CENTRE DE RECHERCHES MATHÉMATIQUES · DIRO · GERAD GROUP FOR RESEARCH IN DECISION ANALYSIS · IVADO INSTITUT DE VALORISATION DES DONNÉES · Université de Montréal

MCM 2017

International Conference on Monte Carlo Methods and Applications

Conference Program and Information

HEC MONTRÉAL

July 3-7, 2017

Montréal skyline from parc Jean-Drapeau
© Parc Jean-Drapeau, Denis Labine

Contents

Welcome to Montreal for MCM 2017

We are pleased and excited to host you in Montreal for the *Eleventh International Conference on Monte Carlo Methods and Applications*. We hope you will have a memorable week here, enjoying both the scientific program and the vibrant life of the city.

Based on my examination of the abstracts, I am very happy with the overall quality and variety of the proposed sessions and presentations. This high quality depends much more on you, the participants, than on the organizers.

The first edition of this biennial conference, formerly known as the *IMACS Seminar on Monte Carlo Methods*, was organized in Brussels in April 1997. Successive editions were then held in Varna, Bulgaria (1999), Salzburg, Austria (2001), Berlin, Germany (2003), Tallahassee, USA (2005), Reading, UK (2007), Brussels, Belgium (2009), Borovets, Bulgaria (2011), Annecy, France (2013), Linz, Austria (2015). We extend our gratitude to the organizers of these remarkable events.

Our scientific program features nine one-hour invited plenary talks from top contributors in our field. We also have 140 regular talks of 30 minutes each (including questions), regrouped in 20 special sessions (in which the speakers were mostly invited by the session organizers) and 18 sessions of contributed talks. The regular talks are organized in four parallel tracks.

Following the tradition, a special issue of *Mathematics and Computers in Simulation* will be devoted to a selection of articles based on presentations made at this conference. Instructions on how and when to submit your manuscripts will be provided on the conference web site, soon after the conference: http://www.crm.umontreal.ca/2017/MCM2017/.

All registered participants are invited to a wine and cheese reception on Monday at Le Salon L'Oréal at HEC Montréal, after the presentations, and to the conference banquet in the Old Port of Montreal on Wednesday evening. You are encouraged to visit Montreal by yourself in your free evenings, and perhaps enjoy the world's famous Montreal Jazz Festival, held from June 28 to July 8.

If you have any problems or special requests during the conference, do not hesitate to ask the organizers.

Pierre L'Ecuyer

Sponsoring Societies

We are very grateful to our sponsors, whose financial contributions cover the expenses of the nine invited plenary speakers and part of the other conference expenses. Without their help, the registration fees would have been significantly higher. These sponsors are the Centre de Recherches Mathématiques (CRM), the Département d'Informatique et de Recherche Opérationnelle (DIRO), the Groupe d'Études et de Recherche en Analyse de Décisions (GERAD), the Institut de Valorisation des Données (IVADO) and the Université de Montréal.

Committees and Organizers

The MCM Conference series has a *Steering Committee* whose main task is to define guidelines on the conference topics and format, select the organizers, and provide advice when needed. The committee members are:

- Ronald Cools, *KU Leuven, Belgium*
- Ivan Dimov, *Bulgarian Academy of Science, Bulgaria*
- Stefan Heinrich, *Universität Kaiserslautern, Germany*
- Christian Lécot, *Université de Savoie, France*
- Pierre L'Ecuyer, *Université de Montréal, Canada*
- Karl K. Sabelfeld, *Novosibirsk State University, Russia*
- Wolfgang Ch. Schmid, *Universität Salzburg, Austria*
- Paula Whitlock, *City University of New York, USA*

The role of the conference *Program Committee* is to make sure that the conference presentations and the proceedings are of the highest possible quality. This includes suggesting plenary speakers, organizing special sessions, reviewing abstracts of contributed talks, and reviewing (or handling the revision) of papers submitted to the proceedings. The members of the Program Committee for MCM 2017 are:

- Pierre L'Ecuyer, *Université de Montréal, Canada* (Chair and conference organizer)
- Fabian Bastin, *Université de Montréal, Canada*
- Jose Blanchet, *Columbia University, USA*
- Zdravko Botev, *UNSW, Australia*
- Nicolas Chopin, *ENSAE, France*
- Ronald Cools, *KU Leuven, Belgium*
- Aaron Courville, *Université de Montréal, Canada*
- Luc Devroye, *McGill University, Canada*
- Josef Dick, *University of New South Wales, Australia*
- Ivan Dimov, *Bulgarian Academy of Science, Bulgaria*
- Arnaud Doucet, *University of Oxford, United Kingdom*
- Emma Frejinger, *Université de Montréal, Canada*
- Stefan Geiss, *University of Jyväskylä, Finland*
- Mike Giles, *University of Oxford, United Kingdom*
- Paul Glasserman, *Columbia University, USA*
- Emmanuel Gobet, *École Polytechnique, France*
- Stefan Heinrich, *Universität Kaiserslautern, Germany*

- Fred Hickernell, *Illinois Institute of Technology, USA*
- Peter Kritzer, *Johann Radon Institute for Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences*
- Dirk Kroese, *Univeristy of Queensland, Australia*
- Frances Kuo, *University of New South Wales, Australia*
- Simon Lacoste-Julien, *Université de Montréal, Canada*
- Gerhard Larcher, *Johannes Kepler Universität Linz, Austria*
- Christian Lécot, *Université de Savoie, France*
- Gunther Leobacher, *University of Graz, Austria*
- Christiane Lemieux, *University of Waterloo, Canada*
- Thomas Müller-Gronbach, *Universität Passau, Germany*
- Makoto Matsumoto, *Hiroshima University, Japan*
- Eric Moulines, *École Polytechnique, Paris, France*
- Harald Niederreiter, *Österreichische Akademie der Wissenschaften, Austria*
- Derek Nowrouzezahrai, *Université de Montréal, Canada*
- Erich Novak, *Friedrich Schiller University, Germany*
- Dirk Nuyens, *KU Leuven, Belgium*
- Art Owen, *Stanford University, USA*
- Raghu Pasupathy, *Virginia Tech, USA*
- Friedrich Pillichshammer, *Johannes Kepler Universität Linz, Austria*
- Klaus Ritter, *TU Kaiserslauten, Germany*
- Christian Robert, *Université Paris-Dauphine, France*
- Gerardo Rubino, *Inria Rennes, France*
- Karl K. Sabelfeld, *Novosibirsk State University, Russia*
- Wolfgang Ch. Schmid, *Universität Salzburg, Austria*
- Bruno Tuffin, *Inria Rennes, France*
- Grzegorz Wasilkowski, *University of Kentucky, USA*
- Paula Whitlock, *City University of New York, USA*
- Henryk Wozniakowski, *Columbia University and University of Warsaw, USA/Poland*

The following persons are taking care of the local organization. They are the ones who end up doing most of the work and they deserve our warmest thanks.

- Karine Hébert, GERAD (program and proceedings)
- Marilyne Lavoie, GERAD (graphism and abstract submission)
- Suzette Paradis, CRM (webmaster)
- Louis Pelletier, CRM (invited speakers)
- Marie Perreault, GERAD (registration, facilities, logistics, and social events)

Information

4.1 Conference Venue

The conference is held at HEC Montréal, 3000 chemin de la Côte-Sainte-Catherine, Montreal (see the map on the back cover or on the conference web site). The HEC building is at short walking distance from the *Université de Montréal* subway station; you can use the entrance near the Eastern corner of the building, which is the closest to the subway station.

Detailed information on Montreal's public transportation system can be found at http://www.stm.info. Maps of the entire network are available for free in the metro stations, where you can also buy a weekly pass, called CAM hebdo, giving unlimited access to the STM network (metro and bus) from Monday to Sunday.

If you prefer to bike to the conference you can use Montreal's bike sharing system, BIXI. You can buy a 1-day access, 3-day access or one-way trip. Trips of 30 minutes or less have no additional fees if you purchased short-term access; usage fees apply for longer trips. For more information visit https://bixi.com/.

If you come by car, pay parking at HEC is available from Monday to Friday (6:00 to 22:30), Saturday and Sunday (7:00 to 18:30), at $ 3.50 per half hour and $ 19 for the day. Bicycles can park for free.

Registration, information desk, coffee breaks, and conference rooms

The registration and information desk will be located in front of the *Amphithéâtre Banque Nationale*, on the main floor (rez-de-jardin) of the HEC building (see the Maps section).

A message board will be located near the registration/information desk. It will provide updates to the program and all other last-minute announcements and information. Messages to participants can be left on the message board.

All the plenary sessions will be held in the Amphithéâtre Banque Nationale situated in the blue section. The coffee breaks will be in the room *Investissement Québec*, next to the Amphithéâtre Banque Nationale. One exception: On Friday morning, the coffee break will be next to the rooms of the regular sessions.

The parallel sessions and the tutorials will be in rooms *Banque de développement du Canada, Banque CIBC, Banque Scotia, and Ernst and Young*, all located on the first floor, in the blue section (see the Maps section).

Lunch breaks

The participants are on their own for lunches. The easiest and least expensive solution is to use the HEC Cafeteria, located on the main floor, not far from the registration desk. It will be open from 7:30 to 20:45, from Monday to Friday.

For more fancy food and a nicer setting (including a nice view), the restaurant *Le Cercle*, located on the 7th floor of HEC Montréal (accessible by the elevator on the south side of the building) offers breakfast from 7:30 and lunch until 14:30, from Monday to Friday. Prices are quite reasonable for the quality. However, room is very limited and reservations are required. They can be made only by phone (514-340-7170) or by going in person to the restaurant during opening hours. Group reservations can be taken for a maximum of 20 people only.

Outside the building, there are many restaurants on Côte-des-Neiges street, as well as on Gatineau (south of Lacombe) and on Lacombe (between Gatineau and Côte-des-Neiges). To get there, you have to walk between one and two kilometers each way, and it will be difficult to make it within the 90 minutes time frame, unless you walk fast and get served quickly.

Some suggestions in the Côte-des-Neiges area:

Bistro Olivieri (local food, nicest food in area), 5219 Côte-des-Neiges, enter through book store
Pub McCarold (burgers, beer), 5400 Côte-des-Neiges
Pho Lien (good vietnamese), 5703 Côte-des-Neiges
Caravan Café (salads and sandwiches, relaxed student's hangout), 3506 Lacombe
La Panthère Verte (local and organic food), 3515 Lacombe
Il Galateo (Italian), 5315 Gatineau
Pocha De Marie (Corean), 5349 Gatineau
La Maisonnée (sports bar, beer, burgers, sandwiches, low price), 5385 Gatineau

Internet access

Wireless Internet access will be available in all classrooms, study rooms, public areas and administrative offices. Detailed information on how to connect your laptop to the wireless network can be found at:

http://www.hec.ca/dti/assistance/reseau_internet/page40500.html

The procedure is quite standard.

– *Network to choose*: Évènements-HEC

– *Password*: Forthcoming in the printed version

Links for tourist information

Montreal official tourist site: http://www.tourisme-montreal.org/

Quebec's Government officiel tourist site: http://www.bonjourquebec.com/

Eating in Montreal

Montreal has a rich variety of restaurants of all kinds, with food from probably any country of the world. A few recommendations if you are looking for fine gourmet dining experiences: Bouillon Bilk, Milos, Toqué, Liverpool House, Joe Beef, Les 400 Coups, Le Damas, Juni (sushi), Chez L'Épicier. One notch less expensive: Pastaga, Léméac, Bistro Olivieri.

Climate

The temperature in Montreal is highly variable. Temperatures above 30° C and high humidity are not uncommon in July. On an average July day, the maximum and minimum temperatures are 26° C and 15° C.

4.2 Social events

Wine and Cheese Reception

A wine and cheese reception will take place on Monday July 7, at 17:45, at Salon L'Oréal, located on the main floor, at the South end of the building. All participants, with their partner if any, are invited. Please wear your badge.

Conference Banquet

The conference banquet will be held on Wednesday evening at Pastaga, Le Marché des Éclusiers, 400 rue de la Commune Ouest, in the old port of Montreal. It can be reached easily by subway and a bit of healthy walking.

The banquet is included in the conference registration fee. Registered participants will receive a ticket for it. Those who have requested a vegetarian menu at registration will receive a different type of ticket. Please make sure you have the right type of ticket, because it will be too late to change at the banquet. Non-registered guests may attend the banquet by purchasing a ticket in advance, either when registering over the Internet for $100 Cdn until June 20, or at the information desk until Monday July 3 at the latest, for $120 Cdn.

4.3 Presentations

Instructions for speakers

Plenary talks are 50 minutes plus 10 minutes for questions and discussion. All other talks are 25 minutes, plus 5 minutes for questions and discussion.

Please make sure that you do not exceed your time. Focus on the essential of your message. Given the short time allowed to each speaker, it is generally not possible to give the full details of your work. You should concentrate on providing a clear explanation of your main results and their significance.

The MCM audience comes from a wide variety of backgrounds: there are mathematicians, computer scientists, statisticians, physicists, financial engineers, and so on. The majority of attendees are unlikely to be familiar with the subject of your talk. So you should reserve some time at the beginning to explain your topic from a broad perspective.

The conference covers both theory and practice. If your talk is theory-oriented, it is important to discuss how your findings can eventually help the practitioners. If you are focusing on a specific application, do not forget to point out issues on which you think theoreticians might help, and problems for which general state-of-the-art techniques could eventually improve efficiency. Fruitful exchanges between theoreticians and practitioners are a key objective of this conference.

Instructions for Session Chairs

Session chairs have the responsibility to make sure the speakers adhere tightly to the schedule. Some participants might want to switch between parallel sessions to attend specific talks. To make sure that this can be done smoothly, session chairs should enforce strict adherence to the schedule. We will provide cards to be shown to the speaker for indicating 5, 3, and 1 remaining minutes of speaking time.

Session chairs should also contact their session speakers ahead of time to verify their presence and inform the organizers of any potential no-shows.

Equipment

Each lecture room will be equipped with a computer and a projector for displaying computer output.

4.4 Special Issue of Mathematics and Computers in Simulation

Selected papers covering the topic of a talk presented at MCM 2017 will appear in a special issue of *Mathematics and Computers in Simulation*. Submission will be through the Elsevier electronic submission system, and the call for papers with the submission instructions will be available on the web site shortly after the conference. The deadline for submission will be October 31, 2017.

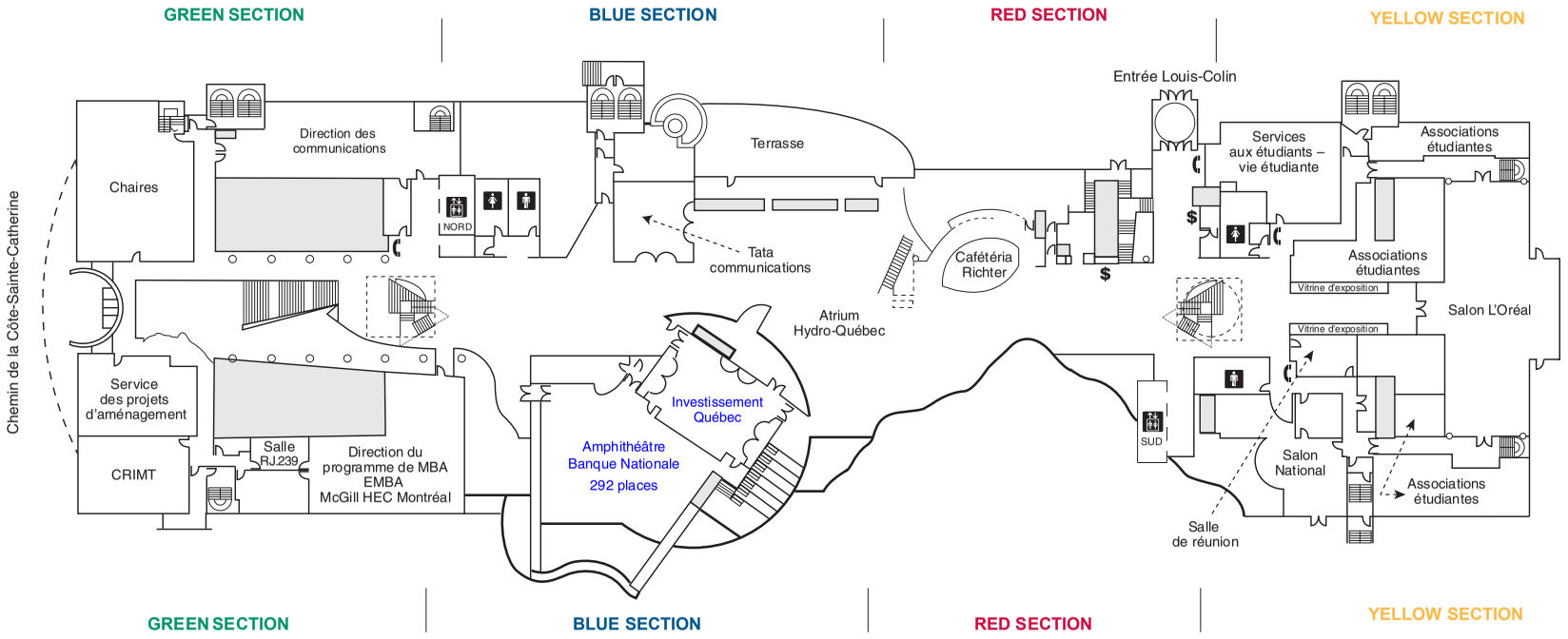
Maps

HEC Montréal
Édifice Côte-Sainte-Cathetine

3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) H3T 2A7

GROUND FLOOR

Faculté de l'aménagement de l'Université de Montréal

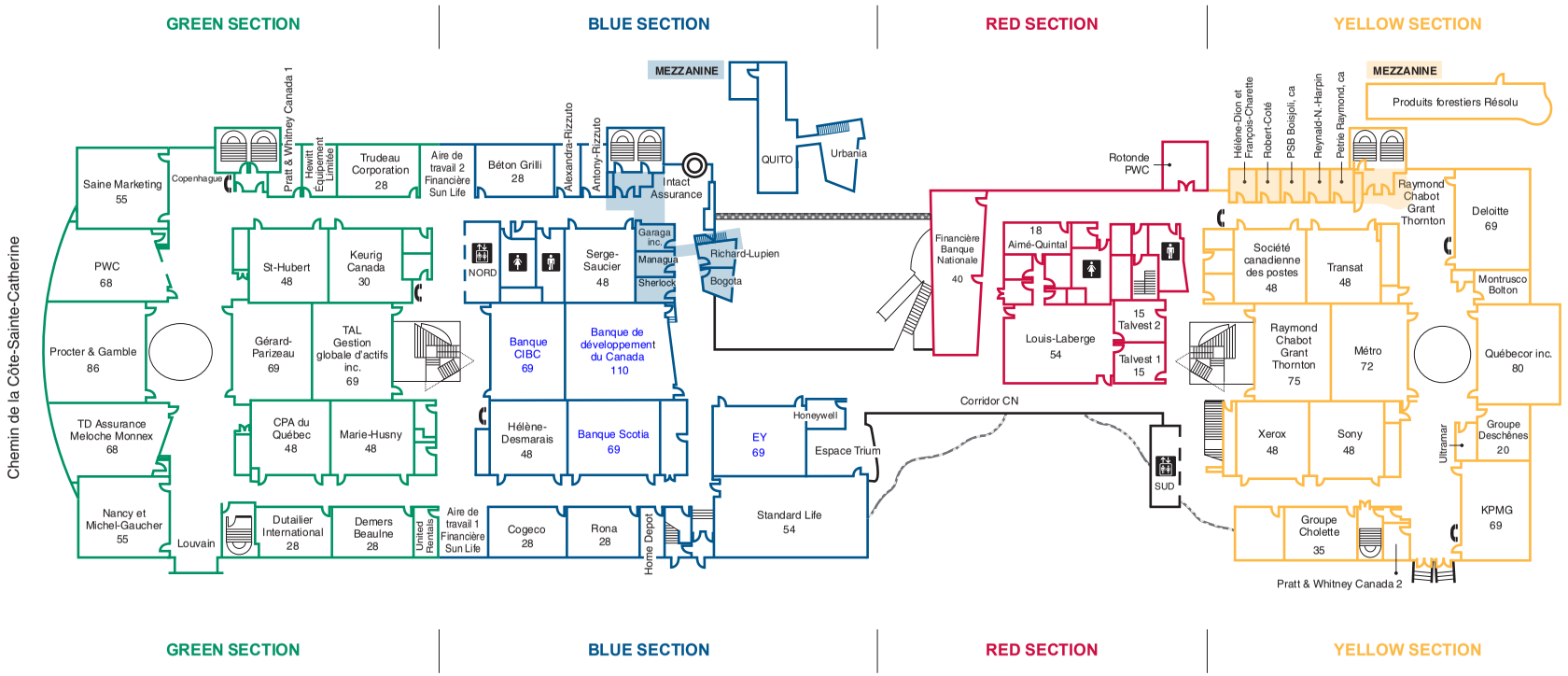


Collège Jean-de-Brébeuf

HEC Montréal
Édifice Côte-Sainte-Catherine

3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) H3T 2A7

1ST FLOOR



Le chiffre inscrit en dessous du nom de la salle indique le nombre de places disponibles.

