul Maurer.

[1] L. Campana, M. Bossy, and J. Bec. Stochastic model for the alignment and tumbling of rigid fibers in two-dimensional turbulent shear flow. *Phys. Rev. Fluids*, 7:124605, Dec 2022. doi:10.1103/PhysRevFluids.7.124605.

[2] L. Campana, M. Bossy, C. Henry, Lagrangian stochastic model for the orientation of inertialess non spherical particles in turbulent flows: an efficient numerical method for CFD approach, https://arxiv.org/abs/2211.10211, November 2022.

Speaker: Evelyn Buckwar

Construction and analysis of splitting methods for Chemical Langevin Equations

Evelyn Buckwar
Johannes Kepler University Linz, evelyn.buckwar@jku.at

Consider modeling the stochastic dynamics underlying different chemical systems, which is usually described by the Gillespie Stochastic Simulation Algorithm (SSA), i.e. the Markov process arising from taking into account every single chemical reaction event. While exact and easy to implement, this algorithm is computationally expensive for chemical reactions involving a large number of molecular species. As an approximation, Chemical Langevin Equations (CLEs) can work for large number of species or/and reactions. In this talk, we construct an explicit splitting method applied to the system of CLEs for a simple example of a reversible bimolecular reaction. The drift term of this stochastic differential equation system satisfies a local one-sided Lipschitz condition and the diffusion term involves square root terms. We then present the main ideas of a mean-square convergence proof, as well as numerical illustrations. The results are joint work with Youssra Souli, Johannes Kepler University, Linz.

Speaker: Khadija Meddouni

Numerical methods for stochastic neural field equations

Gabriel Lord
Radboud University, Nijmegen, NL, gabriel.lord@ru.nl

Khadija Meddouni
Radboud University, Nijmegen, NL, khadija.meddouni@ru.nl

Daniele Avitabile
Vrije Universiteit, Amsterdam, NL, d.avitabile@vu.nl

We consider strong convergence of space-time discretizations of neural field equations with additive noise. Our methods are variants of exponential based methods in time and we examine projection and collocation methods in space. We introduce the neural field equations, some properties of solutions before we outline the proof of strong convergence. Numerically we compare the efficiency of methods. If time permits we will illustrate preliminary results coupling the electrical activity to a porous media.

Speaker: Cónall Kelly

Adaptive Meshes for Stochastic Jump Differential Equations

Cónall Kelly
University College Cork, Ireland, conall.kelly@ucc.ie

Gabriel Lord
Radboud University, Nijmegen, NL, gabriel.lord@ru.nl

Fandi Sun
Heriot-Watt Universsity, Edinburgh, UK, fandi-sun@outlook.com

We consider the use of adaptive timestepping for d-dimensional stochastic differential equations (SDEs) driven by both a Wiener diffusion and a Poisson jump process. In the absence of jumps, adaptive timestepping can be used to ensure strong convergence of explicit Euler-Maruyama and Milstein schemes under conditions where explicit schemes on a uniform mesh fail to converge in either the weak or strong sense. Specifically, when the drift and diffusion satisfy a local Lipschitz condition, a polynomial growth condition, and together a Khasminskii-type monotone condition.
In this talk, we examine how timestepping strategies for non-jump SDEs can be extended to ensure strong convergence of explicit schemes for SDEs with Poisson jumps. We discuss the construction of adaptive meshes, the stochastic framework for a general error analysis, and present numerical examples in the case where the Milstein method is used to discretise the non-jump part of the SDE.

2.2.3 Room 102, corridor 15-25 - MS Recent advances in MCM for forward and inverse problems for stochastic reaction networks

Speaker: Sophia Wiechert

Markovian Projection for Efficient Importance Sampling of Stochastic Reaction Networks

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

Nadhir Ben Rached
University of Leeds, UK, n.benrached@leeds.ac.uk

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

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

We propose a novel importance sampling (IS) approach to improve the efficiency of Monte Carlo (MC) estimators for various statistical quantities of stochastic reaction networks (SRNs). A SRN is a continuous-time, discrete-space Markov chain, which describes the random interaction of d species through reactions. Our special interest lays in estimating rare event probabilities in high dimensional SRNs, in which $d \gg 1$.Therefore, we formulate a parameterized path-dependent IS framework based on the approximate tau-leap scheme. In the same spirit of [1], we derive a highly efficient automated approach to obtain the optimal IS parameters based on a stochastic optimal control (SOC) formulation by solving a backward Hamilton-Jacobi-Bellman (HJB) equation. As an alternative method to the learning-based approach in [1], we mitigate the curse of dimensionality by mapping the problem to a lower dimensional space via the Markovian projection (MP) idea. The output of this model reduction technique is a much lower dimensional SRN that preserves at all time steps the marginal distribution of the original high-dimensional problem. By solving the HJB equation for a projected lower dimensional process, we get projected IS parameters, which can be mapped back to the original d dimensional SRN. The projected process is obtained by solving an optimization problem via a discrete time L^2 regression. Our numerical simulations show that those MP-IS parameters result in a significant variance reduction when used for an IS-MC estimator of the full-dimensional SRN. Especially in the rare event regime, this leads to reduced computational cost compared to the standard TL-MC estimator.

[1] Chiheb Ben Hammouda, Nadhir Ben Rached, Raúl Tempone, and Sophia Wiechert. *Learning-Based Importance Sampling via Stochastic Optimal Control for Stochastic Reaction Networks.* ArXiv preprint arXiv:2110.14335 (2021).

Speaker: Fang Zhou

A scalable approach for solving chemical master equations based on modularization and filtering

Fang Zhou
Department of Biosystems Science and Engineering, ETH-Zürich, Switzerland,
zhou.fang@bsse.ethz.ch

Ankit Gupta
Department of Biosystems Science and Engineering, ETH-Zürich, Switzerland,
ankit.gupta@bsse.ethz.ch

Mustafa Khammash
Department of Biosystems Science and Engineering, ETH-Zürich, Switzerland,
mustafa.khammash@bsse.ethz.ch

Solving the chemical master equation (CME) that characterizes the probability evolution of stochastically reacting processes is greatly important for analyzing intracellular reaction systems. Conventional methods for solving CMEs include the simulation-based Monte-Carlo methods, the direct approach (e.g., the finite state projection), and so on; however, they usually do not scale well with the system dimension either in terms of accuracy or efficiency. To mitigate this problem, this talk presents a new computational method based on modularization and filtering. Our method first divides the whole system into a leader system and several conditionally independent follower subsystems. Then, we solve the CME by applying the Monte Carlo method to the leader system and the direct approach to the filtered CMEs that characterize the conditional probabilities of the follower subsystems. The system decomposition involved in our method is optimized so that all the subproblems above are low dimensional, and, therefore, our approach scales more favorably with the system dimension. Finally, we demonstrate the efficiency and accuracy of our approach in high-dimensional estimation and inference problems using several biologically relevant examples.

Speaker: Ankit Gupta

Frequency domain methods for analysing stochastic reaction networks

Ankit Gupta
Department of Biosystems Science and Engineering, ETH-Zürich, Switzerland,
ankit.gupta@bsse.ethz.ch

Mustafa Khammash
Department of Biosystems Science and Engineering, ETH-Zürich, Switzerland ,
mustafa.khammash@bsse.ethz.ch

The dynamics of intracellular reaction networks are often stochastic and commonly modeled as continuous-time Markov chains (CTMCs). Computational analyses and inference of such stochastic models are fraught with many difficulties, arising due to an inherent curse of dimensionality which causes the underlying state-space to be exorbitantly large. While some of these issues can be mitigated by simulating CTMC trajectories and designing suitable Monte Carlo procedures, these approaches generally impose a heavy computational burden as they often require a large sample of trajectories to achieve the desired statistical accuracy. The aim of this talk is to present novel frequency domain methods that mitigate this issue by efficiently extracting information from a small number of simulated trajectories. Using several examples, we shall demonstrate the efficacy of these methods in producing statistically reliable estimates for various time-varying quantities of interest, like means, variances, autocorrelations, parametric sensitivities, etc.

Speaker: David Warne

Multifidelity multilevel approximate Bayesian computation for stochastic biochemical reaction networks

David J. Warne
School of Mathematical Sciences and Centre for Data Science, Queensland University of Technology, Brisbane, Australia, david.warne@qut.edu.au

Thomas P. Prescott
The Alan Turing Institute, UK, tpprescott@gmail.com

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

Matthew J. Simpson
School of Mathematical Sciences and Centre for Data Science, Queensland University of Technology, Brisbane, Australia,

Many biochemical processes, such as gene expression and regulation, are stochastic in nature. As a result, stochastic models of biochemical kinetics based on continuous-time Markov processes are widely utilised to study cellular functions. However, experimental data based on microscopy and fluorescent tagging represent incomplete observations of reality. This fact, along with the need to solve the chemical master equation, leads to the intractability of the likelihood function. As a result, it is common to apply likelihood-free approaches that replace likelihood evaluations with realisations of the model and observation process. However, likelihood-free techniques are computationally expensive for accurate inference as they may require millions of high-fidelity, expensive stochastic simulations. To address this challenge, we develop a new method based on recent advances in the class of methods for estimation of expectations with respect to posterior distributions for parameter inference with partially observed Markov processes models. Our novel approach combines the multilevel Monte Carlo telescoping summation, applied to a sequence of approximate Bayesian posterior targets, with a multifidelity rejection sampler that learns from low-fidelity, computationally inexpensive, model approximations to minimise the number of high-fidelity, computationally expensive, simulations required for accurate inference. Using several non-trivial examples of biochemical systems, we demonstrate improvements of more than two orders of magnitude over standard rejection sampling techniques.

2.2.4 Room 104, corridor 15-25 - Contributed Talks

Speaker: Till Massing

Simulating Continuous-Time Autoregressive Moving Average Processes Driven By Tempered Stable Lévy Processes

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

We discuss simulation schemes for continuous-time autoregressive moving average (CARMA) processes driven by tempered stable Lévy noises. CARMA processes are the continuous-time analogue to ARMA processes as well as a generalization of Ornstein-Uhlenbeck processes. However, unlike Ornstein-Uhlenbeck processes with a tempered stable driver (see, e.g., [2]) exact transition probabilities for higher order CARMA processes are intractable and not explicitly given. Therefore, we opt for the sample path generation method of [1] and approximate the driving tempered stable Lévy process by truncated series representations. We prove approximation error bounds, conduct Monte Carlo experiments to illustrate the validity, and discuss potential applications. Furthermore, we study several generalizations, for example, driving processes with p-tempered stable distributions for which we derive series representations building on [3].

[1] R. Kawai *Sample path generation of Lévy-driven continuous-time autoregressive moving average processes.* Methodology and Computing in Applied Probability, 19:1, 175–211, 2017.

[2] Y. Qu, A. Dassios and H. Zhao *Exact simulation of Ornstein–Uhlenbeck tempered stable processes.* Journal of Applied Probability, 58:2, 347–371, 2021.

[3] J. Rosiński *Tempering stable processes.* Stochastic processes and their applications, 117:6, 677–707, 2007.

Speaker: Wei Xu

Random Willow Tree with Application in Risk Management

Wei Xu
Department of Mathematics, Toronto Metropolitan University, Canada, wei.xu@torontomu.ca

Derivatives underlying a portfolio is popular on the market to diversify the market risk. However, existing method, the nested simulation, is quite time-consuming for pricing and managing the risk. In this article, we propose an efficient approach, randomized willow tree method. There are three main stages for our approach, portfolio distribution approximation, randomized willow tree construction and managing the risk of derivatives. We first generate some simulated paths to describe the evolution of dynamic portfolio values. Then, the minimal relative entropy (MRE) method is applied to approximate the distribution of portfolio values at each time based on the simulated data. After the approximated distributions are determined, a randomized willow tree can be constructed for pricing and managing the risk of derivatives underlying the portfolio. Finally, we apply the proposed approach to calculate annual dollar delta, 99% VaR and CVaR of a particular derivative, i.e., a 19-year variable annuity with guarantee riders. This application demonstrates the efficiency and accuracy of the proposed approach compared with the common nested simulation technique, especially for a large pool of derivatives underlying the same portfolio.

Speaker: Jonathan Spence

Efficient Risk Estimation for the Credit Valuation Adjustment

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

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

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

Financial derivatives in over-the-counter markets are subject to a series of valuation adjustments. Estimating key risk measures of these adjustments often requires combining several layers of approximation, including Monte Carlo estimation of recursively nested expectations and discretisation of stochastic market factors. As an example, we consider the value-at-risk of the Credit Valuation Adjustment, which depends on the solution, S_t, of an underlying SDE and involves approximating triply nested expectations of the form

$$\mathrm{E}\Big[\mathrm{H}\Big(\mathrm{E}\big[f\big(\mathrm{E}[g(S_T)\mid S_\tau]\big)\mid S_h\big]\Big)\Big], \tag{3}$$

with $0 < h < \tau < T$, the Heaviside function $\mathrm{H}(x) = 1$ if $x \geq 0$ and $\mathrm{H}(x) = 0$ if $x < 0$ and Lipschitz functions f and g. The nested structure of expectations combined with the discontinuity in the Heaviside function and numerical approximation of S_t results in an $\mathcal{O}(\varepsilon^{-5})$ cost to achieve a root mean square error ε using a standard nested Monte Carlo approach.
We consider an alternative estimator for this problem obtained by recursively constructing a hierarchy of unbiased multilevel Monte Carlo averages for the nested expectations appearing within the Heaviside function in (3), allowing for Milstein simulation of S_t and Monte Carlo approximations of each nested expectation. Adaptive multilevel Monte Carlo methods are then used to approximate the outermost expectation in (3), with sample-wise refinement based on proximity to the discontinuity in the Heaviside function. The resulting estimator is shown to have a greatly reduced $\mathcal{O}(\varepsilon^{-2}(\log\varepsilon)^2)$ cost for root mean square error ε.

[1] M. Giles, A.-L. Haji-Ali and J. Spence. *Efficient risk estimation for the Credit Valuation Adjustment.* arXiv preprint arXiv:2301.05886, 2023.

[2] A.-L. Haji-Ali, J. Spence and A. Teckentrup. *Adaptive multilevel Monte Carlo for probabilities.* SIAM Journal on Numerical Analysis, 60(4):2125–2149, 2022.

Speaker: Raaz Dwivedi

Compress Then Test: Powerful Kernel Testing in Near-linear Time

Carles Domingo-Enrich
New York University, USA, cd2754@nyu.edu

Raaz Dwivedi
Harvard and MIT, USA, raaz@mit.edu

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

Kernel two-sample testing provides a powerful framework for distinguishing any pair of distributions based on n sample points. However, existing kernel tests either run in n^2 time or sacrifice undue power to improve runtime. To address these shortcomings, we introduce Compress Then Test (CTT) [1], a new framework for high-powered kernel testing based on sample compression. Leveraging the recent developments in distribution compress [2, 3, 4], CTT provides a cheap approximation to an expensive test by compressing each n point sample into a small but provably high-fidelity coreset. For standard kernels and subexponential distributions, CTT inherits the statistical behavior of a quadratic-time test—recovering the same optimal detection boundary—while running in near-linear time. We couple these advances with cheaper permutation testing, justified by new power analyses; improved time-vs.-quality guarantees for low-rank approximation; and a fast aggregation procedure for identifying especially discriminating kernels. In our experiments with real and simulated data, CTT and its extensions provide 20–200x speed-ups over state-of-the-art approximate MMD tests with no loss of power.

[1] C. Domingo-Enrich, R. Dwivedi, and L. Mackey. Compress Then Test: Powerful kernel testing in near-linear time. *arXiv preprint arXiv:2301.05974*, 2023.

[2] R. Dwivedi and L. Mackey. Kernel thinning. In *Proceedings of Thirty Fourth Conference on Learning Theory*, volume 134 of *Proceedings of Machine Learning Research*, pages 1753–1753. PMLR, 15–19 Aug 2021.

[3] R. Dwivedi and L. Mackey. Generalized kernel thinning. In *International Conference on Learning Representations*, 2022.

[4] A. Shetty, R. Dwivedi, and L. Mackey. Distribution compression in near-linear time. In *International Conference on Learning Representations*, 2022.

2.3 Parallel sessions - Afternoon

2.3.1 Amphitheater 25 - MS Stochastic Computation and Complexity

Speaker: Randolf Altmeyer

Approximation of occupation time functionals and related approximations of Itô processes

Randolf Altmeyer
University of Cambridge, UK, ra591@cam.ac.uk

Integral functionals of continuous time stochastic processes, e.g., occupation and local times, are relevant in many applications. The problem of approximating them from discrete time data appears naturally in numerical analysis and statistics for stochastic processes. In this talk we will review a number of recent approximation results for Itô processes, including optimality results and central limit theorems. We further discuss how they can be applied to obtain generalised Itô formulas for functions f with fractional Sobolev-regularity.

Speaker: Simon Ellinger

Sharp lower error bounds for strong approximation of SDEs with piecewise Lipschitz continuous drift

Simon Ellinger
University of Passau, Germany, simon.ellinger@uni-passau.de

We study the complexity of pathwise approximation of SDEs at the final time in the presence of discontinuities of the drift coefficient. Recently, it has been shown in [1] that for all $p \in [1, \infty)$ a transformed Milstein-type scheme achieves an L^p-error rate of at least $3/4$ if the drift coefficient is piecewise Lipschitz-continuous with a piecewise Lipschitz-continuous derivative and the diffusion coefficient is constant. This rate is known to be sharp if the drift coefficient has a point of discontinuity and, additionally, is bounded and increasing, see [2]. We show that the latter conditions can be dropped.

[1] T. Müller-Gronbach and L. Yaroslavtseva. *A strong order 3/4 method for SDEs with discontinuous drift coefficient.* IMA Journal of Numerical Analysis, 42(1):229–259, 2022.

[2] T. Müller-Gronbach and L. Yaroslavtseva. *Sharp lower error bounds for strong approximation of SDEs with discontinuous drift coefficient by coupling of noise.* To appear in: Annals of Applied Probability, 2022.

Speaker: Máté Gerencsér

Milstein scheme for SDEs with irregular drift

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

Gerald Lampl
TU Wien,

Chengcheng Ling
TU Wien, Austria, chengcheng.ling@tuwien.ac.at

We discuss strong error rates for the Milstein scheme for SDEs with non-Lipschitz drift coefficients. We show that the rates known for the additive case can be recovered in the multiplicative case when using the Milstein scheme, under appropriate conditions on the diffusion coefficient.

Speaker: Larisa Yaroslavtseva

Sharp lower error bounds for strong approximation of SDEs with a drift coefficient of Sobolev regularity $s \in (1/2, 1)$

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

Larisa Yaroslavtseva
University of Graz, larisa.yaroslavtseva@uni-graz.at

We study strong approximation of scalar SDEs $dX_t = \mu(X_t)\,dt + dW_t$ at time $t = 1$ in the case that μ is bounded and has fractional Sobolev regularity $s \in (0, 1)$. Recently, it has been shown in [1] that in this case the equidistant Euler scheme achieves a root mean squared error of order $(1 + s)/2$, up to an arbitrary small ϵ, in terms of the number of evaluations of the driving Brownian motion W. In this talk we show that, for $s \in (1/2, 1)$, this order can not be improved in general.

[1] K. Dareiotis, M. Gerencsér and K. Lê. Quantifying a convergence theorem of Gyöngy and Krylov. arXiv:2101.12185v2 (2022).

2.3.2 Room 101, corridor 24-25 - MS MCM for reactor physics

Speaker: Mathias Rousset

Fluctuations of Rare Event Simulation with Monte Carlo Splitting in the Small Noise Asymptotics

Frédéric Cérou
IRMAR and INRIA, University of Rennes, France,

Sofiane Martel
Ecole Nationale Des Ponts et Chaussées, France,

Mathias Rousset
IRMAR and INRIA, University of Rennes, France, mathias.rousset@inria.fr

Diffusion processes with small noise conditioned to reach a target set are considered. The AMS algorithm is a Monte Carlo method that is used to sample such rare events by iteratively simulating clones of the process and selecting trajectories that have reached the highest value of a so-called importance function. In this paper, the relative variance of the AMS small probability estimator for large sample size is considered. The main result is a large deviations logarithmic equivalent of the latter in the small noise asymptotics, which will be rigorously derived. Interpretations and practical consequences will be discussed.