SDG-6847 - Juillet 2016

Schedule

Special sessions are in green and contributed sessions are in blue.

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

Monday morning, July 3

8^{00}	**Registration opens**
$8^{40} - 9^{00}$	**Opening Session** (Amphithéâtre Banque Nationale) *Pierre L'Ecuyer*
$9^{00} - 10^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Alexander Keller* **Machine Learning and Integral Equations** Chair: *Pierre L'Ecuyer* *p. 30*
$10^{00} - 10^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Stochastic Computation and Complexity I** Chair: *Stefan Heinrich* *p. 40*	Room: Banque CIBC **Stochastic Gradient methods for Monte Carlo and Variational Inference** Chair: *Víctor Elvira* *p. 42*	Room: Banque Scotia **Markov Chain Monte Carlo I** Chair: *Jeff Rosenthal* *p. 44*	Room: Ernst & Young **Stochastic Differential Equations** Chair: *Christian Lécot* *p. 47*
$10^{30} - 11^{00}$	*Raphael Kruse* Error Analysis of Some Randomized Runge-Kutta Methods for ODEs with Time-Irregular Coefficients	*Rajesh Ranganath* Variational Inference: Modern Methods	*Radu V. Craiu* Adaptive Component-Wise Multiple-Try Metropolis Sampling	*Itsaso Hermosilla* Simulating a Stochastic Differential Equation Model by Exact Sampling
$11^{00} - 11^{30}$	*Timo Welti* Convergence in Hölder Norms with Applications to Monte Carlo Methods in Infinite Dimensions	*Qiang Liu* On Stein's Method for Practical Statistical Computation	*Mylène Bédard* Hierarchical Models: Local Proposal Variances for RWM- and MALA-within-Gibbs	*Andreas Thalhammer* Importance Sampling Techniques for Stochastic Partial Differential Equations
$11^{30} - 12^{00}$	*Sotirios Sabanis* Recursive Estimators and MCMC Algorithms	*Ernest K. Ryu* Control Variates, Importance Sampling, and Biconvexity	*Jordan Franks* Importance Sampling versus Delayed Acceptance MCMC when Noisy Approximations are Available	*Samuel Herrmann* Initial-Boundary Value Problem for the Heat Equation—A Stochastic Algorithm
$12^{00} - 12^{30}$	*Jim Calvin* Bounds on the Number of Function Evaluations to Approximate the Global Minimum	*Ingmar Schuster* Adaptive Monte Carlo from the Stochastic Optimization Perspective	*Florian Maire* Locally Informed Adaptive MCMC Algorithm Based on Online PCA	

$12^{30} - 14^{00}$	Lunch break

Monday afternoon, July 3

$14^{00}-15^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Steffen Dereich, WWU Münster, Germany* **Multilevel Adaptations for Stochastic Approximation Algorithms** Chair: *Mike Giles* *p. 31*
$15^{00}-15^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Tractability** Chair: *Peter Kritzer* *p. 50*	Room: Banque CIBC **Quasi-Monte Carlo, Sequential Monte Carlo and Applications in Machine Learning** Chair: *Simon Lacoste-Julien* *p. 52*	Room: Banque Scotia **Markov Chain Monte Carlo II** Chair: *Mylène Bédard* *p. 53*	Room: Ernst & Young **Partial Differential Equations and Random Fields** Chair: *Christian Lécot* *p. 58*
$15^{30}-16^{00}$	*Ian H. Sloan* Geometric Brownian Motion by Brownian Bridge Construction—How Fast is the Uniform Convergence?	*Fredrik Lindsten* Divide-and-Conquer with Sequential Monte Carlo	*Jackson Gorham* Measuring Sample Quality with Kernels	*Pieterjan Robbe* Dimension-Adaptive Multi-Index Monte Carlo for PDEs with Random Coefficients
$16^{00}-16^{30}$	*Henryk Woźniakowski* On Exponential Convergence of Multivariate Problems	*Mathieu Gerber* Sequential Quasi-Monte Carlo	*Oren Mangoubi* Rapid Mixing Bounds for Hamiltonian Monte Carlo on Strongly Log-concave Distributions	*Michael Feischl* Multi-Index Quasi-Monte Carlo and $\mathcal{H}$-Matrices
$16^{30}-17^{00}$	*Erich Novak* Input Sets for Numerical Integration	*Simon Lacoste-Julien* Optimization Tools for Adaptive Monte-Carlo Integration	*Magdalena Strauss* New MCMC Methods for the Ordering and Clustering of Single-Cell Data	
$17^{00}-17^{30}$	*Josef Dick* Quasi-Monte Carlo Methods and PDEs with Random Coefficients	*Deborshee Sen* Communication Efficient Sequential Monte Carlo	*Maksym Byshkin* Efficient Markov Chain Monte Carlo Estimation of Exponential-Family Random Graph Models	

$17^{45}-20^{00}$	Cocktail Salon L'Oréal

Tuesday morning, July 4

$9^{00} - 10^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Emmanuel Gobet* **Recent Advances in Regression Monte-Carlo Methods** Chair: *Christian Robert* *p. 32*
$10^{00} - 10^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Stochastic Computation and Complexity II** Chair: *Thomas Müller-Gronbach* *59*	Room: Banque CIBC **Multilevel Monte Carlo–Theory, Inference and Nested Simulation** Chair: *Mike Giles* *p. 61*	Room: Banque Scotia **Monte Carlo and Quasi-Monte Carlo in Computer Graphics** Chair: *Derek Nowrouzezahrai* *p. 62*	Room: Ernst & Young **Random Number and Random Variate Generation** Chair: *Hiroshi Haramoto* *p. 68*
$10^{30} - 11^{00}$	*Henryk Woźniakowski* Multivariate Approximation for Analytic Functions with Gaussian Kernels	*Kody Law* Multilevel Monte Carlo for Bayesian Inference	*Hélène Perrier* Towards Low-Discrepancy Sequences with Improved Spectral Properties: Applications in Computer Graphics	*Paula Whitlock* The Effect of Pseudorandom Number Generators on the Convergence of Multidimensional Random Walks
$11^{00} - 11^{30}$	*Paweł Przybyłowicz* Lower Bounds for Strong Global Approximation of Solutions of SDEs Under Adaptive Information About Additive Poisson Noise	*Gilles Pagès* Limit Theorems for Weighted or Unweighted Multilevel Estimators: Applications and Comparisons	*Jaroslav Krivanek* Zero-Variance-Based Sampling Schemes in Light Transport Simulation	*Shin Harase* On the Concatenation of Mersenne Twisters
$11^{30} - 12^{00}$	*Taras Shalaiko* Discretization of Occupation Times for Fractional Diffusions	*Abdul-Lateef Haji-Ali* MLMC for Value-at-Risk	*Iliyan Georgiev* Blue-Noise Dithered Sampling for Image Synthesis	*Josef Leydold* A Note on Generating Random Variables with T-concave Densities with the Ratio-of-Uniforms Method
$12^{00} - 12^{30}$	*Thomas Kühn* Complexity of High-Dimensional Approximation in Periodic Function Spaces	*Mike Giles* MLMC for Estimation of Expected Value of Partial Perfect Information	*Ricardo Marques* Bayesian Monte Carlo Spherical Integration for Illumination Integrals	*Aaditya Ramdas* Online Generation of Low-Discrepancy Sequences: From Monte-Carlo to Quasi-Monte-Carlo by Retrying

$12^{30} - 14^{00}$	Lunch break

Tuesday afternoon, July 4

$14^{00}-15^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Christian P. Robert* **Approximate Bayesian Computation: From Convergence Guarantees to Automated Implementation** Chair: *Art B. Owen* *p. 33*
$15^{00}-15^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Sequences for Quasi-Monte Carlo** Chair: *Wolfgang Ch. Schmid* *p. 70*	Room: Banque CIBC **Multilevel Monte Carlo for Stochastic Differential Equations** Chair: *Abdul-Lateef Haji-Ali* *p. 72*	Room: Banque Scotia **Function Approximation and Variance Reduction** Chair: *Henryk Wozniakowski* *p. 75*	Room: Ernst & Young **Simulation of Stochastic Processes** Chair: *Geneviève Gauthier* *p. 78*
$15^{30}-16^{00}$	*Christiane Lemieux* Implementations and Tests for Irreducible Sobol' Sequences	*Fan Zhang* Exact Simulation of Multivariate Itô Diffusions	*Robert J. Kunsch* High-Dimensional Function Approximation—Breaking the Curse with Monte Carlo Methods	*Michael Chiu* Simulation of Multivariate Mixed Poisson Processes
$16^{00}-16^{30}$	*Friedrich Pillichshammer* Metrical Star Discrepancy Bounds for Lacunary Subsequences of Digital Kronecker-Sequences and Polynomial Tractability	*Andreas Stein* An Adaptive Multilevel Monte Carlo Algorithm for Elliptic PDEs with Jump Diffusion Coefficient	*David Barrera* Least Squares Regression for Non-Stationary Designs	*Till Massing* Simulation of Student-Lévy Processes Using Series Representations
$16^{30}-17^{00}$	*Takashi Goda* Optimal Order Quasi-Monte Carlo Integration for Smooth Functions	*Lukas Herrmann* MLQMC with Product Weights for Elliptic PDEs with Lognormal Coefficients Parametrized in Multiresolution Representations	*Massil Achab* SGD with Variance Reduction Beyond Empirical Risk Minimization	*Krzysztof Bisewski* Minimizing Time Discretization Error
$17^{00}-17^{30}$	*Florian Pausinger* Pair Correlations and Equidistribution	*Zhenru Wang* Analysis of Multi-Index Monte Carlo Estimators for a Zakai SPDE	*Alexander Shkolnik* Compactness Approaches for Importance Sampling	*Yiwei Wang* Unbiased Simulation and Parameters Estimation of Distributions with Explicitly Known Fourier Transforms

Wednesday morning, July 5

$9^{00} - 10^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Paul Dupuis* **Competing Sources of Variance Reduction in Parallel Replica Monte Carlo, and Optimization in the Low Temperature Limit** Chair: *Dirk Nuyens* *p. 34*
$10^{00} - 10^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Stochastic Computation and Complexity III** Chair: *Michaela Szölgyenyi* *p. 80*	Room: Banque CIBC **Multivariate Decomposition Methods and Truncation Algorithms** Chair: *Peter Kritzer* *p. 83*	Room: Banque Scotia **Stochastic Optimization** Chair: *Fabian Bastin* *p. 85*	Room: Ernst & Young **Goodness-of-Fit Tests** Chair: *Paula Whitlock* *p. 88*
$10^{30} - 11^{00}$	*Thomas Müller-Gronbach* Lower Error Bounds for Strong Approximation of Scalar SDEs with Non-Lipschitzian Coefficients	*Alexander Gilbert* A New Construction of Active Sets for the Multivariate Decomposition Method	*Uday V. Shanbhag* Variable Sample-Size Stochastic Approximation with Finite Sampling Budget	*Leonard Santana* A Data-Dependent Choice of the Tuning Parameter for Goodness-of-Fit Tests Employing Bootstrapped Critical Values
$11^{00} - 11^{30}$	*Dai Taguchi* On the Euler-Maruyama Scheme for SDEs with Discontinuous Diffusion Coefficient	*Dirk Nuyens* Integration Over $\mathbb{R}^{\mathbf{N}}$ Using the Multivariate Decomposition Method and Higher-Order QMC Rules	*Sanjay Dominik Jena* Tractable Models for Satisficing under Uncertainty	*James S. Allison* A Monte Carlo Evaluation of the Performance of Two New Tests for Symmetry Based on the Empirical Characteristic Function
$11^{30} - 12^{00}$	*Andreas Neuenkirch* The Euler Scheme for SDEs with Discontinuous Drift Coefficient: A Numerical Study of the Convergence Rate	*Peter Kritzer* Truncation Dimension for Linear Problems on Weighted Anchored and ANOVA Spaces	*Joseph L. Durante* Scenario Generation Methods that Replicate Crossing Times in Spatially Distributed Stochastic Systems	*Simon Mak* Support Points—A New Way to Compact Distributions
$12^{00} - 12^{30}$	*Antoine Lejay* Designing and Benchmarking Monte Carlo Methods for Simulating Processes in Discontinuous Media	*Yoshihito Kazashi* Discrete Maximal Regularity and Discrete Error Estimate of a Non-Uniform Implicit Euler–Maruyama Scheme for a Class of Stochastic Evolution Equations	*Julien Keutchayan* A New Framework for Generating Scenario Trees Using Quasi-Monte Carlo Methods	*Hiroshi Haramoto* Testing Soundness of Statistical Tests for Random Number Generators by Using a Three-Level Test

$12^{30} - 14^{00}$	Lunch break

Wednesday afternoon, July 5

$14^{00}-15^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Gunther Leobacher, KFU Graz, Austria* **Two Strongly Convergent Methods for Stochastic Differential Equations with Irregular Coefficients** Chair: *Ivan Dimov* *p. 35*
$15^{00}-15^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Sobol' Indices and Sobol' Sequences** Chair: *Christiane Lemieux* *p. 91*	Room: Banque CIBC **Variance Reduction for Rare-Event Simulation** Chair: *Zdravko I. Botev* *p. 94*	Room: Banque Scotia **Markov Chain Monte Carlo in Bayesian Statistics** Chair: *Radu V. Craiu* *p. 96*	Room: Ernst & Young **Monte Carlo in Particle and Quantum Physics** Chair: *Erich Novak* *p. 98*
$15^{30}-16^{00}$	*Sergei Kucherenko* Sobol' Indices for Constrained Global Sensitivity Analysis	*Hui (Alice) Yao* Rare-Event Simulation for Products of Random Variables	*Matthias Kaeding* Fast, Approximate MCMC for Bayesian Analysis of Large Data Sets: A Design Based Approach	*Ivan T. Dimov* Convergence Probabilistic Analysis of the Monte Carlo Method for Quantum Physics Problems
$16^{00}-16^{30}$	*Lluís Antoni Jiménez Rugama* Automatic Estimation of First-Order Sobol' Indices Using the Replication Procedure	*Nadhir Ben Rached* On the Efficient Simulation of the Left-Tail of the Sum of Correlated Log-Normal Variates	*Christian A. Naesseth* Variational Sequential Monte Carlo	*Natalya Tracheva* A New Monte Carlo Method for Estimation of Time Asymptotic Parameters of Polarized Radiation
$16^{30}-17^{00}$	*Christiane Lemieux* Importance Sampling Techniques for Semi-Parametric Single Index Models	*Zdravko I. Botev* Practical Estimators for the Distribution of the Sum of Correlated Log-Normal Random Variables	*Philippe Gagnon* Weak Convergence and Optimisation of the Reversible Jump Algorithm	*Danhua Shangguan* A Shannon Entropy-Based Strategy for Adjusting History Number of Time-Dependent Transport Problem Automatically

$19^{00}-22^{00}$	Conference Banquet at Le Marché des Éclusiers, in Old Port

Thursday morning, July 6

$9^{00} - 10^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Aicke Hinrichs, Johannes Kepler University, Linz, Austria* **The Complexity of High and Infinite Dimensional Integration** Chair: *Stefan Heinrich* *p. 36*
$10^{00} - 10^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Stochastic Computation and Complexity IV** Chair: *Pawel Przybylowicz* *p. 100*	Room: Banque CIBC **Monte Carlo Methods for Molecular Evolution and Phylogenetics** Chair: *Liangliang Wang* *p. 102*	Room: Banque Scotia **Construction of QMC Point Sets** Chair: *Friedrich Pillichshammer* *104*	Room: Ernst & Young **Applications in Workforce Planning and Scheduling** Chair: *Vassil Alexandrov* *p. 106*
$10^{30} - 11^{00}$	*Fred J. Hickernell* Deterministic, Randomized, and Bayesian Ways to Stop a Simulation	*Simon Laurin-Lemay* Phylogenetic Modeling of CpG Hypermutability Using Exact and Approximate Bayesian Computation	*Mario Neumüller* A Reduced Fast Construction of Polynomial Lattice Point Sets with Low Weighted Star Discrepancy	*Stefka Fidanova* Monte Carlo Method for Workforce Planning
$11^{00} - 11^{30}$	*Wei Fang* Adaptive Multilevel Monte Carlo for Ergodic SDEs with Non-Globally Lipschitz Drift in Infinite Time Interval	*Vu Dinh* Online Bayesian Phylogenetic Inference via Sequential Monte Carlo	*Kosuke Suzuki* Enumeration of the Chebyshev-Frolov Lattice Points in Axis-Parallel Boxes	*Wael Sboui* Evaluation of Aircraft Landing Scheduling Policies Using Monte Carlo Simulation
$11^{30} - 12^{00}$	*Holger Stroot* Strong Approximation of Stochastic Mechanical Systems with Nonlinear Holonomic Constraints	*Nicolas Rodrigue* Parallelizable Monte Carlo Algorithms for Infinite Mixtures in the Detection of Molecular Adaptation	*Ralph Kritzinger* Optimal Discrepancy Rate of Point Sets in Besov Spaces with Negative Smoothness	*Thuy Anh Ta* On the Sample Average Approximation of a Two-Stage Staffing Problem with Chance Constraints and Recourse in Call Centers
$12^{00} - 12^{30}$	*André Herzwurm* Optimal Strong Approximation of Cox-Ingersoll-Ross and Squared Bessel Processes	*Liangliang Wang* A Sequential Monte Carlo Algorithm for Bayesian Phylogenetics	*Takehito Yoshiki* Richardson Extrapolation of Polynomial Lattice Rules	

$12^{30} - 14^{00}$	Lunch break

Thursday afternoon, July 6

$14^{00}-15^{00}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Mark Girolami* **Diffusions and Dynamics on Statistical Manifolds for Monte Carlo Statistical Inference** Chair: *Simon Lacoste-Julien* *p. 37*
$15^{00}-15^{30}$	Coffee break

	Room: Banque Dévelop. Canada **Stochastic Computation and Complexity V** Chair: *Andreas Neuenkirch* *p. 108*	Room: Banque CIBC **Statistical Applications of Monte Carlo Methods** Chair: *Fred J. Hickernell* *p. 110*	Room: Banque Scotia **QMC in Special Settings and Applications** Chair: *Mathieu Gerber* *p. 113*	Room: Ernst & Young **Applications in Business and Operations Management** Chair: *Johan Van Kerckhoven* *116*
$15^{30}-16^{00}$	*Stefan Heinrich* Multilevel Algorithms for Banach Space Valued and Parametric Stochastic Differential Equations	*Mark Huber* Faster Estimates with User-Specified Error for $[0,1]$ Random Variables	*Christian Lécot* Quasi-Monte Carlo Simulation of Coagulation and Fragmentation	*Stephan Prell* A Sequential Design for Gas Storage Optimization Using Kriging Metamodels
$16^{00}-16^{30}$	*Stefan Geiss* Continuous Time Tug-of-War, p-Harmonic Functions, and an Approximation Problem	*Simon Mak* Projected Support Points, with Application to Optimal MCMC Reduction	*Yuya Suzuki* Quasi-Monte Carlo for Calculating Multi-Dimensional Expectations Against Gaussian Distribution	*Laura Hervert-Escobar* Solving the Territorial Design Problem for Business Sales Plan Using Monte Carlo Method
$16^{30}-17^{00}$	*Michaela Szölgyenyi* Optimal Liquidation Under Partial Information with Price Impact	*Kan Zhang* An Adaptive Quasi-Monte Carlo Method for Bayesian Inference with User-Specified Error Tolerance	*Javier Gonzalez-Villa* A New Rotation Invariant Sampling Design on the Sphere	*Nathan Yang* Better Together? Performance Dynamics in Retail Chain Expansion Before and After Mergers
$17^{00}-17^{30}$	*David Krieg* On the Approximation of Tensor Product Operators	*Mark Girolami* Convergence Rates of Control Functional Estimators Based on Stein's Identity	*Amal Ben Abdellah* Density Estimation by Randomized Quasi-Monte Carlo	*Eric Torkia* Networking Simulation Results Across Organizational Boundaries

Friday morning, July 7

	Room: Banque Dévelop. Canada **Accelerated Monte Carlo in Optimization, Statistics, and PDEs with Random Input** Chair: *Raghu Pasupathy* *p. 118*	Room: Banque CIBC **Advanced Monte Carlo Methods in Non-Linear Finance** Chair: *Emmanuel Gobet* *p. 120*	Room: Banque Scotia **Monte Carlo Simulation Using SimJulia** Chair: *Fabian Bastin* *p. 122*	Room: Ernst & Young **Parallel Computation and Codes** Chair: *Ivan Dimov* *p. 125*
$9^{00} - 9^{30}$	*Jeffrey Rosenthal* Adaptive MCMC For Everyone	*Gersende Fort* MCMC Design-Based Non-Parametric Regression for Rare-Event: Application to Nested Risk Computations	*Nicolas Andriessen* SimJulia: The Good, the Bad and the Ugly	*Vassil Alexandrov* On Efficient Parallel Monte Carlo and Quasi-Monte Carlo Hybrid Methods for Matrix Computations
$9^{30} - 10^{00}$	*Raghu Pasupathy* The Adaptive Sampling Gradient Method: Optimizing Smooth Functions with an Inexact Oracle	*Mike Ludkovski* Sequential Design for Estimating Value-at-Risk	*Johan Van Kerckhoven* Simulation Based Manpower Planning: An Introduction Using SimJulia	*Sofiya Ivanovska* Optimal Implementation of Quasi-Monte Carlo Methods for Matrix Computations on Intel MIC Architecture
$10^{00} - 10^{30}$	*Xiaoou Li* Rare Event Analysis and Efficient Simulation for Random Elliptic PDEs with Small Noise	*Isaque Pimentel* Hedging with Non-Quadratic Local Risk Minimization Using Least-Squares Monte Carlo	*Selma Koghee* Simulation of Medical Response to Disasters Using SimJulia	*Gang Li* JMCT: A 3D Monte Carlo Particle Transport Code
$10^{30} - 10^{50}$	Coffee break			

Friday morning, July 7

	Room: Banque Dévelop. Canada **Monte Carlo Simulation and its Applications in Stochastic Dynamic Programming** Chair: *Raghu Pasupathy* *p. 127*	Room: Banque CIBC **Finance Applications** Chair: *Gersende Fort* *p. 128*
$10^{50}-11^{20}$	*Ma Xiang* A Primal-Dual Iterative Monte Carlo Method for Stochastic Dynamic Programs and Its Applications in Finance	*Andrei Cozma* Calibration and Monte Carlo Pricing Under a Hybrid Local-Stochastic Volatility Model
$11^{20}-11^{50}$	*John R. Birge* MCMC Methods for Dynamic Stochastic Optimization	*Geneviève Gauthier* Extracting Latent States from High Frequency Option Prices

$11^{55}-12^{55}$	**Invited Plenary Talk** (Amphithéâtre Banque Nationale) *Art B. Owen, Stanford University, USA* **Quasi-Monte Carlo, Beyond the Unit Cube** Chair: *Frances Kuo* *p. 38*
$12^{55}-13^{00}$	Closing words and end of the conference

Abstracts of Plenary Talks

Monday 9:00–10:00 **Room: Amphi. Banque Nationale**

Machine Learning and Integral Equations

Alexander Keller

NVIDIA

akeller@nvidia.com

As both light transport simulation and reinforcement learning are ruled by the same Fredholm integral equation of the second kind, machine learning techniques can be used for efficient photorealistic image synthesis: Light transport paths are guided by an approximate solution to the integral equation that is learned during rendering [1]. We investigate the application of artificial neural networks to represent this approximate solution in the context of Monte Carlo and quasi-Monte Carlo methods [2, 3] to compute functionals of integral equations.

This is joint work with Ken Dahm.

[1] K. Dahm and A. Keller. Learning light transport the reinforced way. *CoRR*, abs/1701.07403, 2017.

[2] A. Keller. Quasi-Monte Carlo image synthesis in a nutshell. In J. Dick, F. Kuo, G. Peters, and I. Sloan, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2012*, 203–238. Springer, 2013.

[3] A. Keller, K. Dahm, and N. Binder. Path Space Filtering. In R. Cools and D. Nuyens, editors, *Monte Carlo and Quasi-Monte Carlo Methods 2014*, 423–436. Springer, 2016.

Monday 14:00–15:00 **Room: Amphi. Banque Nationale**

Multilevel Adaptations for Stochastic Approximation Algorithms

Steffen Dereich

Westfälische Wilhelms-Universität Münster, Universität Passau

steffen.dereich@wwu.de

In 1951 Robbins and Monro introduced a probabilistic method to compute zeroes of functions $f : \mathbb{R}^d \to \mathbb{R}^d$ that are given in terms of expectations

$$f(\theta) = \mathbb{E}[F(\theta, U)]$$

with U being a random variable attaining values in an arbitrary measurable space and F being a product-measurable $\mathbb{R}^d$-valued function such that the expectation is well-defined. Based on random samples of the random variable $F(\theta, U)$ a stochastic dynamical system was devised that converges to a zero under appropriate contractivity assumptions.

Originally, Robbins and Monro used the *stochastic approximation method* for the computation of quantiles and for solving regression problems. Since the original work, the concept proved to be very useful in various branches of statistics and research remained active until now.

In this talk we focus on the case where $F(\theta, U)$ is not simulatable in which case one relies on approximation. We review previous research and present a new complexity theorem for multilevel stochastic approximation algorithms that is similar to the classical one in Giles (2008). Our approach is universal in the sense that having classical multilevel implementations for a particular application at hand it is straightforward to implement the corresponding stochastic approximation algorithm. Moreover, previous research on multilevel Monte Carlo can be incorporated in a natural way.

This is joint work with Thomas Müller-Gronbach.

Tuesday 9:00–10:00 **Room: Amphi. Banque Nationale**

Recent Advances in Regression Monte-Carlo Methods

Emmanuel Gobet

CMAP, Ecole Polytechnique
and
CNRS, Université Paris Saclay, France

http://www.cmap.polytechnique.fr/~gobet/

Regression Monte-Carlo methods make use of stochastic simulations to evaluate conditional expectation functions. For example, these methods lead to robust numerical schemes for solving non-linear PDEs through Feynman-Kac formulas (like for semi-linear (Stochastic) PDEs or Hamilton-Jacobi-Belman type equations), by computing a sequence of coupled conditional expectations involving some stochastic processes. They are also effective for a variety of problems where one has to design nested simulation algorithms.

In the last 15 years there has been an increasing interest for these techniques, which combine tools from machine learning, stochastic analysis, and which are quite fruitful for applications.

In this talk, I will review recent advances and applications of these approaches: high-dimensional non-linear PDEs, parallel computing (despite coupling non-linearity), nested extreme risks, coupling with MCMC scheme, model-free learning and data-driven scheme. Applications to finance, insurance, population dynamics, smart-grids among others will be presented.

Tuesday 14:00–15:00 **Room: Amphi. Banque Nationale**

Approximate Bayesian Computation: From Convergence Guarantees to Automated Implementation

Christian P. Robert

Université Paris Dauphine,
PSL Research University,
and
University of Warwick

bayesianstatistics@gmail.com

Approximate Bayesian Computation (ABC) has grown into a standard methodology to handle Bayesian inference in models associated with intractable likelihood functions. Due to its approximative nature, it however requires strict constraints its features to ensure consistency, including a connection between the statistics summarising the data and the speed of convergence of the tolerance parameter to zero. [3] and [4] study the asymptotic behaviour of the posterior obtained from ABC and of the ensuing posterior mean, including the asymptotic distribution of the ABC posterior mean. Important implications of the theoretical results for practitioners of ABC will be highlighted.

Since ABC realistic implementations require the selection of a non-sufficient statistic, whose dimension impacts both the convergence properties of the algorithm and the computing requirements, in contrast with [2], [5] propose a version of ABC that both does not impose a selection of the most relevant components and bypasses the derivation of a tolerance. The approach relies on the random forest methodology of [1] when applied to regression. When compared with standard ABC solutions, this technology offers significant gains in terms of robustness to the choice of the summary statistics and of computing time.

[1] L. Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[2] P. Fearnhead and D. Prangle. Constructing summary statistics for Approximate Bayesian Computation: semi-automatic Approximate Bayesian Computation. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 74(3):419–474, 2012.

[3] D. Frazier, G. Martin, C.P. Robert, and J. Rousseau Asymptotic Properties of Approximate Bayesian Computation *arXiv*, 1607.06903, 2016.

[4] M. Li and P. Fearnhead. On the asymptotic efficiency of ABC estimators. *arXiv*, 1506.03481, 2015.

[5] J.-M. Marin, L. Raynal, P. Pudlo, M. Ribatet and C.P. Robert ABC random forests for Bayesian parameter inference *arXiv*, 1605.05537, 2016

Wednesday 9:00–10:00 **Room: Amphi. Banque Nationale**

Competing Sources of Variance Reduction in Parallel Replica Monte Carlo, and Optimization in the Low Temperature Limit

Paul Dupuis

Division of Applied Mathematics, Brown University, USA

dupuis@dam.brown.edu

Computational methods such as parallel tempering and replica exchange are designed to speed convergence of more slowly converging Markov processes (corresponding to lower temperatures for models from the physical sciences), by coupling them through a Metropolis type swap mechanism with higher temperature processes that explore the state space more quickly. It has been shown that the sampling properties are in a certain sense optimized by letting the swap rate tend to infinity. This "infinite swapping limit" can be realized in terms of a process which evolves using a symmetrized version of the original dynamics, and then one produces approximations to the original problem by using a weighted empirical measure. The weights are needed to transform samples obtained under the symmetrized dynamics into distributionally correct samples for the original problem.

After reviewing the construction of the infinite swapping limit, we focus on the sources of variance reduction which follow from this construction. As will be discussed, some variance reduction follows from a lowering of energy barriers and consequent improved communication properties. A second and less obvious source of variance reduction is due to the weights used in the weighted empirical measure that appropriately transform the samples of the symmetrized process. These weights are analogous to the likelihood ratios that appear in importance sampling, and play much the same role in reducing the overall variance. A key question in the design of the algorithms is how to choose the ratios of the higher temperatures to the lowest one. As we will discuss, the two variance reduction mechanisms respond in opposite ways to changes in these ratios. One can characterize in precise terms the tradeoff and explicitly identify the optimal temperature selection for certain models when the lowest temperature is sent to zero, i.e., when sampling is most difficult.

Wednesday 14:00–15:00 **Room: Amphi. Banque Nationale**

Two Strongly Convergent Methods for Stochastic Differential Equations with Irregular Coefficients

Gunther Leobacher

KFU Graz

https://homepage.uni-graz.at/de/guntherleobacher

We consider multidimensional stochastic differential equations (SDEs) with drift that may be discontinuous and diffusion that may be degenerate. The classical theorems on existence and uniqueness as well as the convergence result by Maruyama require Lipschitz continuity of the coefficients. Generalizations exist where Lipschitz continuity is relaxed, but then uniform ellipticity of the diffusion coefficient is essential.

In this talk I will present a concept of piecewise Lipschitz drift where the set of discontinuities is a sufficiently smooth hypersurface in the $\mathbb{R}^d$. This is a situation that frequently appears in optimal control problems.

We show that for existence and uniqueness, as well as numerical approximation of the solution, uniform ellipticity can be relaxed to a geometric "non-parallelity" condition.

We present a numerical method with strong order 1/2 convergence and show strong order 1/5 for the Euler-Maruyama method. The results rely on a transform of the $\mathbb{R}^d$ that maps the SDE to another one with Lipschitz continuous coefficients.

This is joint work with Michaela Szölgyenyi.

Thursday 9:00–10:00 **Room: Amphi. Banque Nationale**

The Complexity of High and Infinite Dimensional Integration

Aicke Hinrichs

Johannes Kepler University, Linz

aicke.hinrichs@jku.at

We present recent results on the complexity of integration. We focus on high and infinite dimensional settings. In particular, we explain how embedding theorems between certain scales of function spaces can be used to transfer complexity results between different settings. Examples of spaces that can be treated are tensor product spaces of Korobov type spaces of increasing smoothness. This approach is equally useful for approximation problems.

Thursday 14:00–15:00 **Room: Amphi. Banque Nationale**

Diffusions and Dynamics on Statistical Manifolds for Monte Carlo Statistical Inference

Mark Girolami

Department of Mathematics, Imperial College London
and
The Alan Turing Institute

m.girolami@imperial.ac.uk

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

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

Friday 11:55–12:55 **Room: Amphi. Banque Nationale**

Quasi-Monte Carlo, Beyond the Unit Cube

Art B. Owen

Stanford University

statweb.stanford.edu/~owen

Quasi-Monte Carlo (QMC) sampling is usually studied with n points arranged nearly uniformly in $[0,1]^d$. There are many techniques to transform samples from the $U[0,1]^d$ distribution into some other distribution on a potentially different space of possibly different dimension. The book by Devroye (1986) is still the definitive reference after more than 30 years. A common practice in QMC is to apply those techniques to low discrepancy point sets from the unit cube. Some of those transformations that work well for Monte Carlo disrupt the smoothness that QMC exploits when it improves on plain Monte Carlo. This talk looks into a set of methods for sampling from triangles, simplices, disks, spheres, spherical triangles, and Cartesian products of such sets. It includes:

- the triangle discrepancy of Brandolini, Colzani, Gigante and Travaglini (2013),
- a van der Corput construction in the triangle that attains discrepancy $O(n^{-1/2})$,
- a Kronecker construction in the triangle that attains the optimal rate of discrepancy $O(\log(n)/n)$,
- a scrambled van der Corput construction that attains MSE $o(n^{-1})$ (given sufficient smoothness),
- generalizations to s-fold products of d-dimensional sets, and some scrambled geometric nets for those spaces,
- a central limit theorem for those generalization (Basu and Mukherjee, 2016),
- a finding that the Fang and Wang (1993) transformations are smooth enough for the QMC rate but not smooth enough for the RQMC rate, and
- an RQMC friendly approach for simplices, which comes with a dimension effect.

This is joint work with Kinjal Basu.

Abstracts of Regular Talks

Monday 10:30–12:30, **Room: Banque de Développement du Canada**

Stochastic Computation and Complexity I

Chair: Stefan Heinrich

Error Analysis of Some Randomized Runge-Kutta Methods for ODEs with Time-Irregular Coefficients

Raphael Kruse

Technische Universität Berlin, Institut für Mathematik, Secr. MA 5-3, Berlin

kruse@math.tu-berlin.de
http://www.tu-berlin.de/?nwg_uq

In this talk we present some recent results on the error analysis of two randomized explicit Runge-Kutta schemes, that are applicable to the numerical solution of ordinary differential equations (ODEs) with time-irregular coefficients. In particular, the methods are applicable to ODEs of Carathéodory type, whose coefficient functions are only integrable with respect to the time variable but are not assumed to be continuous. A further field of application are ODEs with unbounded coefficient functions that contain weak singularities with respect to the time variable.

The main result consists of precise bounds for the discretization error with respect to the norm in $L^p(\Omega)$. In addition, convergence rates are also derived in the almost sure sense. An important ingredient in the analysis are corresponding error bounds for the randomized Riemann sum quadrature rule. The theoretical results are illustrated through a few numerical experiments.

This is joint work with Yue Wu.

[1] R. Kruse and Y. Wu. Error analysis of randomized Runge-Kutta methods for differential equations with time-irregular coefficients. *Comput. Methods Appl. Math.*, 2017. (to appear).

Convergence in Hölder Norms with Applications to Monte Carlo Methods in Infinite Dimensions

Timo Welti

Seminar for Applied Mathematics, ETH Zürich, Switzerland

timo.welti@sam.math.ethz.ch
https://people.math.ethz.ch/~weltit

This talk concerns convergence rates for general stochastic processes in Hölder norms. It is shown that if a sequence of piecewise affine linear stochastic processes converges in the strong sense with a positive rate to a stochastic process which is strongly Hölder continuous in time, then this sequence converges in the strong sense even with respect to much stronger uniform Hölder norms. The convergence rate is essentially reduced by the Hölder exponent of the uniform Hölder norm in which the convergence error is measured. Two applications hereof are presented. On the one hand, estimates in such uniform norms are useful for extending results for stochastic partial differential equations with globally Lipschitz continuous nonlinearities to results for the case where the nonlinearities are only Lipschitz continuous on bounded sets. This is demonstrated for pathwise convergence rates of spectral Galerkin approximations of stochastic partial differential equations. On the other hand, strong convergence rates of multilevel Monte Carlo approximations of expectations of Banach space valued stochastic processes are derived.

This is joint work with Sonja Cox, Martin Hutzenthaler, Arnulf Jentzen, and Jan Van Neerven. [1]

[1] S. Cox, M. Hutzenthaler, A. Jentzen, J. van Neerven, and T. Welti. Convergence in Hölder norms with applications to Monte Carlo methods in infinite dimensions. *arXiv: 1605.00856* (2016), 38 pages. Minor revision requested from *IMA J. Num. Anal.*

Recursive Estimators and MCMC Algorithms

Sotirios Sabanis

School of Mathematics, University of Edinburgh, Edinburgh, Scotland, U.K.

s.sabanis@ed.ac.uk

Some recent advances on recursive estimators with discontinuity in the parameters will be discussed and links with MCMC algorithms will be highlighted.

This is joint work with Huy N. Chau, Chaman Kumar, and Miklós Rásonyi.