Speaker: Davide Mancusi

Variance Reduction and Noise Source Sampling Techniques for Monte Carlo Simulations of Neutron Noise Induced by Mechanical Vibrations

Hunter Belanger
CEA, France, hunter.belanger@gmail.com

Davide Mancusi
CEA, France, davide.mancusi@cea.fr

Amélie Rouchon
CEA, France, amelie.rouchon@cea.fr

Andrea Zoia
CEA, France, andrea.zoia@cea.fr

Neutron noise in nuclear power reactors refers to the small fluctuations around the average neutron flux at steady state resulting from time-dependent perturbations inside the core. The neutron noise equations in the frequency domain can be solved using Monte Carlo simulation codes, which are capable of obtaining reference solutions involving almost no approximations, but are hindered by severe issues affecting the statistical convergence: the simultaneous presence of positive and negative particles, which is required by the nature of the complex noise equations, leads to catastrophically large variance in the tallies. In this work, we consider the important case of neutron noise problems induced by mechanical vibrations. First, we derive a new exact sampling strategy for the noise source. Then, building upon our previous findings in other contexts, we show that weight cancellation methods can be highly beneficial in dealing with the presence of negative weights, enabling extremely large gains in the figure of merit. We successfully demonstrate our results on a benchmark configuration consisting of a fuel assembly with a vibrating pin and we discuss possible pathways for further improvements.

Speaker: Vince Maes

Estimating the statistical error of analog particle-tracing Monte Carlo methods

Vince Maes
KU Leuven, Department of Computer Science, vince.maes@kuleuven.be

Wouter Dekeyser
KU Leuven, Department of Mechanical Engineering, wouter.dekeyser@kuleuven.be

Julian Koellermeier
University of Groningen, Bernoulli Institute, j.koellermeier@rug.nl

Martine Baelmans
KU Leuven, Department of Mechanical Engineering, martine.baelmans@kuleuven.be

Giovanni Samaey
KU Leuven, Department of Computer Science, giovanni.samaey@kuleuven.be

The mesoscopic behaviour of large particle systems is described by kinetic equations that have a high-dimensional phase space. A common approach to deal with the high-dimensionality of the problem is to employ Monte Carlo methods for which the convergence rate is independent of the dimensionality. For kinetic equations with a linear collision operator, modelling for instance collisions of the particles under consideration with a background medium, analog particle-tracing Monte Carlo methods allow for unbiased simulation and estimation of quantities of interest. Such particle-tracing Monte Carlo methods do, however, introduce a statistical error (noise), which depends on the number of particles N used in the simulation. Because of the linearity of the kinetic equation, the N particles can be simulated independently from each other, leading to the well-known result that the statistical error is inversely proportional to the square root of N. The precise value of the statistical error depends on the kinetic equation under consideration.
In this talk, we will show how the statistical error can be calculated (expensive) and estimated (cheap) in the setting of an analog particle-tracing Monte Carlo method for which each particle provides multiple (correlated) contributions to the estimators of quantities of interest. The cheap statistical error estimators can be used to determine the number of particles N that is needed to reach a desired accuracy tolerance. We will apply the theory to a kinetic equation describing neutral particles in the plasma edge of a fusion device and show numerical results.

Speaker: Benjamin Dechenaux

Percolation properties of the neutron population in nuclear reactors

Benjamin Dechanaux
IRSN, France, benjamin.dechenaux@irsn.fr

Thomas Delcambre
CEA, France, thomas.delcambre@cea.fr

Eric Dumonteil
CEA, France, eric.dumonteil@cea.fr

Reactor physics aims at studying the neutron population in a reactor core under the influence of feedback mechanisms, such as the Doppler temperature effect. Numerical schemes to calculate macroscopic properties emerging from such coupled stochastic systems however require to define intermediate quantities (e.g. the temperature field), which are bridging the gap between the stochastic neutron field and the deterministic feedback. By interpreting the branching random walk of neutrons in fissile media under the influence of a feedback mechanism as a directed percolation process and by leveraging on the statistical field theory of birth death processes, we will build a stochastic model of neutron transport theory and of reactor physics. The critical exponents of this model, combined to the analysis of the resulting field equation involving a fractional Laplacian will show that the critical diffusion equation cannot adequately describe the spatial distribution of the neutron population and shifts instead to a critical super-diffusion equation. The analysis of this equation will reveal that non-negligible departure from mean field behavior might develop in reactor cores, questioning the attainable accuracy of the numerical schemes currently used by the nuclear industry.

2.3.3 Room 102, corridor 15-25 - MS PDMPs for high dimensional sampling theory and application

Speaker: Guillaume Chennetier

Adaptive importance sampling based on fault tree analysis for piecewise deterministic Markov process

Guillaume Chennetier
EDF R&D and CMAP, École polytechnique, France, guillaume.chennetier@polytechnique.edu

Hassane Chraibi
EDF R&D, France, hassane.chraibi@edf.fr

Anne Dutfoy
EDF R&D, France, anne.dutfoy@edf.fr

Josselin Garnier
CMAP, École polytechnique, France, josselin.garnier@polytechnique.edu

Piecewise deterministic Markov processes (PDMPs) can be used to model complex dynamical industrial systems. The counterpart of this modeling capability is their simulation cost, because it is necessary to solve the differential equations modeling complex physical phenomena in order to calculate the PDMP flow. That issue makes reliability assessment untractable with standard Monte Carlo methods. Indeed, the failure of a complex system is a rare event and estimating the probability of its occurrence using a Monte Carlo method requires the simulation of a very large number of trajectories of the underlying process. A significant variance reduction can be obtained with a well-calibrated importance sampling method. It is known that the optimal distribution for importance sampling depends explicitly on the committor function of the PDMP. Fault tree analysis offers us elegant tools to approximate this committor function. We present an adaptive importance sampling (AIS) method based on a cross-entropy (CE) procedure for sequentially refining the approximation of the committor function. The method is tested on a system from the nuclear industry: the spent fuel pool, which is described by a PDMP whose discrete variable has more than 32,000 modalities.

[1] Chraibi, H., Dutfoy, A., Galtier, T., & Garnier, J. (2019). On the optimal importance process for piecewise deterministic markov process. *ESAIM: Probability and Statistics, 23*, 893-921.

[2] Chennetier, G., Chraibi, H., Dutfoy, A., & Garnier, J. (2022). Adaptive importance sampling based on fault tree analysis for piecewise deterministic Markov process. *arXiv preprint arXiv:2210.16185.*

[3] Ruijters, E., & Stoelinga, M. (2015). Fault tree analysis: A survey of the state-of-the-art in modeling, analysis and tools. *Computer science review, 15*, 29-62.

Speaker: Joris Bierkens

Analysis of Subsampling Regimes for PDMPs

Sanket Agrawal
University of Warwick, sanket.agrawal@warwick.ac.uk

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

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

Piecewise Deterministic Monte Carlo methods such as the Bouncy Particle Sampler and Zig-Zag have the surprising property that they are capable of asymptotically unbiased subsampling, i.e. we can modify the algorithm so that each iterations only requires the inspection of a subset of the data, without introducing a bias with respect to the target distribution. However important questions remain concerning the computational efficiency of this approach. In this work we aim to clarify some of these questions through the study of scaling limits of PDMC methods in the large data regime.

Speaker: Matthias Sachs

Posterior Computation with the Gibbs Zig-Zag Sampler

Matthias Sachs
University of Birmingham, Birmingham, UK, m.sachs@bham.ac.uk

An intriguing new class of piecewise deterministic Markov processes (PDMPs) has recently been proposed as an alternative to Markov chain Monte Carlo (MCMC). We propose a new class of PDMPs termed Gibbs zig-zag samplers, which allow parameters to be updated in blocks with a zig-zag sampler applied to certain parameters and traditional MCMC-style updates to others. We demonstrate the flexibility of this framework on posterior sampling for logistic models with shrinkage priors for high-dimensional regression and random effects, and provide conditions for geometric ergodicity and the validity of a central limit theorem.

Speaker: Ritabrata Dutta

Efficient Bayesian model averaging for Neural Networks via Gibbs Boomerang

We explore the strength of Markov chain Monte Carlo (MCMC) methods based on piece-wise deterministic Markov processes (PDMP) for the uncertainty quantification of neural network based models. Our sampling machinery builds upon a preconditioned PDMP sampler called Boomerang sampler, which achieves superior mixing due to non-reversibility and can sample exactly while only using the sub-samples of the training data, to deal with the highdimensionality of parameter space and big data bottleneck. To solve the identifiability problem we propose a framework based on Bayesian model averaging using global-local shrinkage priors, more specifically horseshoe priors. We illustrate that in this setting most variational and standard MCMC techniques fail due to the exploding behavior of the gradients of the posterior distribution. Borrowing ideas from the literature of Gibbs PDMPs, we propose Gibbs Boomerang sampler, which does not suffer from the exploding nature of the gradient and further removes the need of estimation of the reference measure to precondition. The only approximation in our method comes while sampling event times from the associated non-homogeneous Poisson process for intractable likelihoods.

2.3.4 Room 104, corridor 15-25 - MS Monte Carlo Methods for Bayesian inference and optimization

Speaker: Savvas Melidonis

Efficient Bayesian computation for low-photon imaging problems

Savvas Melidonis
Heriot-Watt University, United Kingdom, sm2041@hw.ac.uk

Marcelo Pereyra
Heriot-Watt University, United Kingdom, m.pereyra@hw.ac.uk

This talk presents a new and highly efficient MCMC methodology to perform Bayesian inference in low-photon imaging problems, with particular attention to situations involving observation noise processes that deviate significantly from Gaussian noise, such as binomial, geometric and low-intensity Poisson noise. These problems are challenging for many reasons. From an inferential viewpoint, low photon numbers lead to severe identifiability issues, poor stability and high uncertainty about the solution. Moreover, low-photon models often exhibit poor regularity properties that make efficient Bayesian computation difficult; e.g., hard non-negativity constraints, non-smooth priors, and log-likelihood terms with exploding gradients. More precisely, the lack of suitable regularity properties hinders the use of state-of-the-art Monte Carlo methods based on numerical approximations of the Langevin stochastic differential equation (SDE) or other similar dynamics, as both the continuous-time process and its numerical approximations behave poorly. We address this difficulty by proposing an MCMC methodology based on a reflected and regularised Langevin SDE, which is shown to be well-posed and exponentially ergodic under mild

and easily verifiable conditions. This then allows us to derive four reflected proximal Langevin MCMC algorithms to perform Bayesian computation in low-photon imaging problems. The proposed approach is illustrated with a range of experiments related to image deblurring, denoising, and inpainting under binomial, geometric and Poisson noise.

Speaker: Emilie Chouzenoux

PMCnet for Efficient Bayes Inference in Neural Networks

Emilie Chouzenoux
CVN, Inria Saclay, Centrale Supélec, Université Paris-Saclay, France, emilie.chouzenoux@inria.fr

Yunshi Huang
CVN, Inria Saclay, Centrale Supélec, Université Paris-Saclay, France, yunshi.huang@inria.fr

Jean-Chrsitophe Pesquet
CVN, Inria Saclay, Centrale Supélec, Université Paris-Saclay, France, jean-christophe@pesquet.eu

Bayesian neural networks (BNNs) have received an increased interest in the last years. In BNNs, a complete posterior distribution of the unknown weight and bias parameters of the network is produced during the training stage. This probabilistic estimation offers several advantages with respect to point-wise estimates, in particular, the ability to provide uncertainty quantification when predicting new data. This feature inherent to the Bayesian paradigm, is useful in countless machine learning applications. It is particularly appealing in areas where decision-making has a crucial impact, such as medical healthcare. In this work, we introduce a novel algorithm PMCnet, relying on population Monte Carlo methodology, that includes an efficient adaptation mechanism, exploiting geometric information on the complex (often multimodal) posterior distribution. Numerical results on several medical datasets illustrate the good performance of the method.

[1] Y. Huang, E. Chouzenoux, V. Elvira and J.-C. Pesquet. *Efficient Bayes Inference in Neural Networks through Adaptive Importance Sampling.* Tech. Rep. 2022, https://arxiv.org/abs/2210.00993

Speaker: Mohamed Fakhfakh

Hamiltonian Monte Carlo Bayesian Optimization for Sparse Neural Networks

Mohamed Fakhfakh
Toulpouse INP, University of Toulouse, IRIT, France, mohamed.fakhfakh@toulouse-inp.fr

Lotfi Chaari
Toulpouse INP, University of Toulouse, IRIT, France, lotfi.chaari@toulouse-inp.fr

The performance of a deep neural network strongly depends on the optimization method used during the learning process. The essence of most architectures is to build an optimization model and learn the parameters from the available training data.
In this sense, regularization is usually employed for the sake of stability or uniqueness of the solution. When non-smooth regularizers are used, especially to promote sparse networks [1, 2], such as the ℓ_1 norm, this optimization becomes difficult due to the non-differentiability of the target criterion, which may also be non-convex.
We propose a Bayesian optimization framework based on an MCMC scheme that allows efficient sampling even for non-smooth energy function. We demonstrate that using the proposed method for image classification leads to high-accuracy results that cannot be achieved using classical optimizers.

[1] L. Chaari and J.-Y. Tourneret and C. Chaux and H. Batatia. *A Hamiltonian Monte Carlo Method for Non-Smooth Energy Sampling.* IEEE Trans. on Signal Process, vol 64, no. 21, pp. 5585 - 5594, 2016

[2] D. C. Mocanu and E. Mocanu and P. Stone, and P. H. Nguyen and M. Gibescu and A. Liotta. *Scalable training of artificial neural networks with adaptive sparse connectivity inspired by network science.* Communications, vol 9, pp. 1-12, 2018.

Speaker: Nadège Polette

Bayesian Inference for Inverse Problems with Hyperparameters Estimation of the Field Covariance Function

Nadège Polette
CEA, DAM, DIF, F-91297 Arpajon, France/Mines Paris, Université PSL, centre de Géosciences, France,
nadege.polette@cea.fr

Pierre Sochala
CEA, DAM, DIF, F-91297 Arpajon, France, pierre.sochala@cea.fr

Olivier Le Maître
Centre de Mathématiques appliquées, CNRS and INRIA, Ecole Polytechnique, France,
olivier.le-maitre@polytechnique.edu

Alexandrine Gesret
MINES Paris, Université PSL, Centre de Géosciences, France,
alexandrine.gesret@minesparis.psl.eu

Inverse problems are encoutered in many applications whenever one search for information about a physical system based on measurements [6]. The Bayesian inference is an attractive approach for solving such type of problem since it provides a full estimation of the unknown parameters distributions. However, the convergence of the posterior distribution sampled with Markov Chain Monte–Carlo (MCMC) [1] can be difficult to reach especially when dealing with a high-dimensional search space.
In this work, we are interested in estimating a physical field by using a set of undirect observations. The MCMC sampling is accelerated thanks to the replacement of the forward model predictions by surrogate models based on polynomial chaos expansions [7, 2]. In order to reduce the input dimension of the surrogate model, a Karhunen–Loève decomposition is performed, considering that the field of interest is a particular realization of a Gaussian random field. In practice, this representation relies on the choice of hyperparameters, that can be inferred during the MCMC sampling. In [5], a method is developed to mitigate the cost of this supplementary inference by introducing a coordinates transformation. We suggest an alternative method based on a change of measure. The advantage is that the introduced surrogate quantity is smooth along hyperparameters, in contrary to the coordinates transformation one [5, 3].
This change of measure method is applied to a seismic tomography problem, where we infer a seismic wave velocity field with the first-arrival traveltimes at given locations. This application to a continuous velocity model generalizes the work on a layered velocity model realized in traveltime tomography [4]. More realistic predictions than when fixing the hyperparameters at constant values are obtained. Moreover, this method allows for various field shapes, while keeping the implementation computationally tractable.

[1] A. Doucet, A. Smith, N. de Freitas, and N. Gordon. *Sequential Monte Carlo Methods in Practice.* Information Science and Statistics. Springer New York, 2013.

[2] R. G. Ghanem and P. D. Spanos. *Stochastic Finite Element Method: Response Statistics*, pages 101–119. Springer New York, 1991.

[3] A. Siripatana, O. Le Maître, O. Knio, C. Dawson, and I. Hoteit. Bayesian inference of spatially varying Manning's n coefficients in an idealized coastal ocean model using a generalized Karhunen-Loève expansion and polynomial chaos. *Ocean Dynamics*, 70, 2020.

[4] P. Sochala, A. Gesret, and O. Le Maître. Polynomial surrogates for bayesian traveltime tomography. *GEM - International Journal on Geomathematics*, 12(1), December 2021.

[5] I. Sraj, O. P. Le Maître, O. M. Knio, and I. Hoteit. Coordinate transformation and polynomial chaos for the bayesian inference of a gaussian process with parametrized prior covariance function. *Computer Methods in Applied Mechanics and Engineering*, 298:205–228, 2016.

[6] A. Tarantola. *Inverse Problem Theory and Methods for Model Parameter Estimation*, volume xii. Society for Industrial and Applied Mathematics, 01 2005.

[7] N. Wiener. The homogeneous chaos. *American Journal of Mathematics*, 60(4):897–936, 1938.

3 Wednesday 28th June

3.1 Keynote speakers

Speaker: Arnaud Doucet

From Denoising Diffusion Models to Dynamic Transport Models – Generative Modeling and Inference

Arnaud Doucet
Google DeepMind, arnauddoucet@google.com

Denoising diffusion models are a novel powerful class of techniques for generative modeling and inference. These models have superseded generative adversarial networks over the past two years as they are flexible, easy to train and provide state-of-the-art results in numerous application domains such as image synthesis and protein design. In this talk, we will review these methods, illustrate them on a variety of applications and discuss their limitations. We will then show how recent alternative techniques based on dynamic mass transport ideas can resolve some of these limitations. In particular, we will focus on Schrödinger bridges, an entropy-regularized version of the dynamic optimal transport, and present a novel simple method to approximate them numerically.

Speaker: Linda Petzold

The Roles and Consequences of Randomness in Biological Systems

Linda Petzold
University of California Santa Barbara, USA, petzold@ucsb.edu

Stochasticity (randomness) is ubiquitous in biological systems. We will explore some of the ways in which it arises and is used to advantage by biological systems, at a wide range of scales.

3.2 Parallel sessions - Morning

3.2.1 Amphitheater 25 - MS Stochastic Computation and Complexity

Speaker: Daniel Rudolf

Convergence and well-definedness of ellitpical slice sampling

Daniel Rudolf
University of Passau, daniel.rudolf@uni-passau.de

Elliptical slice sampling (ESS) provides a transition mechanism that leads to a Markov chain for approximate sampling of a posterior distribution in a Gaussian prior setting. In applications it is appreciated that ESS avoids local random walk behavior, there is no tuning of a step-size parameter and a richer choice of possible updates is available. We discuss the well-definedness of ESS and provide a geometric convergence result.

Speaker: Stefan Heinrich

Randomized Complexity of Vector-Valued Approximation

Stefan Heinrich
RPTU Kaiserslautern-Landau, heinrich@informatik.uni-kl.de

We study the randomized n-th minimal errors (and hence the complexity) of vector valued approximation. In a recent paper by the author [1] a long-standing problem of Information-Based Complexity was solved: Is there a constant $c > 0$ such that for all linear problems $\mathcal{P}$ the randomized non-adaptive and adaptive n-th minimal errors can deviate at most by a factor of c? That is, does the following hold for all linear $\mathcal{P}$ and $n \in \mathbf{N}$

$$e_n^{\mathrm{ran-non}}(\mathcal{P}) \leq c e_n^{\mathrm{ran}}(\mathcal{P})\,?$$

The analysis of vector-valued mean computation showed that the answer is negative. More precisely, there are instances of this problem where the gap between non-adaptive and adaptive randomized minimal errors can be (up to log factors) of the order $n^{1/8}$. This raises the question about the maximal possible deviation. In this talk we show that for certain instances of vector valued approximation the gap is $n^{1/2}$ (again, up to log factors).