Bounds on the Number of Function Evaluations to Approximate the Global Minimum

Jim Calvin

New Jersey Institute of Technology

calvin@njit.edu

We consider the problem of approximating the minimum of a continuous function by sequentially choosing points at which to evaluate the function. Bounds are described for the number of evaluations required to obtain an ϵ approximation in terms of characteristics, including a norm, of the function. The average number of evaluations is bounded for a class of stochastic models for the function.

Monday 10:30–12:30, **Room: Banque CIBC**

Stochastic Gradient methods for Monte Carlo and Variational Inference

Chair: Víctor Elvira

Variational Inference: Modern Methods

Rajesh Ranganath

Princeton University

rajeshr@cs.princeton.edu

Variational inference is an umbrella term for algorithms which cast Bayesian inference as optimization. Classically, deploying variational inference requires pages of model-specific analysis. This barrier limits variational inference's ability to help quickly explore models for the data being studied. To address this, I will describe recent advances that simplify using variational inference. I will begin with black box variational inference (BBVI). BBVI is based on stochastic optimization of the KL divergence, and can be applied to new models with little model specific work. BBVI is generic, but this ease-of-use comes at a cost. The stochastic gradients used in BBVI can have high variance. We present several methods to address this, including ones that assume differentiability like the reparemeterization gradient.

The new methods of optimization in variational inference allow for new objectives to be studied. Classically, variational inference uses the KL divergence to define the optimization. Though this divergence has been widely used, the resultant posterior approximation can suffer from undesirable statistical properties. To address this, we reexamine variational inference from its roots as an optimization problem. We use operators, or functions of functions, to design variational objectives. As one example, we design a variational objective with a Langevin-Stein operator. We can characterize different properties of variational objectives, such as objectives that admit data subsampling—allowing inference to scale to massive data—as well as objectives that admit variational programs—a rich class of posterior approximations that does not require a tractable density. We illustrate the benefits of this approach on a mixture model and a generative model of images.

On Stein's Method for Practical Statistical Computation

Qiang Liu

Dartmouth College

qiang.liu@dartmouth.edu

Stein's method is a remarkable theoretical tool in probability theory for establishing approximation and limit theorems or error bounds. Although it has been mostly known to theoreticians, recent advances have shown that it can also be extremely useful for practical purposes. In this talk, we will discuss our recent works that leverage the power of Stein's method to address the practical computational challenges in probabilistic graphical models and Bayesian inference, based on a framework that integrates Stein operator with reproducing kernel Hilbert space. At the heart of this framework is a kernelized Stein discrepancy measure that allows us to access the compatibility between data and distributions based on the Fisher's score function, without knowing the normalization constants that are often critically difficult to calculate. We also show that Stein discrepancy corresponds to a type of functional gradient of KL divergence, drawing intriguing connections with variational inference and measure transport. Our framework allows to us derive a number of practical algorithms for various challenging statistical tasks,

including goodness-of-fit tests for evaluating models without knowing the normalization constants, a scalable Bayesian inference algorithm that combines the advantages of variational inference, Monte Carlo and gradient-based optimization, and approximate maximum likelihood algorithms for training deep generative models.

Control Variates, Importance Sampling, and Biconvexity

Ernest K. Ryu

UCLA

eryu@math.ucla.edu

Control variates and importance sampling are two widely used variance reduction methods. For these techniques to be effective, the weights for control variates and the sampling distribution for importance sampling must be chosen judiciously. The weights of control variates are often chosen via regression, and the sampling distribution can be chosen adaptively via stochastic convex optimization when a family of sampling distributions with a log-concave parameterization is used (Ryu 2016). Control variates and importance sampling can be combined. When control variates are used with a family of mixtures for the sampling distribution, the problem of choosing the parameters is convex (He and Owen 2014). More generally, when control variates are used with a family with log-concave parameterization for the sampling distribution, the problem of choosing the parameters is biconvex. In this work, we explore the theoretical and empirical value biconvexity brings to variance reduction.

Adaptive Monte Carlo from the Stochastic Optimization Perspective

Ingmar Schuster

Free University of Berlin

ingmar.schuster@fu-berlin.de

Adaptive Monte Carlo techniques have been a central research topic over the last two decades. We take an in-depth view of existing techniques encompassing them in a framework. We show that the often used minimum-KL-divergence criterion is inappropriate for minimizing the variance of the Monte Carlo estimators due to the poor fit of the tails of the targeted distributions. Instead, we propose stochastic gradient algorithms that directly minimize the estimator's variance in importance sampling, or maximize the Expected Squared Jumping Distance in Metropolis Hastings algorithms.

This is joint work with Víctor Elvira, Heiko Strathmann, and Christian P. Robert.

Monday 10:30–12:30, **Room: Banque Scotia**

Markov Chain Monte Carlo I

Chair: Jeff Rosenthal

Adaptive Component-Wise Multiple-Try Metropolis Sampling

Radu V. Craiu

Department of Statistical Sciences, University of Toronto, Canada

http://www.utstat.toronto.edu/craiu/

One of the most widely used samplers in practice is the component-wise Metropolis-Hastings sampler that updates in turn the components of a vector space Markov chain using accept-reject moves generated from a proposal distribution. More precisely, if we are interested in sampling from the continuous density $\pi(x) : \mathcal{X} \subset \mathbf{R}^d \to \mathbf{R}_+$; the component-wise MH (CMH) will update the ith component of the chain,x_i, using a proposal $y_i \sim T_i(\cdot|x_i)$ and setting the next value of the chain as

$$z_i = \begin{cases} y_i & \text{w.p. } \alpha_i \\ x_i & \text{w.p. } 1-\alpha_i \end{cases}$$

where

$$\alpha_i = \min\left\{1, \frac{T(x_i|y_i)\pi(y_i|x_{[-i]})}{T(y_i|x_i)\pi(x_i|x_{[-i]})}\right\},$$

and $\pi(\cdot|x_{[-i]})$ is the target conditional distribution of the ith component given all the other components $x_{[-i]} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_d)$.

When the target distribution of a Markov chain is irregularly shaped, a 'good' proposal distribution for one part of the state space might be a 'poor' one for another part of the state space. The strategy we propose here aims to close the gap that still exists between Adaptive MCMC and efficient CMH samplers. When contemplating the problem, one may be tempted to try to "learn" each conditional distribution $\pi(\cdot|x_{[-i]})$, but parametric models are likely not flexible enough and nonparametric models will face the curse of dimensionality even for moderate values of d.

For the CMH algorithm imagine that for each component we can propose k candidate moves, each generated from k different proposal distributions. Naturally, the latter will be selected to have a diverse range of variances so that we generate some proposals close to the current location of the chain and some that are further away. If we assume that the transition kernel for each component is such that among the proposed states it will select the one that is most likely to lead to an acceptance, then one can reasonably infer that this approach will improve the mixing of the chain provided that the proposal distributions are reasonably calibrated.

The computational efficiency is increased using an adaptation rule for the CMTM algorithm that dynamically builds a better set of proposal distributions as the Markov chain runs. The ergodicity of the adaptive chain is demonstrated theoretically. The performance is studied via simulations and real data examples.

This is joint work with Jinyoung Yang, Evgeny Levi, and Jeffrey Rosenthal.

Hierarchical Models: Local Proposal Variances for RWM- and MALA-within-Gibbs

Mylène Bédard

Université de Montréal

bedard@dms.umontreal.ca

We study the performance of random walk Metropolis (RWM) and Metropolis-adjusted Langevin (MALA) algorithms within Gibbs for sampling from hierarchical models. For the RWM-within-Gibbs (RWMwG), asymptotically optimal tunings for Gaussian proposal distributions featuring a diagonal covariance matrix are developed using existing scaling analyses. The principal difference with traditional optimal scaling results lies in the local character of the optimal proposal variances obtained, meaning that they vary from one iteration to the next. The concept of local proposal variances has been discussed in [1] and [2]; in the latter, scaling analyses of the RWM algorithm for hierarchical target densities are performed. Although theoretically appealing, local proposal variances had to be obtained numerically in that context, which turned out to be rather impractical. With the RWMwG sampler, these variances may now be found analytically in several cases, leading to a personalized version of the proposal variance in a given iteration. Similar ideas are applied to MALA-within-Gibbs (MALAwG), leading to efficient yet computationally affordable algorithms.

The new approach is predicated on the tractability of the distribution of the conditionally i.i.d. components, given the mixing parameters and (in practice) the observations. It is thus well suited to some hierarchical models; alternatively, we propose a fixed optimal proposal variance, which is shown to be less efficient than the local ones. In an attempt to quantify the benefit, in terms of efficiency, of using local proposal variances rather than a fixed one in the RWMwG and MALAwG, we present numerical illustrations. To add some perspective, we compare these samplers to single-block RWM and MALA, along with some of their variants that include correlation among candidates. In several cases, local versions of RWMwG and MALAwG can outperform fancy variants included in the MCMC toolbox. Local MALAwG is the approach that provides the most convincing results, leading to net efficiency gains in a wide range of situations, compared to a large set of competitors. These gains are however largely influenced by the degree of variability present in the hierarchical model. Even in cases where local samplers do not allow for large gains in terms of theoretical efficiency, the risk associated with these local variances is limited to the extra computational effort required for their implementation, which is usually insignificant compared to a fixed variance.

[1] M. Girolami and B. Calderhead. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society: Series B*, 73:123–214, 2011.

[2] M. Bédard. Hierarchical models and the tuning of RWM algorithms. Submitted, 2015.

Importance Sampling versus Delayed Acceptance MCMC when Noisy Approximations are Available

Jordan Franks

University of Jyväskylä, Finland

jordan.j.franks@jyu.fi
http://users.jyu.fi/~jojofran/

Importance sampling (IS) and delayed acceptance (DA) type algorithms are two popular approaches to Markov chain Monte Carlo (MCMC). Not much has been said, however, about the relative efficiency of these algorithms. Experimental results in a state space model context are promising and show that

the IS type approach can be favourable against a computationally comparable DA scheme (Vihola et al. [arXiv:1609.02541]). Forthcoming theoretical results provide guarantees that under moderate assumptions the IS type approach can not perform much worse than a DA or pseudomarginal type approach. Moreover, when these assumptions on the model are not assured, concrete examples are given which show that an IS (resp. DA) type approach can perform arbitrarily better than a DA (resp. IS) type approach. This shows that an IS type approach may perform better than a DA type approach when good approximate algorithms or data exist, while in contexts where good approximate algorithms or data are not available, a DA type approach may remain favourable to an IS type approach.

Locally Informed Adaptive MCMC Algorithm Based on Online PCA

Florian Maire

University College Dublin, School of Maths and Statistics

florian.maire@ucd.ie

We consider the problem of sampling from a high dimensional distribution π (defined on $\mathbb{R}^n$), whose probability mass is concentrated around lower dimensional subspaces. In such scenarios, Markov chains based on uninformed Random Walk dynamics, such as the Metropolis-Hastings (MH) algorithm, suffer from long mixing time. Some adaptive MCMC methods, such as [1], overcome this challenge by learning *globally* the covariance structure of π using the past history of the chain. We start by noting that Adaptive MH [1] is equivalent to a random walk with uncorrelated perturbation in the space spanned by the eigenvectors $\mu_1, \ldots, \mu_n$ of the empirical covariance matrix of the chain. Instead of a global uncorrelated perturbation, one can wonder if proposing a move along one principal direction μ_d, $d \in \{1, \ldots, n\}$, can be beneficial. This idea has been briefly explored in [2], where the authors suggest selecting a move along μ_d with a probability proportional to the eigenvalue λ_d of μ_d. The adaptive mechanism to estimate the principal directions follows a recursion based on online Principal Component Analysis (PCA) of the past history of the chain.

However, when n is large, selecting a type of move without accounting for the local topology may result in a geometric waiting time before the chain attempts a move in a direction that is locally sensible. This motivates a local exploration, through simulations, of the state space to favour the types of move that are more likely to be accepted. More precisely, given that the chain is at X_k, the first step of a transition $X_k \to X_{k+1}$ specifies a locally informed probability distribution $\{\omega_d(X_k)\}_{d=1}^n$ on $\{1, \ldots, n\}$, defined as a Monte Carlo estimate

$$\omega_d^{(k)}(X) :\propto k^{-1} \sum_{\ell=1}^{k} \pi\left(X + \lambda_d \epsilon_d^{(\ell)} \mu_d\right) \to_{k\to\infty} \mathbb{E}\{\pi(X + \lambda_d \epsilon_d \mu_d)\}\,,$$

where $\epsilon_d^{(1:k)} \sim_{i.i.d.} \mathcal{N}(0,1)$ and the expectation is under $\mathcal{N}(0,1)$. The second step is to propose a move according to $\{\omega_d(X_k)\}_{d=1}^n$ which is then accepted/rejected according to a ratio that maintains the chain π-stationary.

The theoretical analysis is carried out casting our chain as an non-honomogeneous adaptive chain where each kernel is conditioned on a set of particles $\epsilon_{1:n}^{(1:k)}$, regarded as exogenous variables. This allows to couple arguments from [2] for the stability of the global adaptation with some results from non-homogeneous Markov chain theory for the local learning part. We show on a number of examples that our algorithm reduces the asymptotic variance of Monte Carlo estimators compared to methods that only adapt globally.

This is joint work with Pierre Vandekerkhove.

[1] H. Haario, E. Saksman, and J. Tamminen. An adaptive Metropolis algorithm. *Bernoulli*, 2001.

[2] C. Andrieu and J. Thoms. A tutorial on adaptive MCMC. *Stats and Computing*, 18(4), 2008.

Monday 10:30–12:00, **Room: Ernst & Young**

Stochastic Differential Equations

Chair: Christian Lécot

Simulating a Stochastic Differential Equation Model by Exact Sampling

Itsaso Hermosilla

Royal Military Academy, Belgium

itsaso.hermosilla@rma.ac.be

For the study of ballistic performance of materials for protection, the analysis of the most employed ballistic protection limits is a persistent dilemma as it is characterized by dispersion. Nowadays, they are evaluated through the use of numerical methods relying on experimental procedures. However, in order to be sure of their accuracy a large amount of samples need to be collected, which makes the method expensive and time consuming. This is why developing a tool capable of simulating the ballistic limits to substitute the actual firing tests becomes our main objective.

In this study, an algorithm is proposed where we generate sample paths of a stochastic differential equation (SDE) model focused on two main ballistic protection limits. The exact simulation is carried out using scaled Chebyshev points of the second kind for the skeleton of the path and Monte Carlo methods for introducing the Brownian motion. At the first stage we simulated the V50 that estimates the velocity at which 50% of identical projectiles in identical conditions will be defeated (i.e. protection) by an armor system. For that, we assume that the deceleration of a projectile can be modelled by the following SDE which here is expressed in the way of an integral equation:

$$V(t) = V_0 - \int_0^t \mu(V(t))dt + \int_0^t \sigma(V(t))dW(t)$$

where μ is the average deceleration, σ its variance and W(t) is the standard Brownian Motion process. To match both the coefficients μ and σ to the experimental results an inverse problem has been solved. From the literature we know that the use of Monte Carlo methods for the solution of inverse problems was initiated by Keilis-Borok and Yanovskaya (1967) and Press (1968, 1971) and are in regular use these days.

Then, depending on the estimated parameters we are able to compute the more arduous ballistic protection limit V1 where only 1% of impacting projectiles will perforate the target and for which currently there are no experimental means for predicting its expectation and variance. To this purpose, an iterative algorithm is implemented as a function based on the golden search method which assumes that our objective function is unimodal.

The programming language in use for this technical computing is the highlevel dynamic language Julia licensed under the free software MIT.

This is joint work with Ben Lauwens, J. Gallant, and B. Escribano.

Importance Sampling Techniques for Stochastic Partial Differential Equations

Thalhammer Andreas

Johannes Kepler University, Linz, Austria

andreas.thalhammer@jku.at

In this talk, we consider Monte Carlo-based methods for estimating $\mathbb{E}[f(X(T))]$, where $X(T)$ denotes the mild solution of a stochastic partial differential equation (SPDE) at a given time T. It is a well-known result that the resulting Monte Carlo error can be controlled by either enlarging the number of realisations or by applying appropriate variance reduction methods. Obviously, a natural bound on the number of trajectories is imposed by the computational cost of the time integration method, which limits the possibility of increasing the number of numerical trajectories for high dimensional SDEs - especially for systems arising in the numerical treatment of SPDEs.

For this reason, we present two different approaches how importance sampling can be applied to SPDEs in order to reduce the variance of the quantity of interest.

First we consider a finite dimensional approach, where we apply importance sampling to a spatially discretized SPDE and show how this method can be used for rare event simulation purposes.

In the second part, we directly apply a measure transformation in the infinite dimensional setting and show how importance sampling can be decoupled from spatial discretization schemes. Motivated by an optimal measure transformation we construct a class of importance sampling methods and discuss how they can be implemented for numerical simulations. The key advantage of the proposed methods is that the difference in the computational effort between simulating the numerical trajectories for the standard Monte Carlo estimator and for the importance sampling methods is only the approximation of a linear, one-dimensional SODE. We conclude by presenting numerical experiments showing the effectiveness of the proposed techniques due to a remarkable reduction of the Monte Carlo error.

Initial-Boundary Value Problem for the Heat Equation—A Stochastic Algorithm

Samuel Herrmann

Institut de Mathématiques de Bourgogne (IMB) Université de Bourgogne Franche-Comté, Dijon, France

samuel.herrmann@u-bourgogne.fr

In this talk, we focus our attention on the Initial-Boundary Value Problem (IBVP) associated to the heat equation and present a new method of simulation based on the Walk on Moving Sphere Algorithm (WOMS). The main objective is to construct an efficient approximation to the solution of the IBVP. The solution is a $\mathcal{C}^{1,2}$ function u satisfying

$$\left\{\begin{array}{ll} \partial_t u(t,x) = \Delta_x u(t,x), & \forall (t,x) \in \mathbb{R}_+ \times \mathcal{D}, \\ u(t,x) = f(t,x), & \forall (t,x) \in \mathbb{R}_+ \times \partial\mathcal{D}, \\ u(0,x) = f_0(x), & \forall\ x \in \mathcal{D}, \end{array}\right.$$

where f is a continuous function defined on $\mathbb{R}_+ \times \partial\mathcal{D}$, f_0 is continuous on $\mathcal{D}$ and $\mathcal{D}$ denotes a bounded finitely connected domain in $\mathbb{R}^d$.

The foundation stone of our work is the probabilistic representation for the solution of a partial differential equation. Suppose that we are looking for the solution $u(t,x)$ of some PDE defined on the whole space $\mathbb{R}^d$. Under suitable hypothesis we can use the classical form $u(t,x) = \mathbb{E}[f(t,X_t)]$ where $(X_t)_{t\in\mathbb{R}_+}$ is a stochastic process, satisfying a stochastic differential equation, and f a known function.

In order to approximate $u(t,x)$, the Strong Law of large Number allows us to construct Monte Carlo methods once we are able to propose an approximating procedure for the stochastic process $(X_t)_{t\in\mathbb{R}_+}$.

Since we deal here with a bounded domain, the Monte-Carlo method needs a simulation procedure for the couple $(\tau_{\mathcal{D}}, X_{\tau_{\mathcal{D}}})$ where $\tau_{\mathcal{D}}$ stands for the exit time of the domain $\mathcal{D}$. That's why we propose a simulation procedure based on a sophisticated generalisation of the Walk on Spheres (WoS) algorithm first introduced to solve the Dirichlet problem for Laplace's equation, its implementation is rather easy. The definition of the random walk is based on a new mean value formula for the heat equation. The convergence results and different numerical examples permit to emphasize the efficiency and accuracy of the algorithm. This work is based on the manuscript [1] and essentially refers to [2].

This is joint work with Madalina Deaconu.

[1] Madalina Deaconu and Samuel Herrmann. Initial-boundary value problem for the heat equation - A stochastic algorithm. working paper or preprint, hal-01380365, October 2016.

[2] Mervin E. Muller. Some continuous Monte Carlo methods for the Dirichlet problem. *Ann. Math. Statist.*, 27:569–589, 1956.

Monday 15:30–17:30, **Room: Banque de Développement du Canada**

Tractability

Chair: Peter Kritzer

Geometric Brownian Motion by Brownian Bridge Construction—How Fast is the Uniform Convergence?

Ian H. Sloan

UNSW Australia

i.sloan@unsw.edu.au

It is known that the Brownian bridge or Lévy-Ciesielski construction of Brownian paths almost surely converges uniformly to the true Brownian path. But what is the error? In the present project we attempt to show for geometric Brownian motion that at level N (at which there are 2^N points evaluated on the Brownian path) the expected uniform error is of order $O(\sqrt{N}/2^{N/2})$.

This is joint work with Michael Griebel and Frances Y. Kuo.

On Exponential Convergence of Multivariate Problems

Henryk Woźniakowski

Columbia University
and
University of Warsaw

henryk@cs.columbia.edu

We provide a criterion on multivariate problems defined over Hilbert spaces which enjoy exponential convergence. The criterion is provided in terms of the eigenvalues of the corresponding compact operator characterizing the computational complexity of a multivariate problem. We apply this criterion for homogeneous and non-homogeneous tensor products. In particular, we obtain necessary and sufficient conditions on uniform exponential convergence in terms of the exponents of each tensor product factors. It turns out that uniform exponential convergence does not hold for homogeneous tensor products.

Input Sets for Numerical Integration

Erich Novak

FSU Jena, Germany, Math. Institute

erich.novak@uni-jena.de

Let F be a set of integrable functions defined on $[0,1]^d$. We say that F is an *input set for numerical integration* if there exists a (randomized) algorithm that, on input $\varepsilon > 0$ and $f \in F$, computes an approximation of the integral of f such that the error is bounded by ε with high probability.

This definition is motivated by the recent papers [1, 2]. If F is an input set then, of course, the approximate computation of the integral should be as fast as possible.

This is joint work with Daniel Rudolf.

[1] F. J. Hickernell, L. Jiang, Y. Liu and A. B. Owen (2014): Guaranteed conservative fixed width confidence intervals via Monte Carlo sampling. Monte Carlo and Quasi Monte Carlo Methods 2012, 105–128.

[2] L. Gajek, W. Niemiro and P. Pokarowski (2013): Optimal Monte Carlo integration with fixed relative precision, J. Complexity 29, 4–26.

Quasi-Monte Carlo Methods and PDEs with Random Coefficients

Josef Dick

School of Mathematics and Statistics, The University of New South Wales, Australia

josef.dick@unsw.edu.au

In this talk we consider quasi-Monte Carlo rules which are based on (higher order) digital nets and their use in approximating expected values of solutions of partial differential equations (PDEs) with random coefficients. In particular, so-called interlaced polynomial lattice rules have attractive properties when approximating such integrals. These are applied for PDEs with uniform random coefficients, but can also be used in Bayesian inversion. In this talk we give an overview of recent results in this area.

This is joint work with Robert Gantner, Frances Y. Kuo, Quoc T. Le Gia, and Christoph Schwab.

Monday 15:30–17:30, **Room: Banque CIBC**

Quasi Monte-Carlo, Sequential Monte-Carlo and Applications in Machine Learning

Chair: Simon Lacoste-Julien

Divide-and-Conquer with Sequential Monte Carlo

Fredrik Lindsten

Department of Information Technology, Uppsala University

http://www.it.uu.se/katalog/freli660

Probabilistic graphical models are widely used to represent and to reason about underlying structure in high-dimensional probability distributions. We propose a novel class of Sequential Monte Carlo (SMC) algorithms, appropriate for inference in probabilistic graphical models. This class of algorithms adopts a divide-and-conquer approach based upon an auxiliary tree-structured decomposition of the model of interest, turning the overall inferential task into a collection of recursively solved sub-problems. The proposed method is applicable to a broad class of probabilistic graphical models, *including* models with loops. Unlike a standard SMC sampler, the proposed Divide-and-Conquer SMC employs multiple independent populations of weighted particles, which are resampled, merged, and propagated as the method progresses. This method extends the standard chain-based SMC framework to a method that naturally runs on trees. We illustrate empirically that this approach can outperform standard methods in terms of the accuracy of the posterior expectation and marginal likelihood approximations. Divide-and-Conquer SMC also opens up novel parallel implementation options and the possibility of concentrating the computational effort on the most challenging sub-problems. The talk is based on the article [1].

This is joint work with Adam M. Johansen, Christian A. Naesseth, Brent Kirkpatrick, Thomas B. Schön, John Aston, and Alexandre Bouchard-Côté.

[1] F. Lindsten, A. M. Johansen, C. A. Naesseth, B. Kirkpatrick, T. B. Schön, J. A. D. Aston, and A. Bouchard-Côté. Divide-and-conquer with sequential monte carlo. *Journal of Computational and Graphical Statistics (available online)*, 2016.

Sequential Quasi-Monte Carlo

Mathieu Gerber

School of Mathematics, University of Bristol

mathieu.gerber@bristol.ac.uk

Sequential quasi-Monte Carlo (SQMC) is a class of quasi-Monte Carlo (QMC) algorithms for filtering and related sequential problems. Based on N simulations (or 'particles') it is shown in [1] that SQMC converges faster than the classical $N^{-1/2}$ Monte Carlo error rate. However, as it is often the case with QMC techniques, the numerical results in [1] show that SQMC tends to suffer from a curse of dimensionality: the performance gain of SQMC, relative to its Monte Carlo counterpart, tends to vanish for large-dimensional problems.

The objective of this talk is twofold. First, I would like to present results on the approximation error induced by the resampling step of SQMC. More precisely, we show that the variance of this latter converges to zero at a rate faster than $N^{-1-\frac{1}{d}}$, where d is the dimension of the filtering problem. The

interest of this result relies in the fact that the convergence rate of SQMC cannot be, in general, faster than that of its resampling step.

Second, I would like to discuss the application of SQMC to partly observed diffusion models, which are infinitely-dimensional. By exploiting well-known properties of these models, we are able to implement SQMC so that it outperforms significantly standard particle filtering.

This is joint work with Nicolas Chopin and Nick Whiteley.

[1] Mathieu Gerber and Nicolas Chopin. Sequential quasi Monte Carlo. *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 77(3):509–579, 2015.

Optimization Tools for Adaptive Monte-Carlo Integration

Simon Lacoste-Julien

Department of Computer Science and Operational Research (DIRO), Université de Montréal

http://www.di.ens.fr/~slacoste/

I will present some past work appearing in the machine learning community where an *optimization algorithm* was proposed to obtain adaptive quadrature rules for approximating the integrals of functions in a given reproducing kernel Hilbert space (RKHS). More specifically, variants of the venerable Frank-Wolfe optimization algorithm were proposed for this purpose [1]. Two advantages of this perspective are that 1) optimization convergence guarantees can be translated into approximation quality rates for the adaptive quadrature rules and 2) improvements on the optimization algorithm can be translated into improvements of approximation rates.

I will demonstrate the usefulness of this perspective by using the adaptive quadrature rule to improve a particle filter [2]. More specifically, we propose to replace the random sampling step in a particle filter by Frank-Wolfe optimization. By optimizing the position of the particles, we can obtain better accuracy than random or quasi-Monte Carlo sampling. In applications where the evaluation of the emission probabilities is expensive (such as in robot localization), the additional computational cost to generate the particles through optimization can be justified. Experiments on standard synthetic examples as well as on a robot localization task indicate indeed an improvement of accuracy over random and quasi-Monte Carlo sampling.

This is joint work with Fredrik Lindsten, Guillaume Obozinski, and Francis Bach.

[1] F. Bach, S. Lacoste-Julien, and G. Obozinski. On the equivalence between herding and conditional gradient algorithms. In *Proceedings of the 29th International Conference on Machine Learning (ICML)*, https://arxiv.org/abs/1203.4523, 1359–1366, 2012.

[2] Sequential kernel herding: Frank-Wolfe optimization for particle filtering. In *Proceedings of the 18th International Conference on Artificial Intelligence and Statistics (AISTATS)*, https://arxiv.org/abs/1501.02056, 554–552, 2015.

Communication Efficient Sequential Monte Carlo

Deborshee Sen

Department of Statistics and Applied Probability, National University of Singapore

deborshee.sen@u.nus.edu
https://sites.google.com/site/deborsheesen0

Distributed algorithms have become increasingly significant in recent years propelled by fast technological developments in parallel computing. For sequential Monte Carlo methods, the re-sampling step remains the main difficulty in attempting to parallelize them. We consider a recent algorithm, the so-called αSMC [2], which is an attempt at this. Interactions between particles in this algorithm are controlled by a sequence of "α" matrices. Our goal is to minimize interactions while still leading to stable algorithms. We prove that under standard assumptions the stability properties of the algorithm can be ensured by choosing well-connected, yet sparse, graphs. In particular, choosing *Ramanujan graphs* [1] lead to stable-in-time algorithms; and more generally, so do expander graphs. We next prove a central limit theorem when interactions are randomly chosen and we also prove that the asymptotic normalized variance of the filtering estimates produced by the αSMC with random interactions is stable as long as there is a certain minimum level of interaction. An offshoot of this is that the αSMC algorithm with random interaction is asymptotically equivalent to the bootstrap particle filter as long as the level of interaction increases to infinity with the number of particles, even if it is at a very slow rate.

This is joint work with Alexandre Thiery.

[1] Alexander Lubotzky, Ralph Phillips, and Peter Sarnak. Ramanujan graphs. *Combinatorica*, 8(3):261–277, 1988.

[2] Nick Whiteley, Anthony Lee, and Kari Heine. On the role of interaction in sequential monte carlo algorithms. *Bernoulli*, 22(1):494–529, 2016.

Monday 15:30–17:30, **Room: Banque Scotia**

Markov Chain Monte Carlo II

Chair: Mylène Bédard

Measuring Sample Quality with Kernels

Jackson Gorham

Stanford University, Microsoft Research

jgorham@stanford.edu

To improve the scalability of Markov chain Monte Carlo methods, researchers are developing asymptotically biased samplers that lighten the computational burden. Recent work demonstrated that standard MCMC diagnostics are flawed in this setting and defined a computable *graph Stein discrepancy* measure that could provably verify convergence of a sample to its target distribution. This approach was recently combined with the theory of reproducing kernel Hilbert spaces to define a closed-form *kernel Stein discrepancy* (KSD) computable by summing kernel evaluations across pairs of sample points. We develop a theory of weak convergence for KSDs based on Stein's method, demonstrate that commonly used KSDs fail to detect non-convergence even for Gaussian targets, and show that kernels with slowly decaying tails provably determine convergence for a large class of target distributions. The resulting convergence-determining KSDs are suitable for comparing biased, exact, and deterministic sample sequences and simpler to compute and parallelize than graph Stein discrepancies. We use our tools to compare biased samplers, select sampler hyperparameters, and improve upon existing KSD approaches to one-sample hypothesis testing and sample quality improvement.

This is joint work with Lester Mackey.

Rapid Mixing Bounds for Hamiltonian Monte Carlo on Strongly Logconcave Distributions

Oren Mangoubi

University of Ottawa

omangoubi@gmail.com

Markov chain Monte Carlo (MCMC) algorithms are ubiquitous in Bayesian statistics and other areas, and Hamiltonian Monte Carlo (HMC) algorithms are some of the most popular MCMC algorithms. In particular, it is widely believed that HMC outperforms other algorithms in high-dimensional statistical problems. Despite the popularity of HMC, its theoretical properties are not as well-understood as some of its older cousins, such as the Metropolis-Hastings or Langevin MCMC algorithms. This lack of results can make it harder to optimize HMC algorithms, and it means we do not have a good theoretical understanding of when HMC is better than other popular algorithms.

In this talk, we obtain rapid mixing bounds for Hamiltonian Monte Carlo (HMC) in an important class of strongly log-concave target distributions π, showing that HMC is faster than many competitor algorithms including the Langevin MCMC algorithm [1] in this regime. Specifically, we show that an idealized version of the HMC algorithm mixes in $\mathcal{O}^*((\frac{M_2}{m_2})^2)$ steps if the eigenvalues of the Hessian of π are bounded above and below by positive constants M_2 and m_2, respectively. We also show that a k^{th}-order numerical implementation of HMC can sample from π with arbitrary accuracy ϵ and computational cost (measured in gradient evaluations) whose dependence on ϵ and the dimension d of π is at most $d^{\frac{1}{2k}}\epsilon^{-\frac{1}{k}}$.

While most existing methods for analyzing "geometric" MCMC algorithms such as HMC use conductance bounds, conductance bounds cannot capture improvements obtained from momentum. To take advantage of momentum, our analysis of HMC instead uses probabilistic coupling bounds, obtained via Grönwall-type inequalities.

Our results improve on previous work of [2] on the non-asymptotic rate of convergence of HMC algorithms by decreasing their dependence on the dimension, extending from Gaussian to general strongly log-concave targets, and proving convergence in stronger norms.

This is joint work with Aaron Smith.

[1] Alain Durmus and Eric Moulines. Non-asymptotic convergence analysis for the Unadjusted Langevin Algorithm. *The Annals of Applied Probability*, in press.

[2] Christof Seiler, Simon Rubinstein-Salzedo, and Susan Holmes. Positive curvature and Hamiltonian Monte Carlo. In *Advances in Neural Information Processing Systems*, pages 586–594, 2014.

New MCMC Methods for the Ordering and Clustering of Single-Cell Data

Magdalena Strauss

MRC Biostatistics Unit, University of Cambridge

magdalena.strauss@mrc-bsu.cam.ac.uk

We present new MCMC methods for two important applications in systems biology, the ordering and clustering of single-cell data. Recent technologies provide RNA expression levels of large numbers of genes for individual cells, but only a single measurement per cell. As cells progress through changes at different time scales, it is possible to obtain a form of time series data even from these cross-sectional data by means of pseudotemporal ordering. A number of previous approaches have provided point estimates of the order, while more recently, Gaussian process latent variable models and MCMC methods have been applied to understand the uncertainty associated with the pseudotemporal ordering. A good understanding of this uncertainty is crucial both because of measurement noise and the inherent stochasticity of cell development. We present a new type of Gaussian process latent variable model for pseudotemporal ordering, which samples a distribution on the probability space of the orderings, that is on the group of permutations, rather than on the hugely high-dimensional vector space of possible pseudotimes, as done by previous models. We have implemented a Metropolis-Hastings sampler on our sample space, which is still very large and difficult to sample from. We found it necessary to develop novel moves in order to explore the posterior effectively. Our proposal distribution allows the sampler to make long distance moves in this space with a good acceptance rate. By modelling the data as Gaussian processes, we also include the stochasticity that is present in the data irrespective of the ordering, and estimating the parameters jointly with the ordering allows us to understand the posterior distribution in terms of a distribution on a sample space of Gaussian processes, while additionally enabling us to not only order, but also to cluster the single-cell data. Our approach applies Dirichlet process priors and Gaussian process mixture models, previously used to cluster genuine time course data, to pseudotemporal ordering, by integrating these methods into the MCMC sampler.

This is joint work with Paul Kirk, John Reid, and Lorenz Wernisch.

Efficient Markov Chain Monte Carlo Estimation of Exponential-Family Random Graph Models

Maksym Byshkin

Università della Svizzera italiana, Lugano, Switzerland

http://search.usi.ch/it/persone/e3b80ce80182a845f974f7676772506b/Byshkin-Maksym

The increase in the number and size of large network data sets requires novel efficient methods for their analysis. Exponential-family Random Graph Models (ERGMs) are a general form of probability distribution that has been shown to be widely applicable to the analysis of social and other complex networks. However, the empirical scope of ERGMs is limited by the fact that Maximum Likelihood Estimation of the model parameters can be obtained only for relatively small networks with few thousands of nodes. We improve on a recently developed Auxiliary Parameter Markov Chain Monte Carlo method [1] and propose a new MCMC approach for the Maximum Likelihood Estimation (MLE) of ERGMs on large network dataset. In contrast to existing computational approaches for MLE of exponential families (Bayesian, MCMCMLE [2] and Method of Moments [3]) the approach we propose does not carry out a large number of MCMC simulations to draw equilibrium network configurations. The approach is based on properties of equilibrium distributions of Markov chains. Using this approach we design a simple and efficient algorithm to find the MLE when it exists and is unique. In this case the suggested algorithm may be adopted to estimate ERGM parameters several orders of magnitude faster than existing algorithms. The suggested estimator is first tested on small simulated network. We compute the bias and the variance of the estimates and show that the estimates obtained with the proposed method are not less accurate than those obtained with the Method of Moments. We then apply the suggested method to the study of large-scale social and biological networks. We study network of co-authorship of scientists working on condensed matter physics [4]. The network has 40421 nodes and 175693 ties and is much larger than any complex network for which MLE was ever obtained. We show that this network may be estimated in 30 minutes. The implications of the new approach for future studies based on exponential families are discussed.

We acknowledge support from the Swiss National Platform of Advanced Scientific Computing (PASC)

This is joint work with Alex Stivala, Antonietta Mira, Garry Robins, and Alessandro Lomi.

[1] Byshkin, M., Stivala, A., Mira, A., Krause, R., Robins, G., Lomi, A. Auxiliary Parameter MCMC for Exponential Random Graph Models *Journal of Statistical Physics* 165(4), 740–754 (2016)

[2] Geyer, C. J., Thompson, E.A. Constrained Monte Carlo maximum likelihood for dependent data *Journal of the Royal Statistical Society. Series B (Methodological)* 657–699 (1992)

[3] Snijders, T.A.: Markov chain Monte Carlo estimation of exponential random graph models *Journal of Social Structure* 3(2), 1–40 (2002)

[4] Newman, M.E.: The structure of scientific collaboration networks *Proceedings of the National Academy of Sciences* 98(2), 404–409 (2001)

Monday 15:30–16:30, **Room: Ernst & Young (EY)**

Partial Differential Equations and Random Fields

Chair: Christian Lécot

Dimension-Adaptive Multi-Index Monte Carlo for PDEs with Random Coefficients

Pieterjan Robbe

KU Leuven

pieterjan.robbe@cs.kuleuven.be

We present an adaptive version of the Multi-Index Monte Carlo (MIMC) method, introduced by Haji-Ali, Nobile and Tempone (2016), for simulating PDEs with random coefficients. MIMC is an extension of the Multilevel Monte Carlo method from Giles (2008) that generalizes the scalar hierarchy of *levels* to a multi-dimensional hierarchy of *indices*. This is motivated by the observation that in some applications, refining the level of approximation can be done in several ways. Each refinement then corresponds to an index in a multi-dimensional space. The optimal shape of the hierarchy of indices is problem-dependent and can be found by an a priori analysis of the problem. However, in most applications, such analysis is prohibitively expensive. Thus, there is a need for efficient algorithms that automatically detect important dimensions in the problem. Such adaptivity has also been used for deterministic sparse grid cubature in Gerstner and Griebel (2003). We will develop a similar approach for MIMC.

This is joint work with Dirk Nuyens and Stefan Vandewalle.