[1] S. Heinrich, *Randomized Complexity of Parametric Integration and the Role of Adaption I. Finite Dimensional Case* (preprint)

Speaker: Sonja Cox

Infinite-dimensional Wishart processes

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

Christa Cuchiero
University of Vienna, christa.cuchiero@univie.ac.at

Asma Khedher
University of Amsterdam, a.khedher@uva.nl

A Wishart process is a time-homogeneous Markov process $(X_t)_{t\geq 0}$ taking values in the space of positive semi-definite matrices such that X_t has a (generalized) Wishart distribution for every $t \geq 0$. Wishart processes were introduced in the '90s by Bru and have become a popular choice for modelling stochastic covariance. For example, Wishart processes are used in multi-dimensional Heston models to describe the instantaneaous volatility in a multi-dimensional stochastic differential equation. However, models for energy and interest rate markets involve stochastic *partial* differential equations, and thus call for infinite-dimensional covariance models. In our work, we introduce and analyze infinite-dimensional Wishart processes, and discuss some of their advantages and short-comings.

Speaker: Thomas Kühn

High-dimensional approximation in Sobolev spaces of anisotropic mixed smoothness

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

The subject of this talk are embeddings of Hilbert-Sobolev spaces with anisotropic mixed smoothness on the d-dimensional torus. We study the behaviour of the singular numbers for embeddings of these spaces into L_2 and H^1, with special emphasis on the influence of the (arbitrarily high) dimension d of the underlying domain. Asymptotic rates in this setting are known for a long time, but only up to unspecified multiplicative constants. We determine *exact asymptotic constants* and provide *explicit preasymptotic bounds*.
The talk is based on the joint paper [1] with W. Sickel (Jena) and T. Ullrich (Chemnitz).

[1] T. Kühn, W. Sickel and T. Ullrich *How anisotropic mixed smoothness affects the decay of singular numbers for Sobolev embeddings.* J. Complexity 63 (2021), Paper No. 101523, 37 pp.

3.2.2 Room 101, corridor 24-25 - MS High dimensional approximation

Speaker: James A. Nichols

Community detection with entropic regularization

James A. Nichols
Australian National University, Australia, james.nichols@anu.edu.au

Detecting community, or cluster structure, in networks (i.e. graphs) is a common task in a variety of scientific and commercial settings. For example, in the biological sciences, the increasing use of high-throughput assays generate large quantities of genomic data presented in graphical form, and efficiently finding correlative clusters is of paramount importance.
We propose an approach that involves sorting the graph, re-labelling nodes in such a manner as to make neighbouring nodes maximally connected. We use an objective function that penalises off-diagonal components in the re-labelled graph adjacency matrix. Sorting a graph adjacency matrix boils down to a type of *quadratic assignment problem*, known to be equivalent to the travelling salesman problem, and hence NP-hard. We propose a method of adding an entropic regularization term to the objective function, and optimizing with an alternating descent Sinkhorn-type algorithm. This method has been popularized recently in the computational optimal transport literature, but has some history of being used in relation to quadratic assignment problems.
This approach delivers a scalable method that permutes the rows and columns of a matrix, making the matrix *maximally diagonal*. It allows us to do quick community detection, and further results in a permuted matrix that is amenable to a hierarchical (H-)matrix representation.

Speaker: Alexander Gilbert

Density estimation in uncertainty quantification using quasi-Monte Carlo methods with preintegration

Alexander Gilbert
UNSW Sydney, Australia, alexander.gilbert@unsw.edu.au

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

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

Abirami Srikumar
UNSW Sydney, Australia, a.srikumar@student.unsw.edu.au

Quasi-Monte Carlo (QMC) methods have previously shown great success in tackling difficult high-dimensional problems that often occur in uncertainty quantification. However, a key limitation is that they are limited to approximating only the expected value of the quantity of interest. One of the main reasons for this limitation are the smoothness requirements, such as requiring square-integrable mixed first derivatives. In this talk, we present a method for approximating the cumulative distribution function (cdf) and probability density function (pdf) of a quantity of interest coming from the solution of an elliptic PDE with lognormal random coefficients. The key idea is to formulate the cdf (and pdf) as an expected value, or equivalently, a high-dimensional integral, which can then be efficiently approximated by a QMC method. Typically QMC methods struggle to efficiently approximate the cdf because of a lack of smoothness in the

integrand, which for a cdf is an indicator function. We overcome this by using an initial preintegration step to smooth out the integrand. Preintegration, also known as conditional expectation, is a method for smoothing a discontinuous function by integrating with respect to a single specially chosen variable. The result is a function in one dimension less that is now smooth (under appropriate assumptions of course). We will outline the QMC with preintegration method for approximating the cdf and pdf for lognormal PDEs, then present an error analysis and numerical results.

Speaker: Michael Feischl

A quasi-Monte Carlo data compression algorithm for machine learning

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

We present an algorithm from [1] to reduce large data sets using so-called digital nets, which are well distributed point sets in the unit cube. The algorithm efficiently scans the data and computes certain data dependent weights. Those weights are used to approximately represent the data, without making any assumptions on the distribution of the data points. Under smoothness assumptions on the model, we then show that this can be used to reduce the computational effort needed in finding good parameters in machine learning problems which aim to minimize quantities of the form

$$\min_{\theta} \frac{1}{n} \sum_{i=1}^{n} (f_\theta(x_i) - y_i)^2.$$

While the principal idea of the approximation might also work with other point sets, the particular structural properties of digital nets can be exploited to make the computation of the necessary weights extremely fast.

[1] J. Dick, M. Feischl, A quasi-Monte Carlo data compression algorithm for machine learning, Journal of Complexity, Volume 67, 2021.

Speaker: Vesa Kaarnioja

On the Periodic Model of Uncertainty Quantification With Application to Bayesian Inverse Problems

Vesa Kaarnioja
Free University of Berlin, vesa.kaarnioja@fu-berlin.de

The essence of statistical inversion lies in modeling the unknown parameters of a mathematical measurement model using random fields. In the study of forward uncertainty quantification for elliptic partial differential equations (PDEs) with random coefficients, [1] considered a model in which a countable number of independent random variables enter the input random field as periodic functions. This model enables the development of simple quasi-Monte Carlo (QMC) methods, which achieve higher-order cubature convergence rates for high-dimensional numerical integration problems associated with the output of the PDE problem subject to input uncertainties. Moreover, it was shown in [2] that this model also enables the construction of a kernel interpolant for the PDE solution as a function of the stochastic variables in a highly efficient manner. In this talk, I will discuss the application of QMC integration and kernel interpolation for Bayesian inverse problems governed by elliptic PDEs under the periodic paradigm. Numerical experiments are presented to assess the effectiveness of this approach.

[1] V. Kaarnioja, F. Y. Kuo, and I. H. Sloan. *Uncertainty quantification using periodic random variables.* SIAM J. Numer. Anal. 58, 1068–1091, 2020.

[2] V. Kaarnioja, Y. Kazashi, F. Y. Kuo, F. Nobile, and I. H. Sloan. *Fast approximation by periodic kernel-based lattice-point interpolation with application in uncertainty quantification.* Numer. Math. 150, 33–77, 2022.

3.2.3 Room 102, corridor 15-25 - MS Variance reduction techniques for rare events

Speaker: Eya Ben Amar

State-dependent Importance Sampling for Estimating Expectations of Functionals of Sums of Independent Random Variables

Eya Ben Amar
KAUST, Saudi Arabia, eya.benamar@kaust.edu.sa

Nadhir Ben Rached
University of Leeds, UK, n.benrached@leeds.ac.uk

Abdul-Lateef Haji-Ali
Heriot-Watt University, UK, A.HajiAli@hw.ac.uk

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

Estimating the expectations of functionals applied to sums of random variables (RVs) is a well-known problem encountered in many challenging applications. Generally, closed-form expressions of these quantities are out of reach. A naive Monte Carlo simulation is an alternative approach. However, this method requires numerous samples for rare event problems. Therefore, it is paramount to use variance reduction techniques to develop fast and efficient estimation methods. In this work, we use importance sampling (IS), known for its efficiency in requiring fewer computations to achieve the same accuracy requirements. We propose a state-dependent IS scheme based on a stochastic optimal control formulation, where the control is dependent on state and time. We aim to calculate rare event quantities that could be written as an expectation of a functional of the sums of independent RVs. The proposed algorithm is generic and can be applied without restrictions on the univariate distributions of RVs or the functional applied to the sum. We apply this approach to the log-normal distribution to compute the left tail and cumulative distribution of the ratio of independent RVs. For each case, we numerically demonstrate that the proposed state-dependent IS algorithm compares favorably to most well-known estimators dealing with similar problems.

Speaker: Gerardo Rubino

Estimating network resilience, a performability metric

Gerardo Rubino
INRIA, gerardo.rubino@inria.fr

Instead of considering multi-component systems in a binary world, where systems and components are either up (operational, working) or down (unoperational, failed), we move to a multi-variate one, where the system's state space is partitioned into several performance levels, while components are as usual binary ones, subject to failures and possibly repairs. This allows defining performability metrics, where performance and dependability aspects are simultaneously taken into account. Consider an undirected graph whose nodes and/or edges can be failed or working, and where we are interested in the existence of paths composed only of working elements. Using a static (no time variable) setting, we consider the resilience metric that extends the classical connectivity-based ones such as the source-to-terminal reliability, or the all-terminal reliability. Resilience is defined as the expectation of the number of pairs of nodes that can communicate. After analyzing some basic examples, we briefly describe a Monte

Carlo approach where instead of reducing the variance of the estimators, we focus on reducing their time complexities. This view allows a first straightforward way of exploring resilience (as well as many other classical metrics). It also allows easily performing a sensitivity analysis with respect to the individual reliabilities of the components, without a significant overhead of the procedure that estimates the resilience metric alone. The talk is based on the paper cited below.

[1] G. Rubino. *Network Reliability, Performability Metrics, Rare Events and Standard Monte Carlo* In "Advances in Modeling and Simulation: Festschrift for Pierre L'Ecuyer", pp. 401-420, edited by Zdravko I, Botev, Alexander Keller, Christiane Lemieux and Bruno Tuffin, Springer Nature, 2022.

Speaker: Martin Chak

Optimal friction in Langevin dynamics

Martin Chak
LJLL, France, martin.chak@sorbonne-universite.fr

An adaptive method is given for underdamped Langevin dynamics to minimize asymptotic variance with respect to the friction parameter. The procedure is based on a formula for the gradient of the variance in terms of solutions to the Poisson equation. Reduced variance is demonstrated on toy and Bayesian inference problems.

3.2.4 Room 104, corridor 15-25 - MS Recent Progress in Langevin MC

Speaker: Alain Durmus

On the convergence of the Unadjusted and Metropolis Adjusted Langevin Algorithms

Alain Durmus
Ecole Polytechnique, alain.durmus@polytechnique.edu

One of the main questions in the theory of Markov Chains Monte Carlo is to bound or estimate the rate at which the underlying chain, or its ergodic averages, converges to a stationary measure, since this rate determines the efficiency of an algorithm. The primary goal of this talk is to discuss some recent advances in estimating convergence of Langevin MCMC methods including the Unadjusted Langevin Algorithm (ULA) and its Metropolized version MALA. In particular, we will show that using mainstays of MCMC analysis, i.e. drift-and-minorization, imply sharp convergence bounds for ULA, and as a corollary provide a new approach to establish quantitative convergence for Langevin diffusion. Regarding MALA, while it is a popular and widely used Markov chain Monte Carlo method, very few papers derive conditions that ensure its convergence. In particular, assumptions that are both easy to verify and guarantee geometric convergence, are still missing. Here, we provide mild assumptions about the target distribution implying that MALA is V-uniformly geometric convergence. Finally, we pay special attention to the dependence of the bounds we derive on the step size of the Euler-Maruyama discretization, which corresponds to the proposal Markov kernel of MALA.

Speaker: Tyler Farghly

Adaptive Langevin Monte Carlo methods for heavy-tailed sampling via weighted functional inequalities

Tyler Farghly
University of Oxford, tyler.farghly@keble.ox.ac.uk

The non-asymptotic analysis of Langevin Monte Carlo (LMC) is a subject that has received increased attention within computational statistics, with a focus on theoretical guarantees in high-dimensional settings. However, most existing works require that the target distribution is sufficiently light-tailed. When applied to heavy-tailed targets, guarantees for LMC are limited. In fact, existing analyses of the Langevin diffusion, from which LMC is derived, have only established rates with exponential dependence on the dimension. In this work, we propose a simple generalisation of LMC that employs a weighting function that is chosen according to the tail-growth of the target. The algorithm is based on an adaptive Euler-Maruyama (EM) discretisation of the natural diffusion associated with the weighting. Using weighted logarithmic Sobolev inequalities, we establish non-asymptotic rates in both Wasserstein distance and KL divergence that depend polynomially on dimension. Our analysis applies to a class of heavy-tailed targets interpolating between Gaussian and generalised Cauchy distributions. As part of our analysis, we develop a framework for analysing LMC-type algorithms that use adaptive EM schemes based on the theory of random time changes.

Speaker: Sifan Liu

Langevin Quasi-Monte Carlo

Sifan Liu
Stanford University, sfliu@stanford.edu

Langevin Monte Carlo (LMC) and its stochastic gradient versions are powerful algorithms for sampling from complex high-dimensional distributions. To sample from the distribution with density $\pi(x) \propto \exp(-U(x))$, LMC iteratively generates the next sample by taking a step in the gradient direction ∇U with a Gaussian perturbation. Expectations w.r.t. the target distribution π are estimated by averaging over LMC samples. In ordinary Monte Carlo, it is well known that the estimation error can be substantially reduced by replacing independent random samples by quasi-random samples like low-discrepancy sequences. In this work, we show that the estimation error of LMC can also be reduced by using quasi-random samples. Specifically, we propose to use completely uniformly distributed sequences with certain low-discrepancy property to generate the Gaussian perturbations. Under smoothness and convexity conditions, we prove that LMC with quasi-random samples achieves smaller error than standard LMC. We provide rigorous theoretical analysis supported by compelling numerical experiments to demonstrate the effectiveness of our approach.

Speaker: Konstantinos Zygalakis

Accelerating MCMC for imaging science by using an implicit Langevin algorithm

Konstantinos Zygalakis
University of Edinburgh, K.Zygalakis@ed.ac.uk

In this work, we present a highly efficient gradient-based Markov chain Monte Carlo methodology to perform Bayesian computation in imaging problems. Similarly to previous Monte Carlo approaches, the proposed method is derived from a discretisation of the Langevin diffusion. However, instead of a conventional explicit discretisation like Euler-Maruyama, here we use an implicit discretisation based on the θ-method. In particular, each step of the proposed algorithm involves the solution of an optimisation problem. In the case of a log-concave posterior, this optimisation problem is strongly convex and thus can be solved efficiently, leading to effective step sizes that are significantly larger than traditional methods

permit. We can show that even for these large step sizes the corresponding Markov Chain has low bias while also preserving the posterior variance. We demonstrate the proposed methodology on a range of problems including non-blind image deconvolution and denoising. Comparisons with state-of-the-art MCMC methods confirm that the Markov chains generated with our method exhibit significantly faster convergence speeds, achieve larger effective sample sizes, and produce lower mean square estimation errors at equal computational budget.

3.3 Parallel sessions - Afternoon

3.3.1 Amphitheater 25 - MS Numerical methods in statistical physics

Speaker: Juliane U. Klamser

Can Monte Carlo methods be used to simulate active-matter systems?

Juliane U. Klamser
Laboratoire Charles Coulomb, CNRS/ Université de Montpellier, France,
juliane.klamser@umontpellier.fr

Olivier Dauchot
Gulliver, CNRS/ESPCI Paris-PSL, Paris, France, olivier.dauchot@espci.psl.eu

Julien Tailleur
Laboratoire MSC, CNRS/Université de Paris, France, julien.tailleur@univ-paris-diderot.fr

"All known life forms are based on self-propelled entities uniting to create large-scale structures and movements. If this didn't happen, organisms would be limited to using much slower, passive processes such as diffusion to move DNA and proteins around inside cells or tissues, and many of life's complex structures and functions might never have evolved."[1] A central question in the field of active matter concerns the emergent collective phenomena when individual particles have the ability to move persistently, i.e. when particles overcome a characteristic finite distance without changing their direction of motion. Although considerable effort has been put to develop analytical approaches to describe the statistical physics of active matter, the state of the art is far from comparable approaches in equilibrium statistical physics. Our advances therefore mainly rely on numerical studies where many active-matter models have been proposed and simulated. However, little attempts have been made to develop an algorithmic toolbox for those models. In equilibrium, the detailed-balance condition allows to exploit the unphysical moves of Monte Carlo (MC) approaches to efficiently simulate large systems. As there is no analog of detailed balance for active matter, the construction of MC algorithms that faithfully capture continuous-time active-matter models is not straight forward. I will present a realisation of kinetic MC analogues of the work-horse models of self-propelled particles, namely Active-Ornstein Uhlenbeck, Active Brownian, and Run-and-Tumbles particles.[2]

[1] G. Popkin *The physics of life.* Nature, 529, 16–18, 2016.

[2] J. U. Klamser, O. Dauchot, J. Tailleur *Kinetic Monte Carlo Algorithms for Active Matter Systems* Phys. Rev. Lett., 127, 150602, 2021.

Speaker: Noé Blassel

Stochastic Norton Dynamics

Noé Blassel
Ecole des Ponts ParisTech et Inria Paris, France, noe.blassel@enpc.fr

Gabriel Stoltz
Ecole des Ponts ParisTech et Inria Paris, France, gabriel.stoltz@enpc.fr

Transport coefficients quantify the sensitivity of a flux to the magnitude of a force driving a system out of thermodynamic equilibrium, in the small force regime in which the response is linear. Examples of these include the mobility, thermal conductivity or shear viscosity.
Whereas, at the macroscopic level, forces and fluxes play symmetric roles, standard approaches in computational statistical physics fix the force and measure the average response flux. In contrast, the Norton method, proposed by Evans & al., proposes fixing the flux, and measuring the average magnitude of the force needed to induce it.
At the dynamical level, this amounts to consider dynamics constrained to remain on a submanifold of the phase space, defined by the constant-flux condition. However, the constraint is typically non-holonomic, thus existing techniques for constrained dynamics cannot be directly applied.
The original method is defined in a deterministic setting in which crucial properties such as ergodicity may fail. In this work we propose a more general and stochastic formulation of the Norton method, also providing evidence that it is both valid and sometimes more efficient than the standard approach to transport coefficient computations.