Multi-Index Quasi-Monte Carlo and H-Matrices

Michael Feischl

UNSW Sydney

m.feischl@unsw.edu.au
michaelfeischl.net

We consider a new method to generate normal or log-normal random fields from [1] which builds on fast matrix-vector multiplication via H-matrices. The method proves to be robust with respect to the covariance length of the random field, and is particularly efficient for very smooth and very rough random fields. Moreover, the method applies to a fairly general class of covariance functions and is not limited to the stationary case. We use this new method in combination with quasi-Monte Carlo integration, to solve a Poisson equation with random coefficient. Moreover, to exploit the inherent sparsity of the approximation, and to obtain an efficient algorithm, we use the Multi-Index quasi-Monte Carlo approach in three coordinate directions: the finite-element approximation error, the approximation error of the random field, and the integration error of the quasi-Monte Carlo rule. This allows us to significantly reduce the computational time.

This is joint work with Josef Dick, Frances Y. Kuo, and Ian H. Sloan.

[1] M. Feischl, F. Kuo, and I. H. Sloan. Fast random field generation with H-matrices. *ArXiv e-prints*, February 2017.

Tuesday 10:30–12:30, **Room: Banque de Développement du Canada**

Stochastic Computation and Complexity II

Chair: Thomas Müller-Gronbach

Multivariate Approximation for Analytic Functions with Gaussian Kernels

Henryk Woźniakowski

Columbia University
and
University of Warsaw

henryk@cs.columbia.edu

We study d-variate approximation of analytic functions defined on R^d from a tensor product reproducing kernel Hilbert space whose kernel is Gaussian with positive shape parameters γ_j^2. The worst case setting and the class of arbitrary linear functionals is considered. We find necessary and sufficient conditions on various notions of tractability in terms of γ_j^2.

This is joint work with Ian H. Sloan.

Lower Bounds for Strong Global Approximation of Solutions of SDEs Under Adaptive Information About Additive Poisson Noise

Paweł Przybyłowicz

AGH University of Science and Technology, aculty of Applied Mathematics, Krakow, Poland

pprzybyl@agh.edu.pl

We investigate strong global approximation of solutions of SDEs with additive Poisson noise. We consider adaptive information about the driving process, which means that the successively computed sampling points may depend on the particular trajectory of the Poisson process. Under the assumption that the evaluation points are selected in the nondecreasing order we show, by using the Doob's optional stopping theorem, that the adaptive information does not help and the error is asymptotically the same as in the nonadaptive case.

[1] Debowski, J., Przybyłowicz, P., Optimal approximation of stochastic integrals with respect to a homogeneous Poisson process, *Mediterr. J. of Math.* 13 (2016), 3713–3727.

[2] Przybyłowicz, P., Optimal global approximation of stochastic differential equations with additive Poisson noise. *Numer. Algor.* 73 (2016), 323–348.

[3] Przybyłowicz, P., Optimal sampling design for global approximation of jump diffusion SDEs, http://arxiv.org/abs/1701.08311

[4] Hofmann, N., Müller–Gronbach, T., Ritter, K., The optimal discretization of stochastic differential equations. *J. Complexity* 17 (2001), 117–153.

[5] Müller–Gronbach, T., Strong approximation of systems of stochastic differential equations. *Habilitationsschrift*, TU Darmstadt (2002).

Discretization of Occupation Times for Fractional Diffusions

Taras Shalaiko

Institute of Mathematics, Mannheim University

tshalaik@mail.uni-mannheim.de

Consider a solution to an one-dimensional SDE driven by a fractional Brownian motion $\{B_t^H\}_{t\in[0,1]}$ with Hurst index $H \in (0,1)$:

$$X_t = x_0 + \int_0^t b(X_s) dB_s^H,\ t \in [0,1].$$

The integral is understood in the Young sense for $H > 1/2$ or in the sense of Newton-Cotes for $H < 1/2$. Under mild assumptions on a coefficient b and for an arbitrary Borelian set A we establish an L_2-approximation rate of nonsmooth integral-type functionals of the solution by integral sums, e.g.

$$\left(\mathsf{E}\left|\int_0^1 I_A(X_s)ds - \frac{1}{n}\sum_{k=1}^{n} I_A(X_{k/n})\right|^2\right)^{1/2}.$$

The essential tools are properties of generalized Hermite polynomials and the Wiener chaos expansion.

This is joint work with Peter Parczewski.

Complexity of High-Dimensional Approximation in Periodic Function Spaces

Thomas Kühn

Universität Leipzig, Germany

kuehn@math.uni-leipzig.de

This talk is devoted to approximation of periodic functions of finite and infinite smoothness on the d-torus, where the dimension d could be huge. The function spaces considered here are quite general; as special cases they include Sobolev and Gevrey spaces. The error is measured in the L_2-norm and expressed in terms of approximation numbers a_n of the corresponding embeddings.

The asymptotic order of approximation numbers is known in many cases, e.g. for isotropic or dominating mixed Sobolev spaces, but usually only up to *unspecified* multiplicative constants. However, for numerical purposes and for tractability questions in information-based complexity, it is useless to know only the asymptotic rate. In addition one needs precise information on the involved constants, especially their dependence on the dimension d, and on the behaviour of a_n in the preasymptotic range, i.e. for small n.

I will report on recent progress concerning these approximation problems, in particular I will mention asymptotically optimal constants, preasymptotic estimates, and tractabiliy results.

This is joint work with W. Sickel, T. Ullrich, S. Mayer, and F. Cobos.

Tuesday 10:30–12:30, **Room: Banque CIBC**

Multilevel Monte Carlo–Theory, Influence and Nested Simulation

Chair: Mike Giles

Multilevel Monte Carlo for Bayesian Inference

Kody Law

Oak Ridge National Laboratory, TN, USA

lawkj@ornl.gov

Bayesian inference provides a framework for the integration of data into an a priori known distribution. The posterior distribution, however, is known only point-wise (possibly with an intractable likelihood) and up to a normalizing constant. Monte Carlo methods have been designed to sample such distributions, such as Markov chain Monte Carlo (MCMC) and sequential Monte Carlo (SMC) samplers. Recently, the multilevel Monte Carlo (MLMC) framework has been extended to some of these cases. This talk will concern the recent development of multilevel SMC (MLSMC) samplers and the resulting estimators for standard quantities of interest as well as normalizing constants. ML particle filters and ensemble Kalman filters will also be considered, which combine dynamical systems with data in an online fashion.

Limit Theorems for Weighted or Unweighted Multilevel Estimators: Applications and Comparisons

Gilles Pagès

University Pierre and Marie Curie (UPMC)

gilles.pages@upmc.fr
http://simulations.lpma-paris.fr

We analyze in terms of *a.s.* convergence (Strong Law of Large Numbers) and weak rate (Central Limit Theorem) the performances of the Multilevel Monte Carlo estimator (MLMC) introduced in [1, 2] and of its weighted version, the Multilevel Richardson Romberg estimator (ML2R), introduced in [4]. These two estimators permit to compute a very accurate approximation of $I_0 = \mathbb{E}\, Y_0$ by a Monte Carlo type estimator when the random variable $Y_0 \in L^2(\mathbb{P})$ cannot be simulated exactly at a reasonable computational cost, whereas a family of simulatable approximations $(Y_h)_{h\in\{\mathcal{H}}$ is available. We will illustrate these results, carried out in an abstract framework, to two typical fields of interest: discretization schemes of diffusions and nested Monte Carlo. We will conclude by a brief numerical comparison of their performances with a focus on the case "$\beta > 1$" – fast strong convergence rate – where these estimators, at least in their antithetic versions, both behave asymptotically like unbiased ones.

This is joint work with V. Lemaire and D. Giorgi.

[1] M.B. Giles. Multilevel Monte Carlo Path Simulation. *Operations Reserach*, 56(3):607-617, 2008.

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

[3] D. Giorgi, V. Lemaire, G. Pagès. Limit theorems for weighted and regular Multilevel estimators. *Monte Carlo and Applications Journal*, doi.org/10.1515/mcma-2017-0102, 2017.

[4] V. Lemaire, G. Pagès. Richardson-Romberg Multilevel Extrapolation. To appear in *Bernoulli*, January 2014, arXiv:1401.1177v4.

MLMC for Value-at-Risk

Abdul-Lateef Haji-Ali

Mathematical Institute, University of Oxford

abdullateef.hajiali@maths.ox.ac.uk

This talk looks at Monte Carlo methods to estimate the Value-at-Risk (VaR) of a portfolio, which is a measure of the value and probability of the expected total loss of the portfolio in some short time horizon. It turns out that estimating VaR involves approximating a nested expectation where the outer expectation is taken with respect to stock values at the risk horizon and the inner expectation is taken with respect to the option index and stock values at some final time.

Following [1], our approach is to use MLMC to approximate the outer expectation where deeper levels use more samples in the Monte Carlo estimate of the inner expectation. We look at various control variates to reduce the variance of such an estimate. We also explore using an adaptive strategy [2] to determine the number of samples used in estimating the inner expectation. Finally, we discuss using unbiased MLMC [4] when simulating stocks requires time discretization. Our results show that using MLMC to approximate VaR with an error tolerance of ϵ, we are able to get an optimal complexity of approximately $\mathcal{O}(\epsilon^{-2})$ that is independent of the number of options, for a large enough number of options.

This is joint work with Mike Giles.

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

[2] Mark Broadie, Yiping Du, and Ciamac C Moallemi. Efficient risk estimation via nested sequential simulation. *Management Science*, 57(6):1172–1194, 2011.

[3] Wenhui Gou. Estimating value-at-risk using Multilevel Monte Carlo maximum entropy method. Master's thesis, University of Oxford, 2016.

[4] Chang-han Rhee and Peter W Glynn. Unbiased estimation with square root convergence for SDE models. *Operations Research*, 63(5):1026–1043, 2015.

MLMC for Estimation of Expected Value of Partial Perfect Information

Mike Giles

University of Oxford

mike.giles@maths.ox.ac.uk
http://people.maths.ox.ac.uk/gilesm/

Given independent random variables X and Y, the Expected Value of Partial Perfect Information (EVPPI) is defined as

$$\text{EVPPI} = \mathbb{E}\left[\max_d \mathbb{E}\left[f_d(X,Y)\,|\,X\right]\right] - \max_d \mathbb{E}[f_d(X,Y)],$$

where the maximisation is over a finite set of possible decisions d. This arises in a number of contexts, including the funding of medical research (does the value of additional information provided by a research project justify its cost?) and optimisation of oil reservoir recovery (does the additional information obtained from one more test rig justify its cost?).

In this talk we will discuss the way in which Multilevel Monte Carlo (MLMC) can be used, with 2^ℓ samples being used for the inner conditional expectation on level ℓ. Based on previous research [1, 2, 3],

an antithetic estimator is used which greatly reduces the variance of the MLMC estimator. The talk will include an overview of the numerical analysis of this variance, and some numerical experiments demonstrating its effectiveness.

This is joint work with Wei Fang, Takashi Goda, Howard Thom, and Zhenru Wang.

[1] K. Bujok, B. Hambly, and C. Reisinger. Multilevel simulation of functionals of Bernoulli random variables with application to basket credit derivatives. *Methodology and Computing in Applied Probability*, 17(3):579–604, 2015.

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

[3] T. Goda. Unbiased Monte Carlo estimation for the expected value of partial perfect information. ArXiv preprint: 1604.01120, 2016.

Tuesday 10:30–12:30, **Room: Banque Scotia**

Monte Carlo and Quasi-Monte Carlo in Computer Graphics

Chair: Derek Nowrouzezahrai

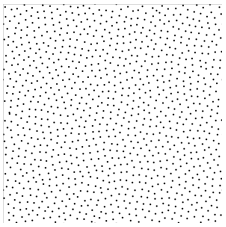
Towards Low-Discrepancy Sequences with Improved Spectral Properties: Applications in Computer Graphics

Hélène Perrier

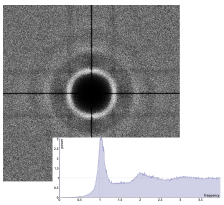
Univesité de Lyon, CNRS, LIRIS UMR 5205, Lyon, France

helene.perrier@liris.cnrs.fr

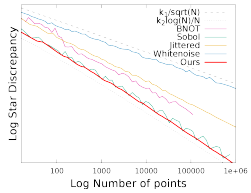
Synthetic images in computer graphics (CG) are produced by integrating light transport through the scene. This problem can be efficiently solved using classical MCQMC approach [4, 3]. However, unlike other MCQMC simulations, sampling in CG must take into account the Human Visual System, extremely sensitive to visual artefacts such as noise or aliasing (structured noise). Also, graphical tasks are often formulated in a relatively low number of dimensions (typically, up to 10 dimensions) [6]. Many sampling algorithms specifically designed for CG have been proposed, based on the notion of Blue Noise, which shapes the distribution of samples based on Fourier spectral properties (see [1] and references therein). In this presentation, we will present samplers that have controlled spectral properties while still being low discrepancy (LD) to insure a high convergence rate. We will present alternative strategies that have been developed to enhance the spectral content of a LD sequence or LD pointset [5, 2]. More precisely, we present a first sampler, limited to dimension 2, allowing us to construct point sets with both LD property and Blue-noise spectral profile (LDBN, [1]). Finally, we present preliminary results of a sampler alleviating the limitations of the LDBN sampler by being sequence, adaptive and extensible to low dimensions > 2.



(a) LDBN pointset containing 1024 samples



(b) Fourier spectrum and its radial mean



(c) Star Discrepancy graph

This is joint work with David Coeurjolly and Victor Ostromoukhov.

[1] Ahmed, A., Perrier, H., Coeurjolly, D., Ostromoukhov, V., Guo, J., Yan, D., Huang, H., and Deussen, O. Low-discrepancy blue noise sampling. *ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH Asia 2016) 35*, 6 (2016).

[2] Grünschloß, L., Hanika, J., Schwede, R., and Keller, A. (t, m, s)-nets and maximized minimum distance. In *Monte Carlo and Quasi-Monte Carlo Methods 2006*. Springer, 2008, 397–412.

[3] Keller, A. Myths of computer graphics. In *Monte Carlo and Quasi-Monte Carlo Methods 2004*. Springer, 2006, 217–243.

[4] Lemieux, C. *Monte Carlo and Quasi-Monte Carlo Sampling*. Springer Series in Statistics. Springer, Dordrecht, 2009.

[5] Owen, A. B. Randomly permuted (t,m,s)-nets and (t, s)-sequences. *Monte Carlo and Quasi-Monte Carlo Methods in Scientific Computing. Lecture Notes in Statistics, Vol. 106* (1995), 299–317.

[6] Pharr, M., Jakob, W., and Humphreys, G. *Physically Based Rendering, Second Edition: From Theory To Implementation*, 3nd ed. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2016.

Zero-Variance-Based Sampling Schemes in Light Transport Simulation

Jaroslav Křivánek

Charles University, Prague

jaroslav.krivanek@mff.cuni.cz

It has been known for a long time in neutron transport simulation, and recently pointed out in light transport calculations in computer graphics [2], that particle paths can be constructed in such a way that the resulting measurement estimators have zero variance. While such zero-variance schemes cannot be achieved without knowing the sought-after solution up front, they still are an invaluable tool for studying and designing various variance reduction sampling techniques, often referred to as 'biasing' in the neutron transport literature.

In this talk, we review our group's recent efforts on variance reduction in photon transport problems (i.e., in light transport simulation for the purpose of realistic computer graphics), based on sampling light transport paths in a way that approximates the zero-variance ideal. The underlying idea of these works is that the information necessary to construct such sampling schemes can be gathered as a side-product of the very sampling procedures used to find the final solution. Each of the works focuses on a different kind of sampling decisions in the entire transport path sampling process. First, we describe our new *direction sampling* (a.k.a. angle selection) scheme. We propose to learn the appropriate directional sampling distributions from particles generated by an adjoint process using parametric density estimation and represent them using the Gaussian mixture model [3, 1]. After that, we turn our attention to *probabilistic path termination and splitting* [4]. We show for the first time a theoretical connection between the zero-variance schemes and path termination and splitting based on the expected contribution to the final result. Finally, we describe our ongoing work on developing a zero-variance-based scheme for *scattering distance sampling* in optically participating media. All the above ideas, when put together, form the basis of an efficient and robust solution for light transport simulation in general environments consisting both of surfaces and participating media with all kinds of scattering properties.

This is joint work with Jiří Vorba, Ondřej Karlík, Martin Šik, Sebastian Herholz, and Oskar Elek

[1] Sebastian Herholz, Oskar Elek, Jiří Vorba, Hendrik Lensch, and Jaroslav Křivánek. Product importance sampling for light transport path guiding. *Comp. Graph. Forum*, 35(4), 2016.

[2] Jaroslav Křivánek and Eugene d'Eon. A zero-variance-based sampling scheme for Monte Carlo subsurface scattering. In *ACM SIGGRAPH 2014 Talks*, SIGGRAPH'14, 2014.

[3] Jiří Vorba, Ondřej Karlík, Martin Šik, Tobias Ritschel, and Jaroslav Křivánek. On-line learning of parametric mixture models for light transport simulation. *ACM Transactions on Graphics (Proceedings of SIGGRAPH 2014)*, 33(4), aug 2014.

[4] Jiří Vorba Vorba and Jaroslav Křivánek. Adjoint-driven russian roulette and splitting in light transport simulation. *ACM Trans. Graph.*, 35(4):1–11, July 2016.

Blue-Noise Dithered Sampling for Image Synthesis

Iliyan Georgiev

Solid Angle

iliyan@solidangle.com

Rendering a photo-realistic image of a virtual three-dimensional scene requires the computation of a complex light transport integral for every image pixel, and the best known method for doing that is Monte Carlo integration. The visual fidelity of such an image depends on the magnitude of the Monte Carlo pixel estimation error as well as on the distribution of this error over the image. To this end, state-of-the-art rendering applications use high-quality stratified sampling patterns which are randomly scrambled or shifted for every pixel to decorrelate the individual pixel estimates.

While the white-noise image error distribution produced by such random pixel decorrelation is eye-pleasing, it is far from being perceptually optimal. We show that visual fidelity can be significantly improved by instead *correlating* the pixel estimates in a way that minimizes the low-frequency content in the image error distribution. Inspired by digital halftoning, our *blue-noise dithered sampling* can produce substantially more faithful images, especially at low sampling rates, without actually affecting the magnitude of the pixel estimation error.

In digital halftoning, dithering is the intentional application of noise to visually randomize the error from quantizing a continuous-tone image. An efficient approach is to threshold the pixels using a blue-noise dither mask tiled over the image, whose scalar values are arranged such that the result of thresholding any constant gray-level image has an isotropic Fourier power spectrum devoid of low frequencies. That is, neighboring pixels get very different thresholds, and similar thresholds are assigned to pixels far apart. Our idea is to apply this concept to correlate pixel estimates in d-dimensional Monte Carlo light transport integration. Given a d-dimensional sampling pattern, we toroidally shift it for every pixel, but rather than choosing the offset randomly, as done traditionally, we look it up in a *blue-noise sample mask* tiled over the image. The value of every pixel in such a mask is a d-dimensional vector, and for $d = 1$ the mask is very similar to a halftoning mask. In this setting, the traditional random-offset pixel decorrelation is equivalent to using a white-noise sample mask.

Similarly to dither masks, our sample masks are scene-independent and can be pre-computed. Starting from a random (white-noise) mask, we repeatedly swap random pixel pairs to optimize the distribution of the offset vectors via simulated annealing. When used for rendering an image of a given scene, the resulting mask makes neighboring image pixels evaluate very different locations in the sampling domain, yielding a blue-noise error distribution. This visually pleasing high-frequency distribution makes the rendered image appear less noisy than that produced by traditional white-noise decorrelation, even though the amount of pixel estimation error remains the same.

This is joint work with Marcos Fajardo.

Bayesian Monte Carlo Spherical Integration for Illumination Integrals

Ricardo Marques

Universitat Pompeu Fabra, Barcelona, Spain

ricardo.marques@upf.edu

The main challenge in photo-realistic rendering lies in the computation of multidimensional integrals involving computing-intensive sampling operations. Classic Monte Carlo (CMC) methods, with its data dimension independence property, provide a straightforward solution to this problem, and are thus particularly well suited for photo-realistic rendering. However the convergence rate of CMC methods towards

the integral value is slow ($n_{-0.5}$), n being the number of samples. Consequently, a large number of samples is required to obtain an accurate estimate which, makes the task of producing a photo-realistic image tedious. Overcoming this limitation has motivated a large body of literature in Computer Graphics. In particular, several works have focused on improving the purely random sampling used in CMC which was identified as one of the causes for the slow convergence rate. Quasi-Monte Carlo (QMC) methods emerged as a more efficient alternative which, by resorting to more regular and deterministic sampling patterns, considerably boost the convergence rate of classical CMC in the case of relatively smooth integrands. Meanwhile, the Bayesian statistician O'Hagan proposed the Bayes-Hermite quadrature [3], a new form of quadrature which is commonly referred to as BMC (Bayesian Monte Carlo). This method consists of a fundamental reconsideration of the problem of Monte Carlo integration which considerably broadens the set of theoretical tools available for a more efficient integral estimate. Rasmussen and Ghahramani [2] have shown that BMC can significantly outperform CMC with importance sampling, partially due to the use of the prior knowledge for optimal sample's placement and weighting. This ability to incorporate prior probabilistic knowledge makes BMC particularly interesting for time-consuming integration problems, such as photo-realistic rendering. In this presentation, we will talk about the application of BMC to photo-realistic rendering. We will go through the existing applications of the method to rendering [1, 2], identifying its advantages and disadvantages. Furthermore, we will present ongoing work on guided sampling with BMC.

This is joint work with C. Bouville, K. Bouatouch, Jaroslav Křivánek, and J. Blat.

[1] A. O'Hagan. Bayes-hermite quadrature. *J. Statist. Plann. Inference*, 29(3), 1991.

[2] C. E. Rasmussen and Z. Ghahramani. Bayesian monte carlo. In Neural Information Processing Systems. *MIT Press*, 2002.

[3] J. Brouillat, C. Bouville, B. Loos, C. Hansen, and K. Bouatouch. A Bayesian Monte Carlo approach to global illumination. *Comp. Graph. Forum*, 28(8), 2009.

[4] R. Marques, C. Bouville, M. Ribardiére, L. P. Santos, and K. Bouatouch. A spherical gaussian framework for bayesian monte carlo rendering of glossy surfaces. *IEEE Trans. on Vis. and Comp. Graph.*, 19(10), 2013.

Tuesday 10:30–12:30, **Room: Ernst & Young (EY)**

Random Number and Random Variate Generation

Chair: Hiroshi Haramoto

The Effect of Pseudorandom Number Generators on the Convergence of Multidimensional Random Walks

Paula Whitlock

CIS Department, Brooklyn College

whitlock@brooklyn.cuny.edu

The Gnu Scientific Library (https://www.gnu.org/software/gsl/) provides a library of pseudorandom number generators for use in modeling and simulation calculations. However, the library has not been updated since the introduction of the Mersenne Twister generator in the late 1990's. In the process of developing software to implement more recently proposed and well-tested prngs, it became apparent in the testing phase that the choice of generator can still affect sensitive simulations. The generators, both existing ones in the GNU Scientific Library and the proposed new ones, were used in multi-dimensional Monte Carlo calculations of hypersphere mixtures of two-components near a phase transition. Differing behaviors were observed depending on the prng used.

This is joint work with Patrick Lempert and Wendy Mei.

On the Concatenation of Mersenne Twisters

Shin Harase

Ritsumeikan University

harase@fc.ritsumei.ac.jp

The 32-bit Mersenne Twister generator MT19937 [3] is one of the most widely used pseudorandom number generators. In practice, we often convert unsigned integers into 53-bit double-precision floating-point numbers in $[0, 1)$ in IEEE 754 format. In the header `<random>` of the C++11 STL in GCC, the following implementation has been adopted. Let $\mathbf{z}_0, \mathbf{z}_1, \mathbf{z}_2, \ldots \in \mathbb{F}_2^{32}$ be a 32-bit unsigned integer sequence generated from MT19937. To obtain 53-bit double-precision floating-point numbers in $[0, 1)$, the GCC implementation generates 64-bit unsigned integers

$$(\mathbf{z}_1, \mathbf{z}_0), (\mathbf{z}_3, \mathbf{z}_2), (\mathbf{z}_5, \mathbf{z}_4), (\mathbf{z}_7, \mathbf{z}_6), \ldots \in \mathbb{F}_2^{64} \tag{1}$$

by concatenating two consecutive 32-bit integer outputs and divides them by 2^{64}.

In this talk, we first assess the concatenation (1) via the dimension of equidistribution with v-bit accuracy, which is a most informative criterion for high dimensional uniformity of the output sequences. In fact, the concatenation (1) degrades the dimensions of equidistribution, compared with simply using 32-bit output values. Next, we analyze such phenomena by using the method in [1]. Roughly speaking, MT19937 has low-weight $\mathbb{F}_2$-linear relations on the bits of outputs, which result in a deviation. We also report that the sequences (1) with lacunary indices are rejected or have suspect p-values in several statistical tests in TestU01 [2], such as overlapping collision tests and Hamming independence tests, which have been previously unknown.

[1] S. Harase. On the $\mathbb{F}_2$-linear relations of Mersenne Twister pseudorandom number generators. *Math. Comput. Simul.*, 100:103–113, 2014.

[2] P. L'Ecuyer and R. Simard. TestU01: a C library for empirical testing of random number generators. *ACM Trans. Math. Software*, 33(4):Art. 22, 40, 2007.

[3] M. Matsumoto and T. Nishimura. Mersenne twister: a 623-dimensionally equidistributed uniform pseudo-random number generator. *ACM Trans. Model. Comput. Simul.*, 8(1):3–30, 1998.

A Note on Generating Random Variables with T-concave Densities with the Ratio-of-Uniforms Method

Josef Leydold

Instiute for Statistics and Mathematics, WU Vienna, Austria

josef.leydold@wu.ac.at

Devroye [1] proposes an acception-rejection algorithm for distributions with given log-concave density f. It requires the exact location of the mode and has a uniformly bounded rejection constant but does not require the normalization constant for f. In this talk we show that the same idea also works for the ratio-of-uniforms method. Thus we get an acception-rejection algorithm with uniformly bounded rejection constant that works for the larger class of all $T_{-1/2}$-concave densities, a generalisation of log-concave densities, that includes unimodal densities with subquadratic tails. The derivation of the algorithm is simpler than the proof in [1]. Moreover, the method can also be extended to densities where the mode is only known approximately.

This is joint work with Wolfgang Hörmann.

[1] Luc Devroye. A note on generating random variables with log-concave densities. *Statistics & Probability Letters*, 82(5):1035–1039, 2012.

Online Generation of Low-Discrepancy Sequences: From Monte-Carlo to Quasi-Monte-Carlo by Retrying

Aaditya Ramdas

Department of EECS & Department of Statistics, University of California, Berkeley

aramdas@berkeley.edu

Given an infinite sequence of i.i.d. uniform random variables on $[0,1)^d$, we construct an efficient online algorithm to select a subsequence that achieves a discrepancy, with respect to the class of axis-aligned hyper-rectangles, of $O(d \log^{2d+1} N/N)$ simultaneously for all $N \in \mathbb{N}$. The algorithm uses a variant of the "power of two choices" that was recently termed "the power of one retry", that adaptively discards at most every second point. In order to make its decision of whether to discard the N-th uniform or not, it needs $O(dN(1+\log N)^{d-1})$ space, and uses $O(d(1+\log N)^d)$ time with high probability. Our discrete-time algorithm is analyzed via the concentration of carefully constructed continuous-time point processes that maintain balance over hierarchies of Haar wavelet functions. We conjecture that a heuristic greedy-retry strategy may achieve an even better discrepancy by a $\log^d N$ factor, and we provide an efficient implementation and detailed simulations in one and two dimensions, that support our main theorem and conjecture.

This is joint work with Ohad Noy Feldheim and Raaz Dwivedi.

Tuesday 15:30–17:30, **Room: Banque de Développement du Canada**

Sequences for Quasi-Monte Carlo

Chair: Wolfgang Ch. Schmid

Implementations and Tests for Irreducible Sobol' Sequences

Christiane Lemieux

University of Waterloo

clemieux@uwaterloo.ca

Irreducible Sobol' (IS) sequences were recently introduced by Faure and Lemieux. They generalize the traditional Sobol' sequences in two main ways: they can be defined in any prime power base rather than being restricted to base 2, and they make use of irreducible polynomials to define the recurrences used to construct their generating matrices, rather than being restricted to primitive polynomials.

In this presentation we will first review IS sequences and explain their relationship to Niederreiter (and other) sequences. We will then present different implementations of IS sequences, some resulting from extensive searches for good "direction numbers", and others based on a naive construction. The proposed implementations will then be assessed using different quality criteria based either on the so-called Property A and A' introduced by Sobol', or the quality parameter t. Numerical tests on different types of integrands will be presented. Comparisons with some of the most popular implementations of Sobol' sequences will also be provided.

This is joint work with Henri Faure.

Metrical Star Discrepancy Bounds for Lacunary Subsequences of Digital Kronecker-Sequences and Polynomial Tractability

Friedrich Pillichshammer

Department of Financial Mathematics and Applied Number Theory, Johannes Kepler University Linz

friedrich.pillichshammer@jku.at

The star discrepancy $D_N^*(\mathcal{P})$ is a quantitative measure for the irregularity of distribution of an N-element point set $\mathcal{P}$ in the multi-dimensional unit cube which is intimately related to the integration error of quasi-Monte Carlo algorithms.

In 2001 it has been shown by Heinrich, Novak, Wasilkowski and Woźniakowski that for every integer $N \geq 2$ there exist point sets $\mathcal{P}$ in $[0,1)^d$ with $|\mathcal{P}| = N$ and $D_N^*(\mathcal{P}) \leq C\sqrt{d/N}$. Although not optimal in an asymptotic sense in N, this upper bound shows that the inverse of star discrepancy depends only linearly on the dimension.

The result by Heinrich et al. and also later variants thereof by other authors are pure existence results and until now no explicit construction of point sets with the above properties is known. Quite recently Löbbe studied lacunary subsequences of Kronecker's $(n\boldsymbol{\alpha})$-sequence and showed a metrical discrepancy bound of the form $C\sqrt{d(\log d)/N}$ with implied absolute constant $C > 0$ independent of N and d.

We discuss this problem and show a corresponding result for digital Kronecker sequences, which are a non-archimedean analog of classical Kronecker sequences.

This is joint work with Mario Neumüller.

Optimal Order Quasi-Monte Carlo Integration for Smooth Functions

Takashi Goda

University of Tokyo

goda@frcer.t.u-tokyo.ac.jp

Quasi-Monte Carlo (QMC) integration using order α digital nets and sequences has been shown to achieve the almost optimal order of convergence $N^{-\alpha}(\log N)^{s\alpha}$ for numerical integration in a reproducing kernel Sobolev space of arbitrary fixed smoothness $\alpha \in \mathbb{N}$, $\alpha \geq 2$. In this talk, building upon the previous existence result on optimal order QMC integration rules by the authors [2], we will prove that order $2\alpha + 1$ digital nets and sequences can achieve the best possible order of convergence $N^{-\alpha}(\log N)^{(s-1)/2}$ in the same function space. Our approach for the proof is to exploit both the decay and the sparsity of the Walsh coefficients of the reproducing kernel at the same time. Further details are available in [3]. We will highlight an analogy to the recent solution of optimal order $\mathcal{L}_2$ discrepancy bounds for order 3 digital nets due to Dick and Pillichshammer [1].

This is joint work with Kosuke Suzuki and Takehito Yoshiki.

[1] J. Dick and F. Pillichshammer, Optimal $\mathcal{L}_2$ discrepancy bounds for higher order digital sequences over the finite field $\mathbb{F}_2$, Acta Arith. 162 (2014), 65–99.

[2] T. Goda, K. Suzuki and T. Yoshiki, Optimal order quasi-Monte Carlo integration in weighted Sobolev spaces of arbitrary smoothness, IMA J. Numer. Anal. 37 (2017), 505–518.

[3] T. Goda, K. Suzuki and T. Yoshiki, Optimal order quadrature error bounds for infinite-dimensional higher-order digital sequences, Found. Comput. Math., 2017. DOI:10.1007/s10208-017-9345-0

Pair Correlations and Equidistribution

Florian Pausinger

TU München, TU Graz

florian.pausinger@gmx.at

A deterministic sequence of real numbers in the unit interval is called *equidistributed* if its empirical distribution converges to the uniform distribution. Furthermore, the limit distribution of the pair correlation statistics of a sequence is called *Poissonian* if the number of pairs $x_k, x_l \in (x_n)_{1 \leq n \leq N}$ which are within distance s/N of each other is asymptotically $\sim 2sN$. A randomly generated sequence has both of these properties, almost surely. There seems to be a vague sense that having Poissonian pair correlations is a "finer" property than being equidistributed. In this talk I will explain why this really is the case, in a precise mathematical sense: a sequence whose asymptotic distribution of pair correlations is Poissonian must necessarily be equidistributed. Furthermore, for sequences which are not equidistributed the square-integral of the asymptotic density of the sequence gives a lower bound for the asymptotic distribution of the pair correlations.

This is joint work with Christoph Aistleitner and Thomas Lachmann.

Tuesday 15:30–17:30, **Room: Banque CIBC**

Multilevel Monte Carlo for Stochastic Differential Equations

Chair: Abdul-Lateef Haji-Ali

Exact Simulation of Multivariate Itô Diffusions

Fan Zhang

Columbia University

fz2222@columbia.edu

We present the first exact simulation algorithm for general multivariate diffusions. All of the previous known algorithms for multivariate diffusions effectively assume a constant diffusion matrix and drift vector field which is conservative. Our algorithm exploits novel ε-strong simulation methods for diffusions based on rough paths, a localization technique, and a multilevel Monte Carlo construction.

This is joint work with Jose Blanchet.

An Adaptive Multilevel Monte Carlo Algorithm for Elliptic PDEs with Jump Diffusion Coefficient

Andreas Stein

SimTech, University of Stuttgart

andreas.stein@mathematik.uni-stuttgart.de

As a simplified model for subsurface flows elliptic equations may be utilized. Insufficient measurements or uncertainty in those are commonly modeled by a random coefficient, which then accounts for the uncertain permeability in a given medium. As an extension of this methodology to multi-phase flows, we incorporate jumps in the diffusion coefficient. These discontinuities then represent transitions in the heterogeneous media. More precisely, we consider a second order elliptic problem where the random coefficient is given by the sum of a (continuous) Gaussian random field and a (discontinuous) jump part. To estimate moments of the solution to the resulting random partial differential equation, we use a pathwise spacial numerical approximation combined with multilevel Monte Carlo sampling. In order to account for the discontinuities and improve the convergence of the pathwise approximation, the spatial domain is decomposed with respect to the jump positions in each sample, leading to varying grids for every path. Hence, it is not possible to create a nested sequence of grids which is suitable for each sample path a-priori. We address this issue by an adaptive multilevel algorithm, where the discretization on each level is sample-dependent and fulfills given refinement conditions.

This is joint work with Andrea Barth.

MLQMC with Product Weights for Elliptic PDEs with Lognormal Coefficients Parametrized in Multiresolution Representations

Lukas Herrmann

Seminar for Applied Mathematics, ETH Zürich

helukas@math.ethz.ch

Parametric diffusions are considered with lognormal coefficients that are given by multiresolution representations. Approximations by QMC with randomly shifted lattice rules for first order are analyzed with dimension independent convergence rates. The local support structure in the multiresolution expansion are known to allow product weights for QMC rules [3]. Product weights allow for linear scaling in the dimension of integration in the cost to create QMC rules by the CBC construction [6]. Multilevel QMC quadratures are considered to reduce the work of the QMC approximation in general polyhedral spatial domains [4, 5]. Analogous results hold for affine-parametric operator equations [1, 2]. This research is supported in part by the Swiss National Science Foundation (SNSF) under grant SNF 159940.

This is joint work with Christoph Schwab.

[1] Gantner, R.N., Herrmann, L., and Schwab, Ch. Quasi-Monte Carlo integration for affine-parametric, elliptic PDEs: local supports imply product weights. Technical Report 2016-32, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2016.

[2] Gantner, R.N., Herrmann, L., and Schwab, Ch. Multilevel QMC with product weights for affine-parametric, elliptic PDEs Technical Report 2016-54, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2016.

[3] Herrmann, L. and Schwab, Ch. QMC integration for lognormal-parametric, elliptic PDEs: local supports imply product weights. Technical Report 2016-39, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2016.

[4] Herrmann, L. and Schwab, Ch. Multilevel QMC with product weights for elliptic PDEs with lognormal coefficients. Technical Report, Seminar for Applied Mathematics, ETH Zürich, Switzerland (2017). (in preparation)

[5] Herrmann, L., Schwab, Ch. QMC algorithms with product weights for lognormal-parametric PDEs. Technical Report 2017-04, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2017.

[6] Nuyens, D. and Cools, R. Fast algorithms for component-by-component construction of rank-1 lattice rules in shift-invariant reproducing kernel Hilbert spaces. *Math. Comp.*, 75(254):903–920 (electronic), 2006.

Analysis of Multi-Index Monte Carlo Estimators for a Zakai SPDE

Zhenru Wang

University of Oxford

zhenru.wang@maths.ox.ac.uk

In this article, we propose a space-time Multi-Index Monte Carlo (MIMC) [1] estimator for a one-dimensional parabolic stochastic partial differential equation (SPDE) of Zakai type. We compare the complexity with the Multilevel Monte Carlo (MLMC) method of Giles and Reisinger (2012) [2], and find, by means of Fourier analysis, that the MIMC method: (i) has suboptimal complexity of $O(\varepsilon^{-2}|\log\varepsilon|^3)$

for a root mean square error (RMSE) ε if the same spatial discretisation as in the MLMC method is used; (ii) has a better complexity of $O(\varepsilon^{-2}|\log\varepsilon|)$ if a carefully adapted discretisation is used; (iii) has to be adapted for non-smooth functionals. Numerical tests confirm these findings empirically.

This is joint work with Christoph Reisinger.

[1] A. L. Haji-Ali, F. Nobile, and R. Tempone. *Multi-index Monte Carlo: when sparsity meets sampling.* Numerische Mathematik, 132(4):767–806, 2015.

[2] M. B. Giles and C. Reisinger. *Stochastic finite differences and multilevel Monte Carlo for a class of SPDEs in finance.* SIAM Journal on Financial Mathematics, 3(1):572–592, 2012.

Tuesday 15:30–17:30, **Room: Banque Scotia**

Function Approximation and Variance Reduction

Chair: Henryk Wozniakowski

High-Dimensional Function Approximation—Breaking the Curse with Monte Carlo Methods

Robert J. Kunsch

Friedrich Schiller University Jena

Robert.Kunsch@uni-jena.de

We study tractability of the information cost $n(\varepsilon, d)$ for d-variate function approximation problems. In the deterministic setting for many unweighted problems the *curse of dimensionality* holds, that means, for some fixed error tolerance $\varepsilon > 0$ the complexity $n(\varepsilon, d)$ grows exponentially in d. For integration problems one can usually break the curse with a standard Monte Carlo method, for function approximation problems however similar effects of randomization have been unknown so far. In this talk I will present two examples from my PhD thesis where randomization breaks the curse of dimensionality.