[1] Evans, D., Hoover, W., Failor, B., Moran, B. & Ladd, A. Nonequilibrium molecular dynamics via Gauss's principle of least constraint. *Physical Review A*. **28**, 1016-1021 (1983,8)

[2] Evans, D. & Ely, J. Viscous flow in the stress ensemble. *Molecular Physics*. **59**, 1043-1048 (1986,12)

[3] Gosling, E., McDonald, I. & Singer, K. On the calculation by molecular dynamics of the shear viscosity of a simple fluid. *Molecular Physics*. **26**, 1475-1484 (1973,12)

[4] Joubaud, R. & Stoltz, G. Nonequilibrium Shear Viscosity Computations with Langevin Dynamics. *Multiscale Modeling & Simulation*. **10**, 191-216 (2012,1)

[5] Lelièvre, T., Rousset, M. & Stoltz, G. Langevin dynamics with constraints and computation of free energy differences. *Mathematics Of Computation*. **81**, 2071-2125 (2012,10)

Speaker: Thomas Pigeon

Adaptive multilevel splitting used to machine learn committor function

Thomas Pigeon
MATHERIALS team-project, Inria Paris, France, thomas.pigeon@inria.fr

Gabriel Stoltz
École des Ponts ParisTech et Inria Paris, France, gabriel.stoltz@enpc.fr

Tony Lelièvre
École des Ponts ParisTech et Inria Paris, France, tony.lelievre@enpc.fr

The study of transitions between metastable states in molecular dynamics is central to the study of chemical and biological systems described at the molecular level. In the context of transition path theory, the committor function is a key object to characterize the system's dynamics [1]. Using the fact that it satisfies a backward Kolmogorov equation, various variational approaches were proposed using artificial neural networks to learn this function. These approaches require a preliminar sampling of the invariant measure of the system, which makes these method difficult to apply in rare event scenarios. To tackle this problem, the methods are generally coupled to an importance sampling method which often requires a prior knowledge of the zones of space in which the sampling should be enhanced [2]. In the present work, we propose an iterative approach to learn the committor function. It uses the fact that the Adaptive Multilevel Splitting (AMS) allows to produce some local approximations of the committor function at the intersection of sampled transition trajectories and level-sets of a putative reaction coordinate. In our approach, this function is dsecribed by a neural network which is successively used to sample committor approximations with the AMS algorithm and then trained to reproduce these approximations. The approach is illustrated on a few representative systems.

[1] W. E. and E. Vanden-Eijnden *Towards a Theory of Transition Paths.* Journal of Statistical Physics, 503:523, 123, 2006.

[2] G. M. Rotskoff and E. Vanden-Eijnden. *Learning with rare data: Using active importance sampling to optimize objectives dominated by rare events.* arXiv preprint arXiv:2008.06334, 2020.

Speaker: Gideon Simpson

Infinite Dimensional Nonlocal Diffusions with Additive Noise

Georgi Medvedev
Drexel University, USA, gsm29@drexel.edu

Gideon Simpson
Drexel University, USA, grs53@drexel.edu

This talk will present recent work on infinite dimensional nonlocal diffusions with additive noise. These may be obtained as the continuum limits of noise driven Kuramoto oscillator systems. A well-posedness theory is developed for this infinite dimensional problem, and convergence results are obtained for both the associated semi-discrete and fully discrete problems. This provides a basis for studying the associated metastability problems of the continuum limit Kuramoto system, which may also be viewed as an approximation of a high, but finite, dimensional problem. Ongoing progress on studying the metastability of exact twisted state solutions, both numerically and analytically will be presented. Novel challenges in the analysis are also highlighted.

3.3.2 Room 101, corridor 24-25 - MS Multilevel MC techniques for discontinuous functionals

Speaker: Chiheb Ben Hammouda

MLMC Combined with Numerical Smoothing for Efficient Probabilities Computation, Density Estimation, and Option Pricing

Christian Bayer
Weierstrass Institute for Applied Analysis and Stochastics (WIAS), Berlin, Germany,
christian.bayer@wias-berlin.de

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

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

The multilevel Monte Carlo (MLMC) is a highly effcient approach to estimate expectations of a functional of a solution to a stochastic differential equation. However, MLMC estimators may be unstable and have a nonoptimal complexity in case of low regularity of the observable. To overcome this issue, we extend our idea of numerical smoothing, introduced in [1] in the context of deterministic quadrature methods, to the MLMC setting. The numerical smoothing technique is based on root finding methods combined with one-dimensional numerical integration with respect to a single well-chosen variable. Motivated by probabilities computation, density estimation, and option pricing problems, our theoretical analysis and numerical experiments show that the employed numerical smoothing significantly improves the complexity and robustness (making the kurtosis at deep levels bounded) of the MLMC method. In particular, we recover the MLMC complexities obtained for Lipschitz functionals. Moreover, our approach effciently estimates density functions, a task that previous methods based on Monte Carlo or MLMC fail to achieve at least in moderate to high dimensions.

[1] Christian Bayer, Chiheb Ben Hammouda and Raúl Tempone. *Numerical smoothing with hierarchical adaptive sparse grids and quasi-Monte Carlo methods for efficient option pricing.* Quantitative Finance, 1-19, 2022.

[2] Christian Bayer, Chiheb Ben Hammouda and Raúl Tempone. *Multilevel Monte Carlo Combined with Numerical Smoothing for Robust and Efficient Option Pricing and Density Estimation.* ArXiv preprint arXiv:2003.05708, 2022.

Speaker: Abdul-Lateef Haji-Ali

Multilevel Path Branching for Digital Options

Michael Giles
Mathematical Institute, University of Oxford, UK, mike.giles@ox.ac.uk

Abdul-Lateef Haji-Ali
School of Mathematical and Computer Sciences, Heriot-Watt University, Edinburgh, UK, a.hajiali@hw.ac.uk

We propose a novel Monte Carlo-based estimator for digital options with assets modelled by stochastic differential equation (SDEs) that are solved approximately using a time-stepping scheme such as Euler-Maruyama or Milstein. The new estimator is based on repeated, hierarchical path splitting and we show that it has an improved strong convergence rate with respect to the time-step. We also show that the computational complexity of Multilevel Monte Carlo (MLMC) using the new estimator is similar to the complexity of a classical MLMC estimator when applied to smoother options and is orders of magnitude less than the computational complexity of a traditional Monte Carlo estimator.

Speaker: Ahmed Kebaier

The interpolated drift implicit Euler scheme Multilevel Monte Carlo method for pricing Barrier options and applications to the CIR and CEV models

Mouna Ben Derouich
Université Sorbonne Paris Nord, LAGA, CNRS, France, benderouiche@math.univ-paris13.fr

Ahmed Kebaier
Laboratoire de Mathématiques et Modélisation d'Evry, CNRS, Université d'Evry, Université Paris-Saclay, France, ahmed.kebaier@univ-evry.fr

Recently, Giles et al. [1] proved that the efficiency of the Multilevel Monte Carlo (MLMC) method for evaluating Down-and-Out barrier options for a diffusion process $(X_t)_{t\in[0,T]}$ with globally Lipschitz coefficients, can be improved by combining a Brownian bridge technique and a conditional Monte Carlo method provided that the running minimum $\inf_{t\in[0,T]} X_t$ has a bounded density in the vicinity of the barrier. In the present work, thanks to the Lamperti transformation technique and using a Brownian interpolation of the drift implicit Euler scheme of Alfonsi [2], we show that the efficiency of the MLMC can be also improved for the evaluation of barrier options for models with non-Lipschitz diffusion coefficients under certain moment constraints. We study two example models: the Cox-Ingersoll-Ross (CIR) and the Constant of Elasticity of Variance (CEV) processes for which we show that the conditions of our theoretical framework are satisfied under certain restrictions on the models parameters. In particular, we develop semi-explicit formulas for the densities of the running minimum and running maximum of both CIR and CEV processes which are of independent interest. Finally, numerical tests are processed to illustrate our results.

[1] Michael Giles, Kristian Debrabant, and Andreas Rössler. *Analysis of multilevel Monte Carlo path simulation using the Milstein discretisation.* Discrete & Continuous Dynamical Systems-B 24.8 (2019): 3881.

[2] Aurélien Alfonsi *Strong order one convergence of a drift implicit Euler scheme: Application to the CIR process* Statistics & Probability Letters 83.2 (2013): 602-607.

Speaker: Andreas Stein

An antithetic multilevel Monte Carlo Milstein scheme for SPDEs

Andreas Stein
ETH Zürich, Switzerland, andreas.stein@sam.math.ethz.ch

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

We develop a novel multilevel Monte Carlo (MLMC) Milstein scheme to estimate quantities of interest for stochastic partial differential equations (SPDEs). Taking the cue from [1], we generalize the antithetic Milstein scheme for SDEs from the finite-dimensional setting to Hilbert space-valued SPDEs. The proposed scheme combines the key advantages of Euler and Milstein discretizations: It is straightforward to implement, as no intractable Lévy area terms arise, while the antithetic correction yields the same variance decay in a MLMC algorithm as the standard Milstein method. Consequently, the computational complexity to achieve a prescribed mean-squared error in the MLMC method is significantly lower as for a corresponding MLMC Euler scheme. We do no require any commutative properties on the diffusion coefficient, as for the classic Milstein scheme, hence our approach is applicable for a broader range of non-linear diffusion coefficients.
The key part of this MLMC algorithm consists of a Milstein-type time stepping scheme for SPDEs, where the intractable higher-order correction term is replaced by a truncated, but tractable version. This

truncated Milstein scheme itself does not improve mean-square convergence rates. However, it accelerates the rate of variance decay in the MLMC method, when combined with an antithetic coupling on the fine scales. To obtain a fully discrete scheme, we combine the truncated Milstein scheme with suitable spatial discretizations and noise approximations on all scales, and show that the antithetic coupling does note introduce an additional bias.

[1] M.B. Giles and L. Szpruch *Antithetic multilevel Monte Carlo estimation for multi-dimensional SDEs without Lévy area simulation* The Annals of Applied Probability, 4:24, 1585–1620, 2014.

3.3.3 Room 102, corridor 15-25 - MS On the power of iid information for (non-linear) approximation

Speaker: Art Owen

Mean dimension of radial basis functions

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

Christopher Hoyt
Stanford University, USA, crhoyt@stanford.edu

A radial basis function (RBF) of $\boldsymbol{x} \in \mathsf{R}^d$ takes the form $\phi(\|\boldsymbol{x} - \boldsymbol{c}\|)$ for a function $\phi : [0, \infty) \to \mathsf{R}$ and a point $\boldsymbol{c} \in \mathsf{R}^d$. RBFs are used to perform multidimensional interpolation and are also used in some machine learning approximations. They are used as covariance functions in Gaussian process models, and an important test function in quasi-Monte Carlo, due to Keister [2], is an RBF.
We show that generalized multiquadric radial basis functions (RBFs) (see Fasshauer [1]) on R^d have a mean dimension that is $1 + O(1/d)$ as $d \to \infty$ with an explicit bound for the implied constant, under moment conditions on their inputs. Under weaker moment conditions the mean dimension still approaches 1. As a consequence, these RBFs become more nearly additive as their dimension increases. Gaussian RBFs by contrast can attain any mean dimension between 1 and d. We also find that the Keister function has a mean dimension that oscillates between approximately 1 and approximately 2 as the nominal dimension d increases.
This work was supported by the National Science Foundation and by a grant from Hitachi, Ltd. We thank Naofumi Hama for comments on the use of RBFs in machine learning.
References
[1] G. E. Fasshauer. Meshfree approximation methods with MATLAB, volume 6. World Scientific, Singapore, 2007.
[2] B. D. Keister. Multidimensional quadrature algorithms. Computers in Physics, 10:119–122, 1996

Speaker: Robert J. Kunsch

Uniform Approximation of Finite Sequences with Randomized Algorithms

Robert J. Kunsch
RWTH Aachen University, Germany, kunsch@mathc.rwth-aachen.de

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

Marcin Wnuk
Osnabrück University, Germany, marcin.wnuk@uni-osnabrueck.de

We study the optimal approximation (with respect to the uniform norm) of finite sequences and compare deterministic and randomized algorithms. We present new upper bounds as well as lower bounds.

Speaker: Fabian Taubert

Dimension-Incremental Function Approximation Using Monte-Carlo Methods

Lutz Kämmerer
Chemnitz University of Technology, Germany, kaemmerer@math.tu-chemnitz.de

Daniel Potts
Chemnitz University of Technology, Germany, potts@math.tu-chemnitz.de

Fabian Taubert
Chemnitz University of Technology, Germany, fabian.taubert@math.tu-chemnitz.de

We present a dimension-incremental algorithm for the nonlinear approximation of high-dimensional functions in an arbitrary bounded orthonormal product basis. Our goal is to detect a suitable truncation of the basis expansion of the function, where the corresponding basis support is assumed to be unknown. Our method is based on point evaluations of the considered function, utilizes efficient algorithms for high-dimensional numerical integration and adaptively builds an index set of a suitable basis support, such that the approximately largest basis coefficients are still included. For the first time, we provide a proof of a detection guarantee for such an index set in the function approximation case under certain assumptions on the sub-methods used within our algorithm, which can be used as a foundation for similar statements in various other situations as well. Some numerical examples in different settings underline the effectiveness and accuracy of our method.

[1] L. Kämmerer, D. Potts and F. Taubert. *Nonlinear Approximation In Bounded Orthonormal Product Bases.* Submitted, arXiv:2211.06071, 2022.

3.3.4 Room 104, corridor 15-25 - Contributed Talks

Speaker: Matti Vihola

Conditional particle filters with bridge backward sampling

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

Mixing of Monte Carlo Markov chains typically deteriorates quickly in increasing dimension. In hidden Markov/state space smoothing, the dimension is proportional to the length of the data record. The conditional particle filter (CPF) with backward sampling [1] is a MCMC for this problem which is guaranteed to admit good mixing properties even with very long data records [2].
When CPF is used to sample a finely time-discretised continuous-time path integral model, its mixing properties deteriorate. The performance degradation is due to two features of the time-discretised model: weakly informative potentials and a slowly mixing dynamic model. Multinomial resampling, which is commonly employed in the (backward sampling) CPF, resamples excessively for weakly informative potentials

thereby introducing extra variance. A slowly mixing dynamic model renders the backward sampling step ineffective.
We detail two conditional resampling strategies suitable for the weakly informative regime: the so-called 'killing' resampling and the conditional version of systematic resampling with mean partial order [3]. To avoid the degeneracy issue of backward sampling, we introduce its generalisation that involves backward sampling with an auxiliary 'bridging' CPF step. Our experiments demonstrate that the CPF with a suitable resampling and the developed 'bridge backward sampling' can lead to substantial efficiency gains in the weakly informative regime.
The talk is based on joint work [4] with Sumeetpal S. Singh and Santeri Karppinen.

[1] N. Whiteley. Discussion on *"Particle Markov chain Monte Carlo methods" by C. Andrieu, A. Doucet and R. Holenstein.* Journal of the Royal Statistical Society: Series B (Statistical Methodology) 72:3, 306–307.

[2] A. Lee, S. S. Singh and M. Vihola. *Coupled conditional backward sampling particle filter.* Annals of Statistics, 48:5, 3066–3089, 2020.

[3] N. Chopin, S. S. Singh, T. Soto and M. Vihola. *On resampling schemes for particle filters with weakly informative observations.* Annals of Statistics, 50:6, 3197–3222, 2022.

[4] S. Karppinen, S. S. Singh and M. Vihola. Conditional particle filters with bridge backward sampling. Preprint arXiv:2205.13898, 2022.

Speaker: Nabil Kahalé

Unbiased time-average estimators for Markov chains

Nabil Kahalé
ESCP Business School, France, nkahale@escp.eu

We consider a time-average estimator f_k of a functional of a Markov chain. Under a coupling assumption, we show that the expectation of f_k has a limit μ as the number of time-steps goes to infinity. We describe a modification of f_k that yields an unbiased estimator $\hat{f}_k$ of μ. It is shown that $\hat{f}_k$ is square-integrable and has finite expected running time. Under certain conditions, $\hat{f}_k$ can be built without any precomputations, and is asymptotically at least as efficient as f_k, up to a multiplicative constant arbitrarily close to 1. Our approach also provides an unbiased estimator for the bias of f_k. We study applications to volatility forecasting, queues, and the simulation of high-dimensional Gaussian vectors. Our numerical experiments are consistent with our theoretical findings.

Speaker: Randolf Altmeyer

Polynomial time guarantees for sampling based posterior inference

Randolf Altmeyer
University of Cambridge, United Kingdom, ra591@cam.ac.uk

The Bayesian approach provides a flexible and popular framework for a wide range of nonparametric inference problems. It relies crucially on computing functionals with respect to the posterior distribution. Examples are the posterior mean or posterior quantiles for uncertainty quantification. In practice, this requires sampling from the posterior distribution using numerical algorithms, e.g., Markov chain Monte Carlo (MCMC) methods. The runtime of these algorithms to achieve a given target precision will typically, at least without additional structural assumptions, scale exponentially in the model dimension and the

sample size. In contrast, in this talk we show that sampling based posterior inference in a general high-dimensional setup is indeed feasible. Given a sufficiently good initialiser, we present polynomial-time convergence guarantees for a widely used gradient based MCMC sampling scheme. The proof exploits the local curvature induced by the Fisher-information of the statistical model near the underlying truth, and relies on results from the non-linear inverse problem literature. We will discuss applications to logistic and Gaussian regression, as well as to density estimation.

[1] R. Altmeyer. *Polynomial time guarantees for sampling based posterior inference in high-dimensional generalised linear models.* arXiv preprint, arXiv:2208.13296.

Speaker: Elena Sofia D'Ambrosio

A Deep Learning Method for computing summary statistics of the filtering equation in the Stochastic Reaction Networks Setting

Elena D'Ambrosio
ETH Zürich, Switzerland, elena.dambrosio@bsse.ethz.ch

Zhou Fang
ETH Zürich, Switzerland, zhou.fang@bsse.ethz.ch

Ankit Gupta
ETH Zürich, Switzerland, ankit.gupta@bsse.ethz.ch

Mustafa Khammash
ETH Zürich, Switzerland, mustafa.khammash@bsse.ethz.ch

Recent advancements in fluorescence and microscopy technologies have propelled systems and synthetic biology by enabling the investigation of biological processes at the single-cell level. However, due to the limited number of fluorescent wavelengths that can be used simultaneously, tracking more than a few molecular species at once using microscopy remains a challenge. As a result, numerical estimation of unmeasured species becomes necessary. This gives rise to a stochastic filtering problem that aims to infer the dynamics of these hidden species from the partial observations by computing a conditional distribution. This problem is especially challenging when the number of accessible states is large and the exact solution is intractable, hindering the analysis and control of single-cell dynamics.
To address this issue, we propose an innovative deep learning approach to solve the filtering equation in the setting of continuous-time noise-free observations, where the conditional distribution of hidden species (given the partial observations) is characterized by a piece-wise deterministic Markov process.[1] We reformulate the filtering equation through a martingale representation using two auxiliary stochastic processes. By exploiting the general approximation properties of deep neural networks, we use a deep neural network to approximate the equivalent formulation of the filtering equation and train it using the Monte-Carlo simulations of the two auxiliary processes as inputs. Our proposed deep learning framework provides an efficient and faster solution to the filtering equation, enabling access to its solution and statistics while avoiding the major drawbacks of particle filtering methods, which are commonly used to solve the filtering problem.

[1] D'Ambrosio, E.S., Fang, Z., Gupta, A. and Khammash, M. *Filtered finite state projection method for the analysis and estimation of stochastic biochemical reaction networks.* 2022, bioRxiv, pp.2022-10.

4 Thursday 29th June

4.1 Keynote speakers

Speaker: Elena Di Bernardino

Geometry and inference of random excursion sets

Elena Di Bernardino
Université Côte d'Azur, Laboratoire J.A. Dieudonné, UMR CNRS 7351, Nice, France,

The excursion set of a smooth random field carries relevant information in its various geometric measures. After an introduction of these geometrical quantities showing how they are related to the parameters of the field, we focus on the problem of discretization (i.e., lattices impact in the inference procedure). From a computational viewpoint, one never has access to the continuous observation of the excursion set, but rather to observations at discrete points in space. It has been reported that for specific regular lattices of points in dimensions 2 and 3, the usual estimate of the surface area of the excursions remains biased even when the lattice becomes dense in the domain of observation. We show that this limiting bias is invariant to the locations of the observation points and that it only depends on the ambient dimension. This talk is based on several joint works with H. Biermé, R. Cotsakis, C. Duval and A. Estrade.

Speaker: Jesús María Sanz Serna

Split Hamiltonian Monte Carlo revisited

Jesús María Sanz Serna
Universidad Carlos III de Madrid, Spain, jmsanzserna@gmail.com

Often, probability distributions may be seen as a perturbation of a Gaussian. When using Hamiltonian Monte Carlo (HMC) algorithms to sample from those targets, it is tempting to split the Hamiltonian H as $H_0(\theta, p) + U_1(\theta)$, where H_0 is quadratic and U_1 small and perform the required numerical integrations of the Hamiltonian dynamics by combining integrations for H_0 and integrations for U_1. This idea is appealing because, if U_1 were to vanish, the integration would be exact so that it may be hoped that for small U_1 the integration would be easy to perform. We will show that, unfortunately, samplers based on the $H_0 + U_1$ splitting suffer from stepsize stability restrictions similar to those of algorithms based on the standard leapfrog integrator. The good news is that those restrictions may be circumvented by preconditioning the dynamics. Numerical experiments show that, when the $H_0(\theta, p) + U_1(\theta)$ splitting is combined with preconditioning, it is possible to construct samplers far more efficient than standard leapfrog HMC.

4.2 Parallel sessions - Morning

4.2.1 Amphitheater 25 - MS MCM for reactor physics

Speaker: Emma Horton

A binary branching model with Moran-type interactions

Emma Horton
INRIA, France, emma.horton94@gmail.com

Alex Cox
University of Bath, UK,

Denis Villemonais
Université de Lorraine, France,

A key result for understanding the behaviour of the solution to the neutron transport equation (NTE) is the Perron-Frobenius decomposition, which allows one to characterise the large time behaviour of the solution via its leading eigenvalue and corresponding left and right eigenfunctions. Thus, obtaining accurate and efficient estimates of these quantities is imperative to nuclear safety and regulation. In this talk, I will introduce a new interacting particle model that combines the natural behaviour of the underlying fission process with a selection and resampling mechanism, allowing one to maintain control over the number of particles in the system. The main result provides an explicit relation between the particle system and the underlying neutron transport process process and also quantifies the L^2 distance between the normalised occupation measures of the two systems.

Speaker: Elöd Pázmán

Comparison of Dynamic and Static Flux Estimators and Their Impact on Group Constant Production in the Time-Dependent Monte Carlo Reactor Code GUARDYAN

Elöd Pázmán
Budapest University of Technology and Economics, Hungary,

Gábor Tolnai
Budapest University of Technology and Economics, Hungary,

Dávid Légrády
Budapest University of Technology and Economics, Hungary, legrady@reak.bme.hu

Recently, group constant production algorithms have been implemented in the Monte Carlo reactor dynamics code GUARDYAN, both for its dynamic and static calculation mode. In this work we outline the algorithms themselves, with an emphasis on the novel time boundary estimator that is specific to the time-dependent Monte Carlo method. The capabilities of the algorithms will be demonstrated, as well as compared, on a simple test geometry. As the flux estimation is a crucial part of group constant generation we will also compare the flux estimation of the algorithms separately, and whether the simulation mode (i.e. dynamic or static) has a notable effect on the results.

Speaker: Alex Cox

Twisted particle filters for neutron transport

Andreas Kyprianou
Warwick University, UK, a.e.kyprianou@gmail.com

Tom Davis
University of Bath, UK,

Alex Cox
University of Bath, UK,

We give a brief overview of some ideas concerning Doob h-transforms and Monte Carlo simulation for criticality problems associated to the neutron transport equation.

Speaker: Andrea Zoia

Analysis of heterogeneous Markov media for particle transport problems

Mikolaj Adam Kowalski
CEA, France,

Coline Larmier
CEA, France,

Andrea Zoia
CEA, France, andrea.zoia@cea.fr

Markov media provide a prototype class of stochastic geometries that are widely used in order to model several complex and disordered systems encompassing e.g. turbulent fluids and plasma, atmospheric layers, or biological tissues, especially in relation to particle transport problems. In several key applications, the statistical properties of random media may display spatial gradients due to material stratification, which means that the typical spatial scale and the probability of finding a given material phase at a spatial location become non-homogeneous. In this work we investigate the main features of spatially heterogeneous Markov media, using Poisson hyperplane tessellations and Arak polygonal fields. We show that both models can generate geometry realisations sharing Markov-like properties, and discuss their distinct advantages and drawbacks in terms of flexibility and ease of use. The impact of these models on the observables related to particle transport will be assessed using Monte Carlo simulations.

4.2.2 Room 101, corridor 24-25 - MS Nonreversible processes theory and applications

Speaker: Giovanni Conforti

A probabilistic view of Sinkhorn's algorithm

Giovanni Conforti
Ecole Polytechnique, France, giovanni.conforti@polytechnique.edu

Alain Durmus
Ecole Polytechnique, France, Alain.Durmus@polytechnique.edu

Giacomo Greco
Eindhoven University of Technology, Netherlands, g.greco@tue.nl

Maxence Noble-Bourillot
Ecole Polytechnique, France, Maxence.Noble-Bourillot@polytechnique.edu

The entropic optimal transport problem (EOT) is obtained by adding an entropic regularisation term in the cost function of the Monge Kantorovich problem. Over the past decade, EOT has found a number of successful applications in statistical machine learning and data science: this is essentially due do the fact that its solutions can be computed by means of Sinkhorn's algorithm. In this talk, we shall propose a new viewpoint on the exponential convergence of Sinkhorn's algorithm that draws inspiration from the theory of Markov processes. In particular, we show how the ergodicity of the reference coupling is related to the exponential rate of convergence.

Speaker: Arnaud Guillin

Pair of run and tumble processes with interactions: invariant measure, jamming and mixing time

Arnaud Guillin
LMBP at University Clermont, France, arnaud.guillin@uca.fr

Léo Hahn
Ecole Normale Supérieure, Frnce, leo.hahn@ens.fr

Manon Michel
LMBP at University Clermont, France, manon.michel@uca.fr

We consider here run and tumble particles, i.e. piece-wise deterministic Markov process, which, when they collide, are stucked in collision until a change of direction sets them free. We will construct this processes and study the long time behavior, as well as exhibit their invariant measure showing jamming at equilibrium. It can be seen as a toy problem to model Motility Induced Phase Separation.

Speaker: Carsten Hartmann

Constraints, strong confinement limits and conditional expecations in nonreversible Langevin dynamics

Carsten Hartmann
Freie University of Berlin, Germany, chartman@mi.fu-berlin.de

Stochastic differential equations that appear in the modelling of real-world systems or in Markov chain Monte Carlo are often endowed with algebraic constraints (e.g. in sampling of Riemannian manifolds). We discuss the realisation of constraints for nonreversible systems by either strong confining forces or projection methods and illustrate the essential difficulties in extending concepts that are well established and understood for reversible dynamics to the nonreversible situation. We discuss invariant measures of constrained systems and outline potential issues that are related to the loss of controllability or ergodicity when constraining nonreversible stochastic dynamics.

Speaker: Daniel Adams

Non-Reversible processes and the relation between GENERIC and Hypocoercivity

Daniel Adams
Heriot-Watt University, UK, da2050@hw.ac.uk

Non-reversible processes are universal in non-equilibrium statistical mechanics, and have recently been adopted by the statistical sampling community to accelerate Markov Chain Monte Carlo algorithms. The framework to analyse non-reversible processes is far less developed than their reversible counterparts, in particular, they do not possess a variational gradient flow formulation. We will explore the decomposition of such processes into their conservative and dissipative parts. Specifically, we will see how the theories of GENERIC and Hypocoercivity (which provide such a decompositon) are related.

[1] D.Adams, M.H.Duong and G.D.Reis. *Operator-splitting schemes for degenerate, non-local, conservative-dissipative systems.* Discrete and Continuous Dynamical systems, 42, 5453–5486, 2022.

[2] M.H.Duong and M.Ottobre. *Non-reversible processes: GENERIC, hypocoercivity and fluctuations.* Nonlinearity, 3, 1617, 2023.

4.2.3 Room 102, corridor 15-25 - MS Sampling Strategies for Bayesian Inference

Speaker: Konstantinos Zygalakis

Bayesian Inference with Data-Driven Image Priors Encoded by Neural Networks

Konstantinos Zygalakis
University of Edinburgh, UK, k.zygalakis@ed.ac.uk

This talk presents a new methodology for performing Bayesian inference in imaging inverse problems where the prior knowledge is available in the form of training data. Following the manifold hypothesis and adopting a generative modelling approach, we construct a data-driven prior that is supported on a sub-manifold of the ambient space, which we learn from training data by using a variational autoencoder or a generative adversarial network. We establish the existence and well-posedness of the associated posterior distribution and posterior moments under easily verifiable conditions, providing a rigorous underpinning for Bayesian estimators and uncertainty quantification analyses. Bayesian computation is performed by using specialised Monte Carlo algorithms. In addition to point estimators and uncertainty quantification analyses, we derive a model misspecification test to automatically detect situations where the data-driven prior is unreliable, and explain how to identify the dimension of the latent space directly from the training data. The proposed approach is illustrated with a range of experiments with the MNIST dataset, where it outperforms alternative image reconstruction approaches from the state of the art. A model accuracy analysis suggests that the Bayesian probabilities reported by the data-driven models are also remarkably accurate under a frequentist definition of probability.

Speaker: Luke Shaw

Rotation-based integrators for HMC

Fernando Casas
Universitat Jaume I, Spain, casas@uji.es

Jesús María Sanz Serna
Universidad Carlos III, Spain, jmsanzserna@gmail.com

Luke Shaw
Universitat Jaume I, Spain, shaw@uji.es

Hybrid Monte Carlo (HMC) is a Markov Chain Monte Carlo (MCMC) technique which numerically integrates Hamiltonian dynamics with Hamiltonian $H = T(p) + U(q)$ to generate proposal samples from a desired probability distribution, $\exp(-U(q))$. $T(p)$ may be freely chosen. The all-purpose standard for the numerical integration remains the Verlet integrator, which composes the exact flows of the sub-Hamiltonians $T(p)$ and $U(q)$. However in many cases, distributions have a Radon-Nikodym density with respect to a Gaussian distribution, or may be numerically approximated by a Gaussian multiplied by a residual perturbation. In either case, the corresponding Hamiltonian then takes the form of an additive perturbation to a harmonic oscillator. We study integrators which compose the exact flows of $H_0 = (1/2)(p^T p + q^T q)$ (a 'rotation') and $U_1(q)$ (a 'kick') for the Hamiltonian $H(q,p) = H_0(q,p) + U_1(q)$, and particularly emphasise the importance of stability for performance in HMC sampling, where moderate to large timesteps are used. The importance of stability and the analysis of a model problem leads us to propose and demonstrate the superiority of a pair of preconditioned modifications of rotation-based integrators in a series of logistic regression problems using simulated and real data [1].

[1] F. Casas, J. M. Sanz-Serna and L. Shaw *Split Hamiltonian Monte Carlo Revisited* Statistics and Computing, 32:5, 1-14, 2022.

Speaker: Nicola Branchini

Generalized Self-Normalized Importance Sampling

Nicola Branchini
University of Edinburgh, UK, nicola.branchini@ed.ac.uk

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

Importance sampling (IS) is an efficient Monte Carlo technique for the approximation of expectations, i.e., integrals involving a probability density function (pdf). The pdf is often known only up to a normalizing constant, which turns the problem even more complicated. In these cases, the *self-normalized* IS (SNIS) estimator is a standard Monte Carlo approach. However, this widely used estimator still keeps fundamental limitations, partly because it has been underlooked in the literature. For instance, its variance is lower bounded by a nonzero constant. Moreover, the SNIS estimator can be interpreted as the ratio of two estimators to approximate both the integral with the unnormalized distribution (numerator) and the missing normalizing constant (denominator). Under this perspective, both estimators use the same set of samples. Thus, the (involuntarily) induced correlation between both estimators can damage the overall performance. In this talk, we give an overview of the SNIS and provide new insights of this popular estimator which is yet not well understood. Then, we take a new approach by duplicating the integration space. We introduce a joint proposal distribution in this extended space, whose two marginals are the proposal distributions used to approximate the integrals of the numerator and denominator, respectively. Moreover, by controlling the dependence of the two estimators, the MSE can be reduced. Our framework generalizes the SNIS estimators and previous works.

Speaker: Lorenzo Nagar
Speaker: Martín Parga Pazos

Adaptive integration approach for sampling with Hamiltonian Monte Carlo based methods

Lorenzo Nagar
Basque Center for Applied Mathematics (BCAM), lnagar@bcamath.org

Martín Parga Pazos
Basque Center for Applied Mathematics (BCAM) and CIC bioGUNE, mparga@bcamath.org

Hamiltonian Monte Carlo (HMC)-based methods have been widely recognized as a powerful sampling tool for Bayesian inference [1, 2, 3, 4]. Key performance factors of HMC are accurate numerical integration of underlying Hamiltonian equations and an appropriate choice of simulation parameters and settings. When applied to physical problems, natural constraints may offer a hint on such choices, which hardly can be translated into Bayesian inference applications. We present a novel adaptive integration approach (we call it s-AIA) that detects a system-specific multistage splitting integrator with a complete set of well-founded integrator-specific parameters to achieve a competitive sampling efficiency in HMC or a Generalized HMC-based Bayesian inference applications. The method relies on analysis of expected

energy errors of multivariate Gaussian models combined with the data generated at the burn-in stage of a (G)HMC simulation. It automatically eliminates those values of simulation parameters, which may cause undesired extreme scenarios, such as resonance artefacts, low accuracy or poor sampling. We use the ideas of s-AIA to develop its advanced variant, s-MAIA, for importance sampling HMC. Testing s-(M)AIA on representative models demonstrates its superiority over popular symplectic integrators, e.g. Verlet, BCSS, ME. An illustrative example of s-(M)AIA application to the study on uncovering novel biomarkers linked to endocrine therapy resistance in breast cancer is provided.

[1] Neal, R. M. *MCMC using Hamiltonian dynamics* Handbook of Markov Chain Monte Carlo, Vol. 2, pp. 113–162, Chapman & Hall / CRC Press, 2011.

[2] Radivojević, T. and Akhmatskaya, E. *Modified Hamiltonian Monte Carlo for Bayesian inference* Statistics and Computing 30, pp. 377–404, 2020.

[3] Akhmatskaya, E., Fernández-Pendás, M., Radivojević, T. and Sanz-Serna, J. M. *Adaptive splitting integrators for enhancing sampling efficiency of modified Hamiltonian Monte Carlo methods in molecular simulation* Langmuir 33.42, pp. 11530–11542, 2017

[4] Akhmatskaya, E. and Reich, S. *New Hybrid Monte Carlo Methods for Efficient Sampling: from Physics to Biology and Statistics* Progress in Nuclear Science and Technology, Vol. 2, pp. 447-462, 2011.

4.2.4 Room 104, corridor 15-25 - Contributed Talks

Speaker: Paul Dobson

Accelerating MCMC using interacting Langevin models

Paul Dobson
University of Edinburgh, Edinburgh, pdobson@ed.ac.uk

Marcelo Pereyra
Heriot-Watt University, Edinburgh, m.pereyra@hw.ac.uk

Konstantinos Zygalakis
University of Edinburgh, Edinburgh, K.Zygalakis@ed.ac.uk

In this talk I will discuss using a hierarchy of Bayesian models given by using simpler priors in order to accelerate the convergence of Sampling algorithms. There has been a lot of success with using neural networks to provide data driven priors for imaging problems. These priors give a more accurate reconstruction but introduce additional challenges in constructing an efficient sampling algorithm. In particular, such models are typically not log concave so algorithms tend to spend a long time in local minima. We extend the model by including a simple log concave model and setting up two Stochastic Differential Equations which have as invariant measures the two posterior distributions and introduce interactions between these two processes to accelerate convergence while preserving the invariant measure.

Speaker: Andreas Eberle

Asymptotic bias of inexact Markov Chain Monte Carlo methods in high dimension

Alain Durmus
CMAP, CNRS, École Polytechnique, Institut Polytechnique de Paris, France,
alain.durmus@polytechnique.edu

Andreas Eberle
University of Bonn, Germany, eberle@uni-bonn.de

Inexact Markov Chain Monte Carlo methods rely on Markov chains that do not exactly preserve the target distribution. Examples include the unadjusted Langevin algorithm (ULA) and unadjusted Hamiltonian Monte Carlo (uHMC). This paper establishes bounds on Wasserstein distances between the invariant probability measures of inexact MCMC methods and their target distributions with a focus on understanding the precise dependence of this asymptotic bias on both dimension and discretization step size. Assuming Wasserstein bounds on the convergence to equilibrium of either the exact or the approximate dynamics, we show that for both ULA and uHMC, the asymptotic bias depends on key quantities related to the target distribution or the stationary probability measure of the scheme. As a corollary, we conclude that for models with a limited amount of interactions such as mean-field models, finite range graphical models, and perturbations thereof, the asymptotic bias has a similar dependence on the step size and the dimension as for product measures.

Speaker: Irene Tubikanec

Network inference in a stochastic multi-population neural mass model via SMC-ABC

Susanne Ditlevsen
University of Copenhagen, susanne@math.ku.dk

Massimiliano Tamborrino
University of Warwick, massimiliano.tamborrino@warwick.ac.uk

Irene Tubikanec
University of Klagenfurt, irene.tubikanec@aau.at

In this talk, we consider a $6N$-dimensional stochastic differential equation, modelling the activity of N coupled populations of neurons in the brain. This neural mass model dates back to the work of Jansen and Rit (1995) and has been introduced to describe human electroencephalography (EEG) rhythms, in particular epileptic signals. Our contributions are threefold: First, we construct an explicit and efficient numerical simulation method for this stochastic N-population model, extending the splitting procedure presented in [2] for the case of one neural population (i.e., $N = 1$). Second, we present a sequential Monte Carlo approximate Bayesian computation (SMC-ABC) method to infer both the continuous model parameters and discrete coupling direction (network) parameters of the equation. This algorithm extends the spectral density based ABC summary statistics proposed in [3] to the case of multiple interacting components and relies on a combination of discrete and continuous perturbation kernels. Third, we apply the derived algorithm on real multi-channel EEG data, aiming to infer the brain's connectivity structure during epileptic seizure. The proposed simulation-based inference method may be used for network estimation also in other types of multi-dimensional coupled stochastic differential equations.

[1] B. H. Jansen and V. G. Rit. *Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns.* Biological Cybernetics, 73:4, 357–366, 1995.

[2] M. Ableidinger, E. Buckwar, and H. Hinterleitner. *A stochastic version of the Jansen and Rit neural mass model: analysis and numerics.* The Journal of Mathematical Neuroscience, 7:8, 2017.

[3] E. Buckwar, M. Tamborrino, and I. Tubikanec. *Spectral density-based and measure-preserving ABC for partially observed diffusion processes. An illustration on Hamiltonian SDEs.* Statistics and Computing, 30, 627–648, 2020.

Speaker: Sara Pérez-Vieites

Adaptive Gaussian nested filter for joint parameter and state estimation in state-space models

Sara Pérez-Vieites
IMT Nord Europe, Institut Mines-Télécom, Centre for Digital Systems, France,
sara.perez-vieites@imt-nord-europe.fr

Víctor Elvira
School of Mathematics, University of Edinburgh, UK, victor.elvira@ed.ac.uk

In this talk, we introduce the adaptive Gaussian nested filter (AGnesF), the first nested method that adapts the number of samples to estimate both the static parameters and the dynamical state variables of a state-space model. The proposed method is based on the nested Gaussian filter (NGF) [1], a probabilistic technique that combines two layers of inference, one inside the other, to compute the joint posterior probability distribution of the static parameters and the state variables given a set of sequentially-collected observations. We propose two novel rules to reduce computational complexity without compromising the performance. The first rule enables the bottom layer techniques to run recursively, thus, the algorithm does not run from scratch every time there is a new observation. The second rule reduces automatically the number of samples in the parameter space. This reduction is based on a statistic that identifies when the parameter samples are redundant. We describe a specific implementation of the new scheme that uses a quadrature Kalman filter (QKF) in the parameter layer, and we study its performance in a stochastic Lorenz 63 model. To appear in [2].

[1] S. Pérez-Vieites and J. Míguez. *Nested Gaussian filters for recursive Bayesian inference and nonlinear tracking in state space models.* Signal Processing, 189, 108295, 2021.

[2] S. Pérez-Vieites and V. Elvira. *Adaptive Gaussian nested filter for parameter estimation and state tracking in dynamical systems.* In 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2023) [Accepted].

4.3 Parallel sessions - Afternoon

4.3.1 Amphitheater 25 - MS Stochastic Computation and Complexity

Speaker: Michaela Szölgyenyi

A higher order approximation method for jump-diffusion SDEs with discontinuous drift coefficient

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

Verena Schwarz
University of Klagenfurt, Austria, verena.schwarz@aau.at

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

We present the first higher-order approximation scheme for solutions of jump-diffusion stochastic differential equations with discontinuous drift. For this transformation-based jump-adapted quasi-Milstein scheme we prove L^p-convergence order 3/4. To obtain this result, we prove that under slightly stronger assumptions (but still weaker than anything known before) a related jump-adapted quasi-Milstein scheme has convergence order 3/4 – in a special case even order 1. Order 3/4 is conjectured to be optimal.

Speaker: Tim Johnston

Nonasymptotic Bounds for EM via Interacting Particle Systems

Tim Johnston
The University of Edinburgh, UK, T.Johnston-4@sms.ed.ac.uk

In this talk we discuss a new interacting particle system used for implementing an expectation maximization (EM) procedure (or more generally, to optimize over the parameters of a latent variable model). This continuous-time interacting particle system has the property that it can be seen as a Langevin diffusion over an extended state space, where the number of particles acts almost like an inverse parameter in classical settings for optimization. This then allow for use of the well-developed theory Langevin diffusions to prove nonasymptotic concentration bounds for the optimization error of the maximum marginal likelihood estimator. In this talk we shall place this new algorithm in the context of existing approaches, and also discuss the structure of our proof and how it naturally lends itself to generalisation.

Speaker: Gunther Leobacher

McKean–Vlasov equations with discontinuous drift

Gunther Leobacher
University of Graz, Austria, gunther.leobacher@ui-graz.at

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

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

In the paper, we first establish well-posedness results for one-dimensional McKean–Vlasov stochastic differential equations (SDEs) and related particle systems with a measure-dependent drift coefficient that is discontinuous in the spatial component, and a diffusion coefficient which is a Lipschitz function of the state only. We only require a fairly mild condition on the diffusion coefficient, namely to be non-zero in a point of discontinuity of the drift, while we need to impose certain structural assumptions on the measure-dependence of the drift. Second, we study Euler–Maruyama type schemes for the particle system to approximate the solution of the one-dimensional McKean–Vlasov SDE. Here, we will prove strong convergence results in terms of the number of time-steps and number of particles. Due to the discontinuity of the drift, the convergence analysis is non-standard and the usual strong convergence order 1/2 known for the Lipschitz case cannot be recovered for all presented schemes.

[1] G. Leobacher and C. Reisinger and W. Stockinger *Well-posedness and numerical schemes for one-dimensional McKean–Vlasov equations and interacting particle systems with discontinuous drift.* BIT Numerical Mathematics 62, pages 1505–1549 (2022)

Speaker: Oleg Butkovsky

Strong rate of convergence of the Euler scheme for SDEs with irregular drift driven by Levy noise

Oleg Butkovsky
Weierstrass Institute, Berlin, Germany, butkovsky@wias-berlin.de

We study the strong rate of convergence of the Euler–Maruyama scheme for a multidimensional stochastic differential equation (SDE)

$$dX_t = b(X_t)dt + dL_t,$$

with irregular β-Hölder drift, $\beta > 0$, driven by a Lévy process with exponent $\alpha \in (0, 2]$. For $\alpha \in [2/3, 2]$ we obtain strong L_p and almost sure convergence rates in the whole range $\beta > 1 - \alpha/2$, where the SDE is known to be strongly well–posed. This significantly improves the current state of the art both in terms of convergence rate and the range of α. In particular, the obtained convergence rate does not deteriorate for large p and is always at least $n^{-1/2}$; this allowed us to show for the first time that the the Euler–Maruyama scheme for such SDEs converges almost surely and obtain explicit convergence rate. Furthermore, our results are new even in the case of smooth drifts. Our technique is based on a new extension of the stochastic sewing arguments.

[1] O. Butkovsky, K. Dareiotis, and M. Gerencsér. *Strong rate of convergence of the Euler scheme for SDEs with irregular drift driven by Levy noise.* arXiv:2204.12926 (2022), to appear in Annales de l'Institut Henri Poincaré.

4.3.2 Room 101, corridor 24-25 - MS Convergence results for kinetic samplers

Speaker: Lionel Riou-Durand

Metropolis Adjusted Langevin Trajectories: a robust alternative to Hamiltonian Monte Carlo

Lionel Riou-Durand
University of Warwick, UK, lionel.riou-durand@warwick.ac.uk

Hamiltonian Monte Carlo (HMC) is a widely used sampler, known for its efficiency on high dimensional distributions. Yet HMC remains quite sensitive to the choice of integration time. Randomizing the length of Hamiltonian trajectories (RHMC) has been suggested to smooth the Auto-Correlation Functions (ACF), ensuring robustness of tuning. We present the Langevin diffusion as an alternative to control these ACFs by inducing randomness in Hamiltonian trajectories through a continuous refreshment of the velocities. We connect and compare the two processes in terms of quantitative mixing rates for the 2-Wasserstein and L2 distances. The Langevin diffusion is presented as a limit of RHMC achieving the fastest mixing rate for strongly log-concave targets. We introduce a robust alternative to HMC built upon these dynamics, named Metropolis Adjusted Langevin Trajectories (MALT). Studying the scaling limit of MALT, we obtain optimal tuning guidelines similar to HMC, and recover the same scaling with respect to the dimension without additional assumptions. We illustrate numerically the efficiency of MALT compared to HMC and RHMC.

[1] Lionel Riou-Durand and Jure Vogrinc. Metropolis Adjusted Langevin Trajectories: a robust alternative to Hamiltonian Monte Carlo. *arXiv:2202.13230*, 2022.

Speaker: Katharina Schuh

Convergence of unadjusted Hamiltonian Monte Carlo and Langevin dynamics via couplings

Katharina Schuh
Technical University of Vienna, Austria, katharina.schuh@asc.tuwien.ac.at

We analyse the long-time behaviour of both the unadjusted Hamiltonian Monte Carlo method applied to mean-field particle models and the nonlinear Langevin dynamics of McKean- Vlasov type. We establish contraction bounds in L^1 Wasserstein distance with dimension-free rates. In both cases, the results are not restricted to strongly convex confining potentials. They rather include multi-well potentials. The proofs are based on coupling approaches and on the construction of appropriate distance functions.

[1] K. Schuh *Global contractivity for Langevin dynamics with distribution-dependent forces and uniform in time propagation of chaos.* ArXiv Preprint arXiv:2206.03082, 2022.

[2] N. Bou-Rabee, K. Schuh *Convergence of unadjusted Hamiltonian Monte Carlo for mean-field models.* ArXiv Preprint arXiv:2009.08735, 2020.

Speaker: Lucas Journel

Sampling of singular Gibbs measure

Lucas Journel
Laboratoire Jacques-Louis Lions, Paris, France, lucas.journel@sorbonne-universite.fr

We are interested in sampling Gibbs measures associated to singular potentials using the kinetic Langevin diffusion. We show a 'Talay-Tubaro' expansion of the invariant measure of numerical schemes based on this process, under conditions which hold for instance for a system of N particles in a confining potential with Lennard-Jones pairwise interactions. A key point of the proof is to establish in this context hypocoercive long-time convergence estimates for the continuous-time process in terms of high order Sobolev norms.

[1] Lucas Journel and Pierre Monmarché. Convergence of the kinetic annealing for general potentials. *Electronic Journal of Probability*, 27(none):1 – 37, 2022.

Speaker: Lihan Wang

Convergence Rates of Kinetic Sampling Dynamics via Space-time Poincaré-type Inequality

Yu Cao
New York University, USA, yucaoyc@outlook.com

Jianfeng Lu
Duke University, USA, jianfeng@math.duke.edu

Lihan Wang
Max Planck Institute for Mathematics in the Natural Sciences, lihan.wang@mis.mpg.de

We will discuss some results on the analysis of convergence rates of some hypocoercive kinetic sampling dynamics, including underdamped Langevin dynamics, randomized Hamiltonian Monte Carlo, zigzag process and bouncy particle sampler. The analysis is based on the Armstrong-Mourrat variational framework for hypocoercivity which combines a Poincaré-type inequality in time-augmented state space and an L^2 energy estimate.

4.3.3 Room 102, corridor 15-25 - MS Sampling Schemes - Quality Measures, Point Generation, and Applications

Speaker: François Clément

Subset Sampling for Low Discrepancy Point Sets

François Clément
Sorbonne Université, CNRS, LIP6, France, francois.clement@lip6.fr

Carola Doerr
Sorbonne Université, CNRS, LIP6, France, carola.doerr@lip6.fr

Luís Paquete
University of Coimbra, CISUC, Department of Informatics Engineering, Portugal, paquete@dei.uc.pt

Low-discrepancy sequences such as Halton or Sobol' were designed to have small discrepancy values asymptotically, when the number of points tends to infinity. However, practical applications will often have a finite number of points, too small to reach the asymptotic setting. The lack of low-discrepancy constructions tailored to specific point set size and dimension combinations led us to introduce the Star Discrepancy Subset Selection Problem (SDSSP) in [1]. The SDSSP consists in choosing from a set P of n points a subset P_k of size $k \leq n$ such that the L_∞ star discrepancy of P_k is minimized.
For dimensions two and three, we provided two exact algorithms to solve the SDSSP. In this talk, we provide heuristics to solve the SDSSP in higher dimensions with a greater number of points. This new approach allows us to provide point sets of much lower discrepancy for all dimensions for which the discrepancy can be reliably computed. For example, in dimension 6 with between 100 and 200 points, the discrepancies obtained are between 20% and 50% smaller than that of the initial Sobol' sequence. These improvements naturally lead to better inverse star discrepancies, the smallest number of points needed to obtain a given discrepancy value. We also provide a comparison of the point sets obtained with the SDSSP with those obtained by an energy functional introduced by Steinerberger [2] to build low-discrepancy points sets.

[1] F. Clément and C. Doerr and L. Paquete. *Star discrepancy subset selection: Problem formulation and efficient approaches for low dimensions*, Journal of Complexity, 70, 101645, 2022, https://doi.org/10.1016/j.jco.202

[2] S. Steinerberger. *A non-local functional promoting low-discrepancy point sets*, Journal of Complexity, 54, 101410, 2019.

Speaker: Jasmin Fiedler

Maximal inequalities in Discrepancy theory

Jasmin Fiedler
Hochschule Ruhr West, Germany, jasmin.fiedler@hs-ruhrwest.de

Discrepancy theory is one of the cornerstones of Quasi-Monte Carlo methods. In recent years, stochastic methods have gained importance and have been used to prove the existence of sequences whose discrepancy fulfills certain desired criteria. This talk aims at providing an overview over recent developments in this area, specifically on the usage of stochastic inequalities such as Bernstein-type inequalities, Hoeffding-type inequalities and their maximal versions as well as Lévy-Ottaviany-type inequalities. Many of the known results are tailored to the iid case and we discuss challenges when trying to generalize these results to a more general setting. Finally, we show possible applications in discrepancy theory.

Speaker: Nathan Kirk

Stratified Sampling of the Unit Cube

Nathan Kirk
Queen's University Belfast, nkirk09@qub.ac.uk

Randomised quasi-Monte Carlo (RQMC) sampling is a popular method for constructing point sets which have improved uniform distribution properties when compared to Monte Carlo samples, while still possessing the advantages of being 'random' in theoretical analysis. Classical jittered sampling is an example of RQMC and the starting point for our exploration into more general stratified sampling, in which we partition the d-dimensional unit cube into N arbitrary sets and a stratified point set is generated by choosing one point uniformly (and stochastically independent of the other points) from each set.
In this talk, we begin by presenting results from [2] where we show closed formulae for several discrepancy measures of jittered sampling. Our discussion then focuses on a particular family of equivolume partitions, first introduced in [1] in dimension 2 and generalised to arbitrary dimension in [3], which yield an N-element stratified point set which has better distribution qualities than a Monte Carlo sample. This family of partitions also has the benefit of being constructed for arbitrary N unlike classical jittered sampling.
Lastly, supported by numerical results we reveal how the family of equivolume partitions from [1] can be altered to give an optimal configuration with respect to the discrepancy – perhaps giving some insight into how partitioning sets should be positioned to yield the optimal partition of the unit cube with respect to the discrepancy.

[1] Kiderlen, M., Pausinger, F.: *Discrepancy of stratified samples from partitions of the unit cube.* Monatsh. Math. 195, 267–306, 2022.

[2] Kirk, N., Pausinger, F.: *On the expected discrepancy of jittered sampling* arXiv: 2208.08924

[3] Clement, F., Kirk, N., Pausinger, F.: On the construction of equivolume partitions of the d-dimensional unit cube arXiv: 2204.09340

Speaker: Zexin Pan

Super-polynomial Accuracy of Median-of-means

Zexin Pan
Department of Statistics, Stanford University, USA, zep002@stanford.edu

Art Owen
Department of Statistics, Stanford University, USA, owen@stanford.edu

Digital net is an important class of Quasi-Monte Carlo methods used for multidimensional integration. In the talk, I am going to show digital net randomized by linear scrambling and digital shift exhibits surprising concentration behavior. Taking the median of several digital net averages can exclude outliers and boost the convergence rate from nearly cubic to super-polynomial when the integrand is smooth. I will focus on the intuition behind the super-polynomial asymptotic rate and describe an empirical rule about how many samplings are needed for such fast rate to occur.

[1] Z. Pan and A. B. Owen. *Super-polynomial accuracy of multidimensional randomized nets using the median-of-means.* arXiv preprint arXiv:2208.05078, 2022.

4.3.4 Room 104, corridor 15-25 - MS MS (Quasi-)MC Software

Speaker: Aleksei Sorokin

Collaborative Integrations with the QMCPy Framework

Aleksei Sorokin
Illinois Institute of Technology, USA, asorokin@hawk.iit.edu

Monte Carlo methods are a staple of scientific computing. Researchers have developed a number of specialized Monte Carlo techniques for increased efficiency and practicality. We have extracted some common components of such methods into abstract classes in the QMCPy Python library. This talk highlights subclasses which integrate QMCPy with existing software libraries from across the scientific computing community. Examples include sequence generators for Quasi-Monte Carlo methods, variable transformations for uncommon distributions, adaptive stopping criteria for infinite dimensional problems, and complex function evaluations in containerized environments.

Speaker: Anne Reinarz

UM-Bridge

Linus Seelinger
Heidelberg University, Germany, linus.seelinger@iwr.uni-heidelberg.de

Anne Reinarz
Durham University, UK, anne.k.reinarz@durham.ac.uk

In this talk I will present UM-Bridge (the UQ and Model Bridge), an open source project that provides a unified interface for numerical models that is accessible from virtually any programming language or framework. It is primarily intended for coupling advanced models (e.g. simulations of complex physical processes) to advanced statistical or optimization methods. Containarization of models using UM-Bridge is trivial, allowing for unified, portable, fully reproducible and black-box models.
As an example I will present a parallelized multilevel Markov chain Monte Carlo (MLMCMC) method, an algorithmically scalable UQ algorithm for Bayesian inverse problems. The integration between UQ and modeling software is facilitated by UM-Bridge and allows for large-scale parallelism across forward model evaluations and the UQ algorithms themselves. The main scalability challenge presents itself in the form of strong data dependencies introduced by the MLMCMC method, prohibiting trivial parallelization. I will demonstrate the method by using it to infer the most probable locations for the initialisation of a tsunami from buoy data.

[1] N. Metropolis and S. Ulam. *The Monte Carlo Method.* Journal of the American Statistical Association, 44:247, 335–341, 1949.

Speaker: Mikkel Bue Lykkegaard

TinyDA: Multilevel Delayed Acceptance MCMC for Human Beings

Mikkel Bue Lykkegaard
digiLab, Exeter, United Kingdom, mikkel@digilab.co.uk

In this talk, I will present `tinyDA`, an open-source `Python` package for MCMC sampling, which is particularly well-suited for analysing large-scale Bayesian inverse problems.
`tinyDA` implements the Multilevel Delayed Acceptance (MLDA) MCMC algorithm as described in Lykkegaard, et al. (2022). It includes a selection of gradient-free MCMC proposals, including Adaptive Metropolis, preconditioned Crank-Nicolson and Differential Evolution Adaptive Metropolis (DREAM), as well as a Multilevel Adaptive Error Model. Additionally, `tinyDA` integrates seamlessly with `ArviZ` for posterior inference, analysis, and diagnostics. It is modular, extensible, and easy to use.
I will give a brief overview of the core MLDA algorithm and its properties, a walk-through of setting up `tinyDA` for an arbitrary Bayesian inverse problem and provide a few examples of practical applications.

[1] M. B. Lykkegaard, T. J. Dodwell, C. Fox, G. Mingas and R. Scheichl *Multilevel Delayed Acceptance MCMC.* Preprint, https://arxiv.org/abs/2202.03876, 2022.

Speaker: Pieterjan Robbe

Multilevel Delayed Acceptance MCMC for the prediction of xenon diffusion in UO2 nuclear fuel

Pieterjan Robbe
Sandia National Laboratories, `ppmrobbe@sandia.gov`

Predicting the diffusivity of xenon diffusion in UO2 is crucial for assessing the performance of nuclear fuels. We apply the Multilevel Delayed Acceptance (MLDA) method to an atomistic model used to predict xenon diffusivity in order to estimate the model parameters, given data reported as summary statistics. MLDA uses a coarse model approximation to efficiently generate proposals that will be accepted with high probability in an MCMC method. In our setting, this coarse model approximation is a higher-order polynomial chaos surrogate model. We discuss challenges in implementation, and formulate a strategy to generate multiple samples simultaneously on a parallel machine.

5 Friday 30th June

5.1 Keynote speakers

Speaker: Michael Gnewuch

High- and Infinite-Dimensional Integration and Function Recovery: Randomized Algorithms and Their Analysis

Michael Gnewuch
University of Osnabrück, Germany, michael.gnewuch@uni-osnabrueck.de

Consider an integration or a function recovery problem where the input functions depend on a very large or even infinite number of variables and belong to some reproducing kernel Hilbert space. The following questions arise naturally:

1. What kind of structure of the functions or of the function space, respectively, avoids the curse of dimensionality and ensures that the problem is computationally tractable?
2. What kind of algorithms give us almost optimal convergence rates?
3. How does their analysis depend on the specific function space and its norm?

In this talk we want to discuss these questions in the setting where we are allowed to use randomized algorithms, where the error criterion is given by the randomized worst-case error (i.e., the worst-case root mean square error over the norm unit ball of the function space), and where the cost of evaluating a function in some point x depends in a reasonable way on the number of "active variables" of x. In the course of the talk, we will see that the questions above may be answered in the following way:
Helpful concepts to lift the curse of dimensionality are tensor product spaces, weights that moderate the importance of different groups of variables or increasing smoothness of the input functions (where "increasing" is meant with respect to the ordered variables).
Depending on the chosen cost model, multilevel algorithms or multivariate decomposition methods, based on good building block algorithms that take care of lower dimensional sub-problems, may achieve convergence rates arbitrarily close to the optimal order.
Furthermore, we want to present an elaborate framework for the embedding of different (scales of) Hilbert spaces, which enables us to transfer tractability results from specific Hilbert spaces to larger classes of spaces.

Speaker: Tamara Broderick

The Bayesian Infinitesimal Jackknife for Variance

Tamara Broderick
Massachusetts Institute of Technology, Boston, USA, tamarab@mit.edu

The frequentist variability of Bayesian posterior expectations can provide meaningful measures of uncertainty even when models are misspecified. Classical methods to asymptotically approximate the frequentist covariance of Bayesian estimators such as the Laplace approximation and the nonparametric bootstrap can be practically inconvenient, since the Laplace approximation may require an intractable integral to compute the marginal log posterior, and the bootstrap requires computing the posterior for many different bootstrap datasets. We develop and explore the infinitesimal jackknife (IJ), an alternative method for computing asymptotic frequentist covariance of smooth functionals of exchangeable data, which is based on the "influence function" of robust statistics. We show that the influence function for

posterior expectations has the form of a posterior covariance, and that the IJ covariance estimate is, in turn, straightforward to compute from a single set of posterior samples. Under conditions similar to those required for a Bayesian central limit theorem to apply, we prove that the corresponding IJ covariance estimate is asymptotically equivalent to the Laplace approximation and the bootstrap. In the presence of nuisance parameters that may not obey a central limit theorem, we argue heuristically that the IJ covariance can remain a good approximation to the limiting frequentist variance. We demonstrate the accuracy and computational benefits of the IJ covariance estimates with simulated and real-world experiments.

5.2 Parallel sessions - Morning

5.2.1 Amphitheater 25 - MS Stochastic Computation and Complexity

Speaker: Łukasz Stępień

Adaptive step-size control for global approximation of SDEs driven by countably dimensional Wiener process

Łukasz Stępień
AGH University of Science and Technology, Poland, lstepie@agh.edu.pl

In this paper we deal with global approximation of solutions of stochastic differential equations (SDEs)

$$\begin{aligned} dX(t) &= a(t, X(t))dt + \sigma(t)dW(t), \quad t \in [0, T], \\ X(0) &= x_0 \end{aligned}$$

driven by countably dimensional Wiener process $W(t) = [W_1(t), W_2(t), \ldots]^T$. Under certain regularity conditions imposed on the coefficients a, σ, we show lower bounds for the exact asymptotic error behaviour. To this end, we analyse separately two classes of admissible algorithms: based on equidistant, or arbitrary meshes. Our results indicate that in both cases, decrease of a method error requires significant increase of the cost term, which is illustrated by the product of cost and global error diverging to infinity. This is, however, not visible in the finite dimensional case. In addition, we propose an implementable, path-independent Euler algorithm with adaptive step-size control, which is asymptotically optimal among algorithms using specified truncation levels of the underlying Wiener process. Our theoretical findings are supported by numerical experiments.
In the aforementioned work, we extend the asymptotic results for global approximation from [1] to SDEs with countably dimensional noise structure by utilising strategy presented in [2] where pointwise setting was investigated.

[1] N. Hofmann, T. Müller–Gronbach, and K. Ritter. *Optimal approximation of stochastic differential equations by adaptive step-size control.* Mathematics of Computation, **69**, 1017–1034, 1999.

[2] P. Przybyłowicz, M. Sobieraj, Ł. Stępień. *Efficient approximation of SDEs driven by countably dimensional Wiener process and Poisson random measure.* SIAM Journal on Numerical Analysis , **60**, 824–855, 2022.

Speaker: Monika Eisenmann

Domain decomposition methods for SPDEs

Evelyn Buckwar
Johannes Kepler University Linz, Austria, evelyn.buckwar@jku.at

Ana Djurdjevac
FU Berlin, Germany, adjurdjevac@zedat.fu-berlin.de

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

In recent years, SPDEs have become a well-studied field in mathematics. With their increase in popularity, it becomes important to efficiently approximate their solutions. Thus, modern time-stepping methods need to be developed.
Operator splitting schemes are a powerful tool for deterministic and stochastic differential equations. An example is given by domain decomposition schemes, where we split the domain into sub-domains. Instead of solving one expensive problem on the entire domain, we deal with cheaper problems on the sub-domains. This is particularly useful in modern computer architectures, as the sub-problems may often be solved in parallel. While splitting methods have already been used to study domain decomposition methods for deterministic PDEs, this is a new approach for SPDEs.
We provide an abstract convergence analysis of a splitting scheme for SPDEs and state a domain decomposition scheme as an application of the setting. The theoretical results are verified through numerical experiments.

Speaker: Sotirios Sabanis

Adaptive stochastic optimizers, Euler-Krylov's polygonal approximations and the training of neural nets

Sotirios Sabanis
University of Edinburgh, Edinburgh, UK // Alan Turing Institute, London, UK // National Technical University of Athens, Athens, Greece, s.sabanis@ed.ac.uk

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]. We named this new form Euler-Krylov's polygonal approximations and used it to create new, stochastic (adaptive) optimization algorithms with superior performance in the training of artificial neural networks, see [3] and [4]. Key findings of this new methodology will be reviewed and their links to diffusion generative models will be highlighted.

[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] 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.

[4] D.-Y. Lim, A. Neufeld, Y. Zhang and S. Sabanis *Langevin dynamics based algorithm e-THεO POULA for stochastic optimization problems with discontinuous stochastic gradient.* arXiv preprint, arXiv:2210.13193, 2022.

Speaker: Tomasz Bochacik

On error bounds, optimality and exceptional sets for selected randomized schemes for ODEs

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

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

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

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

We will discuss error bounds and optimality (in the Information-Based Complexity sense) for selected randomized schemes approximating solutions of ODEs. In particular, we will investigate randomized Euler schemes (explicit and implicit) and the randomized midpoint scheme under inexact information and mild assumptions about the right-hand side function (local Hölder and Lipschitz continuity in time and space variables, respectively), cf. [2, 3, 4, 6]. Furthermore, we will discuss the concept of exceptional set for randomized schemes for ODEs [5] and we will establish an upper bound on the probability of exceptional set for a particular class of randomized methods, including Euler and midpoint schemes, under inexact information [1, 2]. Finally, we will report the results of numerical experiments which illustrate theoretical findings discussed in this talk.

[1] T. Bochacik. *On the properties of the exceptional set for the randomized Euler and Runge-Kutta schemes.* Advances in Computational Mathematics, 2023 (accepted).

[2] T. Bochacik. *Randomized algorithms approximating solutions of ordinary differential equations.* PhD Thesis, 2023+ (in preparation).

[3] T. Bochacik, M. Goćwin, P. M. Morkisz, and P. Przybyłowicz. *Randomized Runge-Kutta method – Stability and convergence under inexact information.* Journal of Complexity, 65, 101554, 2021.

[4] T. Bochacik and P. Przybyłowicz. *On the randomized Euler schemes for ODEs under inexact information.* Numerical Algorithms, 91, 1205–1229, 2022.

[5] S. Heinrich and B. Milla. *The randomized complexity of initial value problems.* Journal of Complexity, 24, 77–88, 2008.

[6] R. Kruse and Y. Wu. *Error analysis of randomized Runge-Kutta methods for differential equations with time-irregular coefficients.* Computational Methods in Applied Mathematics, 17, 479–498, 2017.

5.2.2 Room 101, corridor 24-25 - MS PDMPs for high dimensional sampling theory and application

Speaker: Alexandre Bouchard-Côté

PDMPs as Monte Carlo algorithms models

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

Models are widely used to gain insight from data, but there is also a large potential for using models to understand and improve complex algorithms, in particular, Monte Carlo algorithms. In this talk I will show some examples where Piecewise Deterministic Markov Processes (PDMPs) are useful models for Monte Carlo algorithms based on distribution continua.
The PDMP that emerges from this analysis has an event rate that we call the communication barrier. We show that the communication barrier is a useful object to study and estimate, for a variety of reasons ranging from methodological (adaptation, Monte Carlo standard error bounds, distributed algorithms) to theoretical (geometric ergodicity, dimensional scaling).
Our methodological findings are implemented in an open source package, Pigeons, allowing the user to leverage clusters of 1000s of nodes to speed-up difficult Monte Carlo problems without requiring knowledge of distributed algorithms. Pigeons is available at https://github.com/Julia-Tempering/Pigeons.jl

Speaker: Nawaf Bou-Rabee

Randomized Time Integrator for Unadjusted Hamiltonian MCMC

Nawaf Bou-Rabee
Rutgers, The State University of New Jersey, USA, nawaf.bourabee@rutgers.edu

A simple randomized time integrator is suggested for integrating the underlying Hamiltonian dynamics in unadjusted Hamiltonian Monte Carlo (uHMC) in place of the usual Verlet integrator. For absolutely continuous target distributions μ with K-strongly convex and L-gradient Lipschitz potential, and an initial distribution with finite second moment ν, uHMC with this randomized time integrator is able to achieve ϵ-accuracy in L^2-Wasserstein distance with $O((d/K)^{1/3}(L/K)^{5/3}\epsilon^{-2/3}\log(W^2(\nu,\mu)/\epsilon))$ gradient evaluations; whereas the corresponding complexity of uHMC with Verlet is

$$O((d/K)^{1/2}(L/K)^2\epsilon^{-1}\log(W^2(\nu,\mu)/\epsilon)).$$

Duration randomization is shown to additionally improve the complexity for a reason analogous to time integrator randomization. This is joint work with Milo Marsden (Stanford).

Speaker: Peter A. Whalley

Randomized Time Riemannian Manifold Hamiltonian Monte Carlo

Peter A. Whalley
University of Edinburgh, Edinburgh, UK, p.a.whalley@sms.ed.ac.uk

Daniel Paulin
University of Edinburgh, Edinburgh, UK, dpaulin@ed.ac.uk

Benedict Leimkuhler
University of Edinburgh, Edinburgh, UK, b.leimkuhler@ed.ac.uk

Hamiltonian Monte Carlo (HMC) algorithms which combine numerical approximation of Hamiltonian dynamics on finite intervals with stochastic refreshment and Metropolis correction are popular sampling schemes, but it is known that they may suffer from slow convergence in the continuous time limit. A recent paper of Bou-Rabee and Sanz-Serna (Ann. Appl. Prob., 27:2159-2194, 2017) demonstrated that this issue can be addressed by simply randomizing the duration parameter of the Hamiltonian paths. In this article, we use the same idea to enhance the sampling efficiency of a constrained version of HMC, with potential benefits in a variety of application settings. We demonstrate both the conservation of the stationary distribution and the ergodicity of the method. We also compare the performance of various schemes in numerical studies of model problems, including an application to high-dimensional covariance estimation.

Speaker: Torben Sell

Gradient-Based Markov Chain Monte Carlo for Bayesian Inference With Non-Differentiable Priors

Torben Sell
University of Edinburgh, Edinburgh, UK, torben.sell@ed.ac.uk

Jacob Vorstrup Goldman
University of Cambridge, Cambridge, UK, jdv22@eng.cam.ac.uk

Sumeetpal Sidhu Singh
University of Cambridge, Cambridge, UK, sss40@cam.ac.uk

The use of non-differentiable priors in Bayesian statistics has become increasingly popular, in particular in Bayesian imaging analysis. Current state of the art methods are approximate in the sense that they replace the posterior with a smooth approximation via Moreau-Yosida envelopes, and apply gradient-based discretized diffusions to sample from the resulting distribution. We characterize the error of the Moreau-Yosida approximation and propose a novel implementation using underdamped Langevin dynamics. In misson-critical cases, however, replacing the posterior with an approximation may not be a viable option. Instead, we show that Piecewise-Deterministic Markov Processes (PDMP) can be utilized for exact posterior inference from distributions satisfying almost everywhere differentiability. Furthermore, in contrast with diffusion-based methods, the suggested PDMP-based samplers place no assumptions on the prior shape, nor require access to a computationally cheap proximal operator, and consequently have a much broader scope of application. Through detailed numerical examples, including a non-differentiable circular distribution and a non-convex genomics model, we elucidate the relative strengths of these sampling methods on problems of moderate to high dimensions, underlining the benefits of PDMP-based methods when accurate sampling is decisive.

5.2.3 Room 102, corridor 15-25 - MS Exploring the intersections of importance sampling, MCMC, and optimization

Speaker: Juan Kuntz

Particle-Based Algorithms for Maximum Likelihood Estimation of Latent Variable Models

Juan Kuntz
Department of Statistics, University of Warwick, UK, juan.kuntz-nussio@warwick.ac.uk

Jen Ning Lim
Department of Statistics, University of Warwick, UK, jen-ning.lim@warwick.ac.uk

Adam M. Johansen
Department of Statistics, University of Warwick, UK, a.m.johansen@warwick.ac.uk

Neal and Hinton (1998) recast maximum likelihood estimation of any given latent variable model as the minimization of a free energy functional F on an infinite-dimensional space, and the EM algorithm as coordinate descent applied to F. Here, we explore alternative ways to optimize the functional. In particular, we identify various gradient flows associated with F and show that their limits coincide with F's stationary points. By discretizing the flows, we obtain practical particle-based algorithms for maximum likelihood estimation in broad classes of latent variable models. The novel algorithms scale to high-dimensional settings and perform well in numerical experiments.

[1] J. Kuntz, J. N. Lim, and A. M. Johansen. *Scalable Particle-Based Alternatives to EM.* arXiv preprint arXiv:2204.12965, 2022.

Speaker: Jimmy Olsson

PaRISian particle Gibbs samplers for state and parameter learning in nonlinear state-space models

Gabriel Cardoso
Electrophysiology and Heart Modeling Institute, Pessac, France,
gabriel.victorino-cardoso@polytechnique.edu

Yazid Janati
Télécom SudParis, Palaiseau, France, janati.yazid@gmail.com

Sylvain Le Corff
Sorbonne Université, Paris, France, sylvain.lecorff@gmail.com

Eric Moulines
Ecole polytechnique, Palaiseau, France, eric.moulines@polytechnique.edu

Jimmy Olsson
KTH Royal Institute of Technology, Stockholm, Sweden, jimmyol@kth.se

Nonlinear state-space models, being the most classical generative models for serial data and sequences in general, are ubiquitous in statistical machine learning. The particle-based, rapid incremental smoother (PaRIS) is a sequential Monte Carlo technique allowing for efficient online approximation of expectations of additive functionals under joint-state posteriors in these models. Such expectations appear naturally in several learning contexts, such as maximum likelihood estimation (MLE) and Markov score climbing (MSC). The PaRIS has linear computational complexity, limited memory requirements, and comes with non-asymptotic bounds, convergence results, and stability guarantees. Still, being based on self-normalised importance sampling, the PaRIS is biased; thus, In this talk we discuss a novel approach, the PaRISian particle Gibbs (PPG) sampler, which can be viewed as a PaRIS algorithm driven by conditional SMC moves, resulting in bias-reduced estimates of the targeted quantities. We substantiate the PPG algorithm with theoretical results, including new bounds on its bias and variance as well as deviation inequalities. Moreover, we show how the PPG can be successfully applied in a general learning framework covering stochastic-gradient MLE and MSC as special cases. In this context, we establish, under standard assumptions, non-asymptotic bounds highlighting the value of bias reduction and the implicit Rao–Blackwellisation of PPG. These are the first non-asymptotic results of this kind in this setting. The proposed methodology and theoretical claims are illustrated by numerical experiments.

[1] G. Cardoso, E. Moulines, and J. Olsson. *Particle-based, rapid incremental smoother meets particle Gibbs.* https://arxiv.org/abs/2209.10351.

[2] J. Olsson and J. Westerborn. *Efficient particle-based online smoothing in general hidden Markov models: the PaRIS algorithm.* Bernoulli, 23:3, 1951–1996, 2017.

Speaker: Víctor Elvira

MCMC-driven adaptive importance samplers

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

In this work, we first introduce an adaptive importance sampling framework that combines the benefits of the importance sampling (IS) and Markov chain Monte Carlo (MCMC) approaches. The framework is called *layered adaptive importance sampling* (LAIS). Different parallel MCMC chains (upper layer) provide the location parameters of the proposal probability density functions (pdfs) used in an IS method (lower layer). Then, we consider several specific implementations of LAIS, e.g., the parallel random walk Metropolis variant of LAIS and the *anti-tempered* LAIS where the MCMC algorithms sample from an *anti-tempered* version of the posterior distribution (with a reduced scaling). We also provide a theoretical support to explain the different strategies, including the anti-tempered variant. Numerical results also confirm the advantages of the proposed schemes.

[1] L. Martino, V. Elvira, D. Luengo, and J. Corander. Layered adaptive importance sampling. *Statistics and Computing*, 27(3): 599–623, 2017.

[2] L. Martino, V. Elvira, and D. Luengo. Anti-tempered layered adaptive importance sampling. In *2017 22nd International Conference on Digital Signal Processing (DSP)*, pages 1–5. IEEE, 2017.

Speaker: Dootika Vats

Comparing apples to oranges: a universal effective sample size

Medha Agarwal
University of Washington, USA, medhaaga@uw.edu

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

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

Effective sample size (ESS) is a popular, powerful, and practical numerical summary for assessing the performance of a Markov chain Monte Carlo (MCMC) sampler. ESS estimates the number of iid samples that would return the same variance of an estimator, as a given MCMC sample. The idea of an ESS is also used in importance sampling to determine the quality of the proposal, although it is often remarked that comparing the ESS in importance sampling to the ESS in MCMC would be like comparing apples to oranges. In this talk, I will compare apples to oranges! I present a unifying framework for ESS that allows users to compare MCMC and importance sampling for a given estimation problem. I further discuss how the ESS can be employed to arrive at principled stopping rules for simulations. Some open problems and practical concerns will be presented in addition to a few examples.

[1] M. Agarwal, D. Vats, V. Elvira. A principled stopping rule for importance sampling. *Electronic Journal of Statistics*, 16(2): 5570-5590, 2022.

5.2.4 Room 104, corridor 15-25 - Contributed Talks

Speaker: Emil Løvbak

Adjoint Monte Carlo particle methods with reversible random number generators

Emil Løvbak
KU Leuven, Belgium, emil.loevbak@kuleuven.be

Giovanni Samaey
KU Leuven, Belgium, giovanni.samaey@kuleuven.be

When solving optimization problems constrained by high-dimensional PDEs, Monte Carlo particle methods are often the only practical approach to simulate the PDE. Unfortunately, these methods introduce noise in the computed particle distributions and, as a consequence, evaluations of the objective function. Through an adjoint-based approach, we can compute the corresponding gradient down to machine precision through a similar Monte Carlo simulation. However, this approach requires retracing the particle trajectories from the constraint simulation, backwards in time, when computing the gradient. When storing these paths for large-scale simulations, one quickly runs into memory issues. In this talk, we solve these memory issues by regenerating particle trajectories backward in time. To do so, we use of the reversible pseudorandom number generator described in [1] to generate the paths in the constraint simulation. After describing our reversible approach, we demonstrate how it outperforms prior approaches on some concrete test-problems.

[1] E. Løvbak, F. Blondeel, A. Lee, L. Vanroye, A. Van Barel, S. Samaey *Reversible random number generation for adjoint Monte Carlo simulation of the heat equation.* arXiv:2302.02778, 2023.

Speaker: Giorgos Vasdekis

Pseudo-marginal Piecewise Deterministic Monte Carlo

Richard Everitt
University of Warwick, U.K., richard.everitt@warwick.ac.uk

Giorgos Vasdekis
University College London, U.K, G.Vasdekis@ucl.ac.uk

Piecewise Deterministic Markov Processes (PDMPs) have recently caught the attention of the MCMC community for having a non-diffusive behavior, potentially allowing them to explore the state space efficiently. This makes them good candidates to generate MCMC algorithms. One important problem in Bayesian computation is inference for models where pointwise evaluation of the posterior is not available, but one has access to an unbiased estimator of the posterior. A technique to deal with this problem is the Pseudo-marginal Metropolis Hastings. In this talk we describe a PDMP algorithm that can be used in the same posterior free setting and can be seen as the analogue of Pseudo-marginal for Piecewise Deterministic Monte Carlo. We show that the algorithm targets the posterior of interest and we provide numerical examples, focusing on the case of Approximate Bayesian Computation (ABC), a popular method to deal with problems in the setting of likelihood free inference .

[1] R. Everitt and G. Vasdekis *Pseudo-marginal Piecewise Deterministic Monte Carlo* Under preparation.

[2] C. Andrieu and G. O. Roberts *The pseudo-marginal approach for efficient Monte Carlo computations* Annals of Statistics, 2009. https://doi.org/10.1214/07-AOS574

[3] M. Sachs and D. Sen and J. Lu and D. Dunson. *Posterior Computation with the Gibbs Zig-Zag Sampler* Bayesian Analysis, 2022. *https://doi.org/10.1214/22-BA1319*

[4] G. Vasdekis and G. O. Roberts *Speed Up Zig-Zag* To appear in Annals of Applied Probability, 2023+. Available at https://arxiv.org/abs/2103.16620.

Speaker: Yu Guang Wang

Applied Harmonic Analysis and Particle Dynamics for Designing Neural Message Passing on Graphs

Yu Guang Wang
Shanghai Jiao Tong University, China, yuguang.wang@sjtu.edu.cn

Graph representation learning has broad applications applications from recommendation systems to drug and protein designs. In this talk, I will talk about using harmonic analysis and particle systems to design useful neural message passing with theoretically guaranteed separability and efficient computation. These message passings are proved to have strictly positive lower bounded Dirichlet energy and thus to circumvent the oversmoothing problem appearing in many spatial GNNs, when the node features are indistinguishable as the network deepens.

Speaker: Josef Leydold

A Transformed Density Rejection Based Algorithm for Densities with Poles and Inflection Points

Wolfgang Hörmann
Boğaziçi University, Türkiye, hormannw@boun.edu.tr

Josef Leydold
WU Vienna University of Economics and Business, Austria, josef.leydold@wu.ac.at

Generation of non-uniform random variates is a crucial part for any stochastic simulation and Monte Carlo integration. Besides specialized algorithms for standard distribution there also exist quite a few automatic methods that work for large classes of distributions. Among them the so called transform density rejection method has very fast marginal generation times but requires a suitable transformation such that the transformed density is either concave or convex. It has been shown that the method also works efficiently if only rough estimates of the inflection points of the transformed density are known. In this contribution we present the main idea behind a algorithm called "Tinflex" and show that it can also be used for unbounded densities where the location of the pole is known.

[1] W. Hörmann and J. Leydold. A generalized transformed density rejection algorithm. In Z. Botev, A. Keller, C. Lemieux, and B. Tuffin, editors, *Advances in Modeling and Simulation: Festschrift for Pierre L'Ecuyer*, pages 283–300. Springer International Publishing, 2022.

5.3 Parallel sessions - Afternoon

5.3.1 Amphitheater 25 - MS Numerical methods in statistical physics

Speaker: Michela Ottobre

Uniform in time approximations of stochastic dynamics

Michela Ottobre
Heriot Watt University, Edinburgh, UK, m.ottobre@hw.ac.uk

Complicated models, for which a detailed analysis is too far out of reach, are routinely approximated via a variety of procedures; this is the case when we use multiscale methods, when we take many particle limits and obtained a simplified, coarse-grained dynamics, or when we use numerical methods. While approximating, we make an error which is small over small time-intervals but it typically compounds over longer time-horizons. Hence, in general, the approximation error grows in time so that the results of our "predictions" are less reliable when we look at longer time-hormizons. However this is not necessarily the case and one may be able to find dynamics and corresponding approximation procedures for which the error remains bounded, uniformly in time. We will discuss a very general approach to understand when this is possible. I will show how the approach we take is very broad and how it can be used for all of the approximation procedures mentioned above. This is based on a series of joint works with a number of people: L. Angeli, J. Barre', D. Crisan, P. Dobson, B. Goddard, I. Souttar and E. Zatorska.

Speaker: Matthew Dobson

Steady-State Solutions for Nonequilibrium Langevin Dynamics

Matthew Dobson
University of Massachusetts Amherst, USA, dobson@umass.edu

Abdel Kader Geraldo
University of Massachusetts Amherst, USA, ageraldo@umass.edu

Nonequilibrium molecular dynamics simulation of incompressible flows under periodic boundary conditions involve deforming boundary conditions and specialized nonequilibrium forcing. For planar background flows, we use the automorphism remapping periodic boundary conditions Lees-Edwards PBCs and Kraynik-Reinelt PBCs to treat respectively shear flow and planar elongational flow. We prove that the dynamics converges exponentially fast to a steady-state limit cycle, and show the convergence of numerical integrators for the dynamics.

[1] M. Dobson and A. K. Geraldo. *Convergence of Nonequilibrium Langevin Dynamics for Planar Flows.* arXiv:2208.14358.

Speaker: Renato Spacek

Extending the regime of linear response with synthetic forcings

Renato Spacek
École des Ponts ParisTech et Inria Paris, France, renato.spacek@enpc.fr

Gabriel Stoltz
École des Ponts ParisTech et Inria Paris, France, gabriel.stoltz@enpc.fr

Transport coefficients, such as the mobility, thermal conductivity and shear viscosity, are quantities of prime interest in statistical physics. At the macroscopic level, transport coefficients relate an external forcing of magnitude η, with $\eta \ll 1$, acting on the system to an average response expressed through some steady-state flux. In practice, steady-state averages involved in the linear response are computed as time averages over a realization of some stochastic differential equation. Variance reduction techniques are of paramount interest in this context, as the linear response is scaled by a factor of $1/\eta$, leading to large statistical error.
One way to limit the increase in the variance is to allow for larger values of η by increasing the range of values of the forcing for which the nonlinear part of the response is sufficiently small. In theory, one can add an extra forcing to the physical perturbation of the system, called synthetic forcing, as long as this extra forcing preserves the invariant measure of the reference system. The aim is to find synthetic perturbations allowing to reduce the nonlinear part of the response as much as possible. In this talk, I will present a mathematical framework for quantifying the quality of synthetic forcings, in the context of linear response theory, and discuss various possible choices for them. I will illustrate my analysis with numerical results on the computation of the mobility in low dimensional systems.

Speaker: Grigorios Pavliotis

Optimal Langevin Samplers

Grigorios Pavliotis
Imperial College London, UK, pavl@ic.ac.uk

Sampling from a probability distribution in a high dimensional space is a standard problem in computational statistical mechanics, Bayesian statistics and other applications. A standard approach for doing this is by constructing an appropriate Markov process that is ergodic with respect to the measure from which we wish to sample. In this talk we will present a class of sampling schemes based on Langevin-type stochastic differential equations. We will show, in particular, how the performance of the sampling scheme can be improved by breaking detailed balance [6, 3], by modifying the underdamped Langevin dynamics [4] and by preconditioning the dynamics [2]. In addition, we make the link with the theory of (weakly) interacting diffusions [1], and we also introduce an appropriate notion of optimality (for Langevin samplers), based on the decision-theoretic framework of Glynn and Witt [5].

[1] A. Borovykh, N. Kantas, P. Parpas, and G. A. Pavliotis. On stochastic mirror descent with interacting particles: convergence properties and variance reduction. *Phys. D*, 418:Paper No. 132844, 21, 2021.

[2] M. Chak, N. K., T. Leliévre, and G. A. Pavliotis. Optimal friction matrix for underdamped Langevin sampling, 2021.

[3] A. B. Duncan, T. Lelièvre, and G. A. Pavliotis. Variance Reduction Using Nonreversible Langevin Samplers. *J. Stat. Phys.*, 163(3):457–491, 2016.

[4] A. B. Duncan, N. Nüsken, and G. A. Pavliotis. Using Perturbed Underdamped Langevin Dynamics to Efficiently Sample from Probability Distributions. *J. Stat. Phys.*, 169(6):1098–1131, 2017.

[5] P. W. Glynn and W. Whitt. The asymptotic efficiency of simulation estimators. *Oper. Res.*, 40(3):505–520, 1992.

[6] T. Lelievre, F. Nier, and G. A. Pavliotis. Optimal non-reversible linear drift for the convergence to equilibrium of a diffusion. *J. Stat. Phys.*, 152(2): 237–274 , 2013.

5.3.2 Room 101, corridor 24-25 - MS High dimensional approximation

Speaker: Abi Srikumar

Multi-level quasi-Monte Carlo methods for kernel interpolation in uncertainty quantification

Alexander Gilbert
UNSW Sydney, alexander.gilbert@unsw.edu.au

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

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

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

Abi Srikumar
UNSW Sydney, a.srikumar@unsw.edu.au

As high-dimensional problems become increasingly prevalent in many applications, the effective evaluation of these problems within the limits of our current technology poses a great hurdle due to the exponential increase in computational cost as dimensionality increases. One class of strategies for evaluating such problems efficiently are quasi-Monte Carlo (QMC) methods.
In this talk, we explore the effectiveness of multi-level quasi-Monte Carlo methods for approximating solutions to elliptic partial differential equations with stochastic coefficients that depend periodically on the stochastic parameters. In particular, we are interested in fast approximation using kernel-based lattice point interpolation. We present some theoretical results regarding the error convergence of such approximations and the results of numerical experiments.

Speaker: Dirk Nuyens

Higher order lattice rules on $\mathbf{R}^d$

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

Lattice rules are traditionally used to approximate integrals on the unit cube. If the function can be expanded into an absolutely converging Fourier series, then it is particularly convenient to obtain error bounds in terms of the decay of the Fourier coefficients.
When the integral is on $\mathbf{R}^d$, then there are several ways of still making use of lattice rules. Some of them are able to retain the smoothness of the integrand and allow us to get higher order of convergence...

Speaker: Yoshihito Kazashi

Density estimation in RKHS with application to Korobov spaces in high dimensions

Yoshihito Kazashi
University of Strathclyde, UK, y.kazashi@strath.ac.uk

Fabio Nobile
EPF Lausanne, Switzerland, fabio.nobile@epfl.ch

I will be talking about a kernel-based method for estimating a probability density function (pdf) from an i.i.d. sample. Our estimator is a linear combination of kernel functions, the coefficients of which are determined by a linear equation. I will present an error analysis for the mean integrated squared error in a general reproducing kernel Hilbert space setting. This theory is then applied to estimate pdfs belonging to weighted Korobov spaces, for which a dimension independent convergence rate is established. Under a suitable smoothness assumption, our method attains a rate arbitrarily close to the optimal rate.

[1] Y. Kazashi and F. Nobile. Density estimation in RKHS with application to Korobov spaces in high dimensions. *SIAM J. Numer. Anal.*, to appear

Speaker: Ian H. Sloan

Periodic kernel-based high-dimensional approximation

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

This talk describes a recent periodic-kernel-based QMC method for high-dimensional approximation. In this work, jointly with Vesa Kaarnioja, Yoshihito Kazashi, Frances Kuo and Fabio Nobile, an elliptic partial differential equation with a random input field is modelled in a non-standard way with periodic random variables together with kernel-based interpolation, giving a computational cost that grows merely linearly with dimensionality. The method is feasible even with hundreds of parameters in an input random field.

5.3.3 Room 102, corridor 15-25 - MS Multilevel MC techniques for discontinuous functionals

Speaker: Sebastian Krumscheid

Multilevel Monte Carlo methods for parametric expectations: distribution and robustness measures

Quentin Ayoul-Guilmard
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Switzerland,
quentin.ayoul-guilmard@epfl.ch

Sundar Ganesh
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Switzerland,
sundar.ganesh@epfl.ch

Sebastian Krumscheid
Karlsruhe Institute of Technology, Germany, sebastian.krumscheid@kit.edu

Fabio Nobile
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Switzerland,
fabio.nobile@epfl.ch

The multilevel Monte Carlo (MLMC) method is an efficient sampling method for the approximation of expected system outputs and is applicable to a wide range of applications. However, its use for quantities of interest that cannot be expressed as expected values is not straightforward. In this talk, we will discuss recent advances that allow for a more informative characterization of a system output's distribution using the MLMC method. Specifically, we will introduce MLMC techniques for approximating generic parametric expectations, that is, expected values depending on a parameter. The resulting MLMC estimator for functions allows for the derivation of efficient approximations to further characterize a system output's distribution. First, we illustrate its use for estimating the characteristic function of a quantity of interest with a cumulative distribution that is only piecewise continuous (i.e., a mixed distribution). We will then outline a procedure for constructing MLMC estimators for robustness indicators, focusing on those quantifying "tail" risks using quantiles (value-at-risk) or conditional values-at-risk. While these techniques offer near-optimal computational complexity, their direct application may be hindered in practice when relying on the theoretical a priori error estimates. To address this gap, we will present novel computable error estimators for optimally tuning the MLMC methods, leading to an efficient and robust continuation type adaptive algorithm.

Speaker: Fabio Nobile

MLMC for the computation of CVaR and its sensitivities in PDE-constrained risk-averse optimization

Sundar Ganesh
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Switzerland,
sundar.ganesh@epfl.ch

Fabio Nobile
Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Switzerland,
fabio.nobile@epfl.ch

In this work, we tackle the problem of minimising the Conditional-Value-at-Risk (CVaR) of output quantities of complex differential models with random input data, using gradient- based approaches in combination with the Multi-Level Monte Carlo (MLMC) method. In particular, we consider the framework of multi-level Monte Carlo for parametric expecta- tions introduced in [1] and propose modifications of the MLMC estimator, error estimation procedure, and adaptive MLMC parameter selection to ensure the estimation of the CVaR and sensitivities for a given design with a prescribed accuracy. We then propose combining the MLMC framework with an alternating inexact minimisation-gradient descent algorithm, for which we prove exponential convergence in the optimisation iterations under the assumptions of strong convexity and Lipschitz continuity of the gradient of the objective function. We demonstrate the performance of our approach on two numerical examples of practical relevance, which evidence the same optimal asymptotic cost-tolerance behaviour as standard MLMC methods for fixed design computations of output expectations.

[1] Sebastian Krumscheid and Fabio Nobile. *Multilevel Monte Carlo approximation of functions.* SIAM/ASA Journal on Uncertainty Quantification 6.3 (2018), pp. 1256–1293.

Speaker: Cedric Beschle

CLMC techniques for elliptic PDEs with random discontinuities

Andrea Barth
University of Stuttgart, IANS, andrea.barth@mathematik.uni-stuttgart.de

Cedric Beschle
University of Stuttgart, IANS, cedric.beschle@mathematik.uni-stuttgart.de

The accurate and efficient estimation of stochastic moments of (functionals of) solutions to random problems is of high interest in the field of uncertainty quantification. Adding discontinuities to the underlying stochastic model may lead to large local effects on the solutions and inefficiencies in the moment estimation. Standard multilevel Monte Carlo methods are robust, but not able to efficiently account for these local effects, resulting in high computational cost. To this end, we consider the continuous level Monte Carlo method (CLMC) allowing for samplewise adaptivity to local solution features with the potential of a high cost reduction. However, this computational advantage is not necessarily reflected in the estimation of moments via standard CLMC. We motivate a variant of CLMC that leverages the low samplewise approximation cost onto the estimation of stochastic moments. Its improved performance is demonstrated via an application to a random elliptic PDE with discontinuous diffusion coefficient.

Speaker: Daniel Roth

Multilevel Monte Carlo Learning

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

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

Daniel Roth
Qontigo GmbH, droth@qontigo.com

Martin Simon
Frankfurt University of Applied Sciences, martin.simon@fb2.fra-uas.de

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. [2] with the idea of multilevel Monte Carlo path simulation of Giles [1]. 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 [1]. 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] Giles, Michael B. *Multilevel monte carlo path simulation* Operations Research, 56, 607–617, 2008

[2] 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

5.3.4 Room 104, corridor 15-25 - Contributed Talks

Speaker: Nicolas Chopin

Higher-order stochastic integration through cubic stratification

Nicolas Chopin
ENSAE, IPP Paris, nicolas.chopin@ensae.fr

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

We propose two novel unbiased estimators of the integral $\int_{[0,1]^s} f(u)du$ for a function f, which depend on a smoothness parameter $r \in \mathbb{N}$. The first estimator integrates exactly the polynomials of degrees $p < r$ and achieves the optimal error $n^{-1/2-r/s}$ (where n is the number of evaluations of f) when f is r times continuously differentiable. The second estimator is computationally cheaper but it is restricted to functions that vanish on the boundary of $[0,1]^s$. The construction of the two estimators relies on a combination of cubic stratification and control variates based on numerical derivatives. We provide numerical evidence that they show good performance even for moderate values of n.

[1] N. Chopin and M. Gerber. *Higher-order stochastic integration through cubic stratification.* arXiv:2210.01554

Speaker: Pierre L'Ecuyer

Improved Versions of the Lattice Tester and LatMRG Software Tools

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

Mamadou Thiongane
University Cheikh Anta Diop, Dakar, Senegal, mdthiongane@yahoo.fr

Christian F. Weiss
Ruhr West University of Applied Sciences, Germany, Christian.Weiss@hs-ruhrwest.de

LatMRG is a large software library offering tools to study the lattice structure of linear random number generators and search for generators with a good lattice structure. It was originally written in Modula-2 [1], a language no longer supported. A partial C++ implementation was made available recently, but it was still missing important features which we now offer in our latest version. LatNet Builder use Lattice Tester, another C++ software tool whose purpose is to measure the quality of lattices in general, and which has been improved as well.
The aim of this talk is to give an overview of the latest versions of Lattice Tester and LatMRG, show examples of what they can do, and highlight the recent improvements. These new versions will be available on GitHub at the time of the conference, together with detailed user guides.

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

Speaker: David Métivier

The Robust Quasi Monte Carlo Method

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

Matthieu Lerasle
Ensae-CREST, CNRS, France, matthieu.lerasle@ensae.fr

David Métivier
MISTEA, INRAe, France, david.metivier@inrae.fr

Given a simulation budget of B points to calculate an expectation/integral, a Monte Carlo method achieves a mean squared error as $1/\sqrt{B}$. Quasi Monte Carlo methods and randomized versions are asymptotically faster. The question addressed in this presentation is, given a budget B and some confidence level δ, what is the optimal confidence interval size one can build? For which estimator? We show that a judicious choice of "robust" aggregation methods coupled with Quasi Monte Carlo techniques allows to reach the optimal error bound. In this talk, I will present Quasi Monte Carlo methods, different concentration inequalities and robust mean estimators (old and new) to get to the solution, with supporting numerical experiments.

[1] E. Gobet, M. Lerasle, D. Métivier. *Mean estimation for randomized quasi Monte Carlo method* (Submitted) https://hal.science/hal-03631879.

Speaker: Sergei Kucherenko

Active Subspaces for Problems with Dependent Variables using QMC Sampling

Sergei Kucherenko
Imperial College London, SWT 2AZ, UK, s.kucherenko@imperial.ac.uk

Peter Yatsyshin
The Alan Turing Institute, NW1 2DB, UK, pyatsyshin@turing.ac.uk

Nilay Shah
Imperial College London, SWT 2AZ, UK, n.shah@imperial.ac.uk

The output of a model may only depend on a limited number of directions in the input space. The active subspaces method (ASM) identifies important directions in the parameter space which can allow significant dimension reduction [1]. Each direction corresponds to a linear combination of the original input variables, which is referred to as the active subspace. The active subspace is identified by performing the eigenvalue decomposition of the covariance-like matrix C of response gradients.
Traditional methods of global sensitivity analysis rank variables and allow dimension reduction only for models whose parameters are aligned along the axes of the parameter space. The ASM allows dimension reduction regardless of the orientation of the important directions in the parameter space. Applying ASM to high-dimensional problems, metamodels can be built in low dimensional active subspaces at low computational costs. We show that using such metamodels variance based Sobol' indices can be accurately computed for all the inputs at the reduced costs.
We generalised the ASM for the case of models with dependent variables and compared the generalised ASM with that of derivative based global sensitivity measures extended for models with dependent variables [2] and generalised Sobol' indices [3].
We show that application of QMC sampling instead of traditionally used MC significantly increases both the accuracy of computing parameters of ASM (the elements of matrix C and the values of eigenvalues and eigenvectors) and metamodels built in low dimensional active subspaces.

[1] P.G. Constantine. *Active subspaces. Emerging Ideas for Dimension Reduction in Parameter Studies.* Published in SIAM spotlights, 2015.

[2] M. Lamboni and S. Kucherenko. *Multivariate sensitivity analysis and derivative-based global sensitivity measures with dependent variables.* Reliability Engineering and System Safety, 212, 107519, 2021.

[3] S. Kucherenko, S. Tarantola and P. Annoni. *Estimation of global sensitivity indices for models with dependent variables.* Computer Physics Communications, 183, 937-–946, 2012.

Index of Authors

Index of Speakers