The first example is about L_∞-approximation of functions from unweighted periodic tensor product Hilbert spaces where the initial error is constant 1, thus properly normalized. For particular problems of this kind (e.g. Korobov spaces of smoothness greater than 1) we can prove the curse of dimensionality in the worst case setting, in contrast to polynomial tractability with $n^{\mathrm{ran}}(\varepsilon, d) \preceq d\,(1 + \log d)\,\varepsilon^{-2}$ in the randomized setting. The algorithm is linear and uses n independent "Gaussian functionals" as random information, that way being a function approximation analogon of standard Monte Carlo integration. The basic idea of the algorithm originates from [1] and has been used to show that randomization can help to improve the order of convergence for several approximation problems.

The second example is the L_1-approximation of *monotone functions* $f : [0,1]^d \to [-1,+1]$, where $\mathbf{x} \le \mathbf{y}$ implies $f(\mathbf{x}) \le f(\mathbf{y})$. Here we allow only function values for information. While the curse of dimensionality holds in the deterministic setting [2], this is not the case in the randomized setting. In detail, for fixed ε the complexity "only" depends exponentially on $\sqrt{d}\,(1 + \log d)$, however still, the problem is *not* weakly tractable. The algorithm is based on standard Monte Carlo approximation of certain Haar wavelet coefficients, a similar approach for monotone Boolean functions $f : \{-1,+1\}^d \to \{-1,+1\}$ can be found in [3].

[1] P. Mathé. Random Approximation of Sobolev Embeddings. *Journal of Complexity*, 7(3):261–281, 1991.

[2] A. Hinrichs, E. Novak, H. Woźniakowski. The curse of dimensionality for the class of monotone functions and for the class of convex functions. *Journal of Approximation Theory*, 163(8):955–965, 2011.

[3] N. H. Bshouty, C. Tamon. On the Fourier spectrum of monotone functions. *Journal of the ACM*, 43(4):747-770, 1996.

Least Squares Regression for Non-Stationary Designs

David Barrera

École Polytechnique, Palaiseau, France
and
Institut de Mathématiques de Toulouse, Toulouse, France

david.barrera@polytechnique.edu

The main goal of this talk is to present a series of new results concerning the rate of convergence for least square estimates in the case in which the sampling process (the "design") is not i.i.d.

Our results are given in the setting of nonparametric regression, and they cover the corresponding estimates in the i.i.d. case (as given, for instance, in [1]) without any essential loss in the respective rates of convergence nor the introduction of additional hypotheses in order to carry out the proofs. They justify also a more general -but very natural- interpretation of the least-squares regression function as the "best" approximation to the response function in a given statistical experiment, and provide further theoretical ground for the research on numerical methods in which a non-stationary evolution has to be considered.

We illustrate these results and their aforementioned interpretation in the numerical context by looking at estimation problems in which the i.i.d. setting is either not satisfiable or not convenient, emphasizing in particular the Markovian setting. We also illustrate the relevance of these tools in the error analysis of Monte Carlo algorithms like the one in [2].

This is joint work with Gersende Fort and Emmanuel Gobet.

[1] Györfi, L.; Kohler,M.; Krzyzak, A. and Walk, H. (2002). A Distribution-Free Theory of Nonparametric Regression. *Springer Ser. Statist.*

[2] Fort, G.; Gobet, E. and Moulines, E. (2017) MCMC Design-Based Non-Parametric Regression for Rare Event. Application for Nested Risk Computations. To appear in *Monte Carlo Methods Appl.*

SGD with Variance Reduction Beyond Empirical Risk Minimization

Massil Achab

Centre de Mathématiques Appliquées, CNRS, Ecole Polytechnique, Palaiseau, France

massil.achab@cmap.polytechnique.fr

We introduce a doubly stochastic proximal gradient algorithm for optimizing a finite average of smooth convex functions, whose gradients depend on numerically expensive expectations. Indeed, the effectiveness of SGD-like algorithms relies on the assumption that the computation of a subfunction's gradient is cheap compared to the computation of the total function's gradient. This is true in the Empirical Risk Minimization (ERM) setting, but can be false when each subfunction depends on a sequence of examples. Our main motivation is the acceleration of the optimization of the regularized Cox partial-likelihood (the core model in survival analysis), but other settings can be considered as well.

The proposed algorithm is doubly stochastic in the sense that gradient steps are done using stochastic gradient descent (SGD) with variance reduction, and the inner expectations are approximated by a Monte-Carlo Markov-Chain (MCMC) algorithm. We derive conditions on the MCMC number of iterations guaranteeing convergence, and obtain a linear rate of convergence under strong convexity and a sublinear rate without this assumption.

We illustrate the fact that our algorithm improves the state-of-the-art solver for regularized Cox partial-likelihood on several datasets from survival analysis.

This is joint work with Agathe Guilloux, Stéphane Gaiffas, and Emmanuel Bacry.

Compactness Approaches for Importance Sampling

Alexander Shkolnik

Department of Statistics, UC Berkeley

ads2@berkeley.edu

A common measure of optimality of an estimator Q_a under an importance measure $\mathbf{Q}$ is defined by the ratio of its second moment to its first moment on a logarithmic scale.

$$\rho = \liminf_{a \uparrow \infty} \left\{ \frac{\frac{1}{a} \log \mathbf{E_Q}\, Q_a^2}{\frac{1}{a} \log \mathbf{E_Q}\, Q_a} \right\} \tag{1}$$

Here, a defines some problem dependent parametrization and we assume the sequences in the numerator and denominator are bounded. When $\rho = 2$, the number of importance sampling trials to approximate $\mathbf{E_Q}\, Q_a$, to some fixed level of precision, is subexponential in a. Such an estimator is called asymptotically optimal. A plain Monte Carlo scheme exhibits $\rho = 1$. A standard approach to selecting the optimal importance measure $\mathbf{Q}$ relies on large deviations theory. That is, large deviation techniques are employed to derive upper and lower bounds on the sequences appearing in the numerator and denominator of (1). Such an approach requires the development of an appropriate large deviations principle (LDP) with a good rate function. This program is frequently too difficult to implement for practical problems.

We provide an analysis that circumvents many obstacles in the standard program. The key tool is the large deviations counterpart of Prohorov's (relative compactness) theorem from weak convergence theory, demonstrated by A. Puhalskii [1]. That is, when the laws of $\{Q_a\}$ are exponentially tight, from any subsequence for which which $\rho < 2$, we can extract a further subsequence along which an LDP with a good rate function is guaranteed to exist. The approach reduces the optimality analysis to a verification of exponential tightness in some suitable topology, a much easier task. It also facilitates the identification of importance sampling schemes that are suboptimal but much superior to plain Monte Carlo, i.e. $\rho > 1$.

Using our approach we derive a simpler set of conditions for optimality than those stemming from the usual application of the Gärtner-Ellis theorem. We provide an alternative proof of optimality for the dynamic importance sampling estimators of Dupuis & Wang, [2], extending the scope of their application to affine Markov processes. Time permitting, we demonstrate the effectiveness of the approach on some challenging problems in event timing simulation.

[1] Anatolii Puhalskii. On functional principle of large deviations. *New Trends in Probability and Statistics, Vol. 1 (Bakuriani, 1990)* 198-219, VSP, Utrecht, 1991.

[2] Paul Dupuis and Hui Wang. Importance sampling, large deviations, and differential games. *Stochastics: An International Journal of Probability and Stochastic Processes*, 76(6):481–508, 2004.

Tuesday 15:30–17:30, **Room: Ernst & Young (EY)**

Simulation of Stochastic Processes

Chair: Geneviève Gauthier

Simulation of Multivariate Mixed Poisson Processes

Michael Chiu

University of Toronto, IBM

chiu@cs.toronto.edu

Multivariate mixed Poisson processes have many important applications in Insurance, Finance, and many other areas of Applied Probability. In this paper, we extend the Backward Simulation (BS) approach [1] to modelling multivariate mixed Poisson processes and analyze the connection to the extreme measures describing the joint distribution of the processes at the end of the time horizon. The BS approach relies on the conditional uniformity of the arrival times, given the number of events and allows for the correlation coefficient between Poisson processes to take extreme values. In an earlier work [2], the forward continuation of the BS was introduced for Poisson processes in order to achieve richer correlation profiles. It was also shown that the forward continuation of the BS is asymptotically stationary. Along with the extension of the BS to multivariate mixed Poisson processes, we investigate the forward contiunation of the BS approach for multivariate mixed Poisson processes.

This is joint work with Alexander Kreinin and Ken Jackson.

[1] Kreinin, A.: Correlated Poisson processes and their applications in financial modeling. Financial Signal Processing and Machine Learning (2016)

[2] Chiu, M., Jackson, K. and Kreinin, A.: Correlated Multivariate Poisson Processes and Extreme Measures. ArXiv e-prints 1702.00376

Simulation of Student-Lévy Processes Using Series Representations

Till Massing

Department of Economics, University of Duisburg-Essen

till.massing@uni-due.de

Lévy processes have become very popular in many applications in finance, physics and beyond. The Student-Lévy process is one interesting special case where increments are heavy-tailed and, for 1-increments, Student t distributed. Although theoretically available, there is a lack of path simulation techniques in the literature due to its complicated form. We address this issue using series representations of Rosiński [2] with the inverse Lévy measure method and the rejection method and prove upper bounds for the mean squared approximation error. Furthermore, we extend the numerical inverse Lévy measure method of Imai and Kawai [1] to incorporate explosive Lévy tail measures. Monte Carlo studies verify the error bounds and the effectiveness of the simulation routine. As a side result we obtain series representations of the so called inverse gamma subordinator which is used to generate paths in this model.

[1] Junichi Imai and Reiichiro Kawai. Numerical inverse Lévy measure method for infinite shot noise series representation. *J. Comput. Appl. Math.*, 253:264–283, December 2013.

[2] Jan Rosiński. *Series Representations of Lévy Processes from the Perspective of Point Processes*, pages 401–415. Birkhäuser Boston, Boston, MA, 2001.

Minimizing Time Discretization Error

Krzysztof Bisewski

Centrum Wiskunde en Informatica, Kortweg de Vries Institute for Mathematics University of Amsterdam

bisewski@cwi.nl

We consider the error arising from time discretization when estimating the tail of the distribution of a supremum of a real stochastic process $(X_t)_{t\in[0,1]}$, namely $w(b) := P\big(\sup_{t\in[0,1]} X_t > \ b\ \big)$. For the standard Brownian Motion we prove that if the discretization is equidistant, then in order to control the error as b grows large, the number of grid-points has to grow at least quadratically in b. As an alternative to equidistant discretization, we derive an explicit family of grids, with grid-points adaptive in b. For the adaptive family the required number of grid-points to control the error is independent of b, providing a significant computational improvement. The adaptive grids that we develop can be used to construct a strongly efficient algorithm for the estimation of $w(b)$.

This is joint work with Daan Crommelin and Michel Mandjes.

Unbiased Simulation and Parameters Estimation of Distributions with Explicitly Known Fourier Transforms

Yiwei Wang

Dept. of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Hong Kong

ywwang@se.cuhk.edu.hk

The characteristic functions of Affine and Lévy processes have explicit forms, which help us simulate and estimate parameters from the "frequency-domain". In this paper, we propose an importance-sampling based method of obtaining unbiased estimators for evaluating expectations involving random variables with explicitly known Fourier transforms. In contrast to existing methods, our approach avoids time-consuming numerical Fourier inversion and can be applied effectively to high dimensional option pricing. We also provide a method-of-moments estimator based on the characteristic functions using observations of both stock and option prices. Consistency and asymptotic normality of the estimator are proved under some regular conditions. Through some numerical examples, we show that the proposed methods are particularly effective.

This is joint work with Nan Chen.

Wednesday 10:30–12:30, **Room: Banque de Développement du Canada**

Stochastic Computation and Complexity III

Chair: Michaela Szölgyenyi

Lower Error Bounds for Strong Approximation of Scalar SDEs with Non-Lipschitzian Coefficients

Thomas Müller-Gronbach

University of Passau

thomas.mueller-gronbach@uni-passau.de

We study the problem of pathwise approximation of the solution of a scalar SDE, either at the final time or globally in time, based on n sequential evaluations of the driving Brownian motion on average. We present lower error bounds in terms of n under mild local smoothness assumptions on the coefficients of the SDE. This includes SDEs with superlinearly growing or piecewise Lipschitz continuous coefficients and also certain types of CIR-processes.

This is joint work with Mario Hefter and André Herzwurm.

On the Euler-Maruyama Scheme for SDEs with Discontinuous Diffusion Coefficient

Dai Taguchi

Osaka University

dai.taguchi.dai@gmail.com
https://sites.google.com/site/daitaguchihomepage/

Let $X = (X_t)_{0\leq t\leq T}$ be the solution of one-dimensional stochastic differential equation (SDE) $\mathrm{d}X_t = \sigma(X_t)\mathrm{d}W_t$, where $W = (W_t)_{0\leq t\leq T}$ is a standard one-dimensional Brownian motion. The solution X is rarely analytically tractable, so one often approximates X by using the Euler-Maruyama scheme $X^{(n)} = (X_t^{(n)})_{0\leq t\leq T}$ given by $\mathrm{d}X_t^{(n)} = \sigma(X_{\eta_n(t)}^{(n)})\mathrm{d}W_t$, where $\eta_n(t) = kT/n$ if $t \in [kT/n, (k+1)T/n)$. It is well-known that if σ is Lipschitz continuous, the Euler-Maruyama scheme for X converges at the strong rate of order $1/2$, that is $E[|X_T - X_T^{(n)}|] \leq Cn^{-1/2}$ for some $C > 0$. The strong rate in the case of non-Lipschitz coefficient has been studied recently. Gyöngy and Rásonyi [1] prove that if σ is β-Hölder continuous, then

$$E[|X_T - X_T^{(n)}|] \leq \begin{cases} Cn^{-\beta+1/2} & \text{if } \beta \in (1/2, 1), \\ C(\log n)^{-1} & \text{if } \beta = 1/2. \end{cases}$$

In this talk, we assume that the diffusion coefficient σ satisfies that $\sigma := \rho \circ f$ where ρ is $1/2$-Hölder continuous and there exists $0 < \underline{\rho} < \overline{\rho}$ such that $\underline{\rho} \leq \rho(x) \leq \overline{\rho}$ and $f = f_1 - f_2$, f_1 and f_2 are bounded, strictly increasing with finite discontinuous points. Note that under the above assumption, X has a unique strong solution, (see [2]). Under the above assumption, we will show that there exists $C > 0$ such that

$$E[|X_T - X_T^{(n)}|] \leq C(\log n)^{-1}.$$

Moreover, we will prove that if σ is monotone and β-Hölder continuous with $\beta \in (0,1)$ then

$$E[|X_T - X_T^{(n)}|] \leq Cn^{-\beta/2}.$$

The idea of proof is to use the Yamada-Watanabe approximation technique, and "tightness" and some estimations of the local time of the Euler-Maruyama approximation.

[1] Gyöngy, I. and Rásonyi, M.: A note on Euler approximations for SDEs with Hölder continuous diffusion coefficients. Stochastic. Process. Appl. 121, 2189–2200 (2011).

[2] Le Gall, JF.: One-dimensional stochastic differential equations involving the local times of the unknown process. In Stochastic analysis and applications 1984, 51–82, Springer Berlin Heidelberg.

[3] Ngo, H-L., and Taguchi, D.: Strong convergence for the Euler-Maruyama approximation of stochastic differential equations with discontinuous coefficients. Statist. Probab. Lett., 125 (2017) 55–63.

The Euler Scheme for SDEs with Discontinuous Drift Coefficient: A Numerical Study of the Convergence Rate

Andreas Neuenkirch

Universität Mannheim

andreas.neuenkirch@googlemail.com

The Euler scheme is one of the standard schemes to obtain numerical approximations of stochastic differential equations (SDEs). Its convergence properties are well-known in the case of globally Lipschitz continuous coefficients. However, in many situations, relevant systems do not show a smooth behaviour, which results in SDE models with discontinuous coefficients. In this manuscript, we will carry out intensive numerical tests for the convergence properties of the Euler scheme. The test equations under consideration are scalar SDEs with a piecewise constant drift coefficient and a constant diffusion coefficient. Our tests reveal that for inward pointing drift coefficients convergence rates are higher and independent of the initial conditions and furthermore the estimates are stable. This seems to be due to the ergodicity of the Euler scheme and the underlying SDE.

This is joint work with Simone Göttlich and Kerstin Lux.

Designing and Benchmarking Monte Carlo Methods for Simulating Processes in Discontinuous Media

Antoine Lejay

Inria Nancy Grand-Est/Institut Élie Cartan de Lorraine, Nancy, France
and
Inria Paris

Antoine.Lejay@inria.fr

Simulating diffusion processes in discontinuous media is a challenging problem with important practical applications. Among others, geophysics, population ecology, biology, oceanography, molecular chemistry are domains in which one encounters diffusion phenomena where the diffusivity presents sharp changes. For example, in the Darcy's law, the hydraulic conductivity depends on the rock properties. Therefore, in the case of layered aquifers or fractured media for instance, it may be modelled as discontinuous in space.

Our presentation deals with a diffusion process evolving in one-dimensional discontinuous media.

In a first time, we present how we designed an exact simulation method for dealing with such a problem. This method combines stochastic analysis and expressions of the involved distributions [2].

In a second time, we present a set of benchmark tests for simulation techniques of diffusion processes in 1D discontinuous media based on both physical and mathematical relevance. These tests aim at investigating the qualities and defaults of existing numerical methods. We use them on four different methods proposed in [1, 2, 4]. From these empirical studies, we emphasize that the quality of the numerical methods relies essentially on a criterium of symmetry. The better the symmetry criterium is respected, the better the method. We think that this criterium should probably be true in higher dimension and could serve as a guide to design numerical methods for the 2D and the 3D cases [3].

This is joint work with Géraldine Pichot.

[1] H. Hoteit, R. Mose, A. Younes, F. Lehmann, and Ph. Ackerer. Three-dimensional modeling of mass transfer in porous media using the mixed hybrid finite elements and the random-walk methods. *Math. Geology*, 34(4):435–456, 2002.

[2] A. Lejay and G. Pichot. Simulating diffusion processes in discontinuous media: a numerical scheme with constant time steps. *J. Comput. Phys.*, 231(21):7299–7314, 2012.

[3] A. Lejay and G. Pichot. Simulating diffusion processes in discontinuous media: benchmark tests. *J. Comput. Phys.*, 314:384–413, 2016.

[4] G. J. M Uffink. *Analysis of dispersion by the random walk method.* Phd thesis, Delft University, The Netherlands, 1990.

Wednesday 10:30–12:30, **Room: Banque CIBC**

Multivariate Decomposition Methods and Truncation Algorithms

Chair: Peter Kritzer

A New Construction of Active Sets for the Multivariate Decomposition Method

Alexander Gilbert

UNSW Sydney, Australia

alexander.gilbert@student.unsw.edu.au

An important component of the Multivariate Decomposition Method is constructing an *active set*, which is, loosely speaking, a set consisting of the subsets of variables which contribute most to the integral. In this talk we introduce a new construction of "optimal" active sets. We will assume that the integrands of interest belong to a weighted function class, where the weights in some sense describe the importance of collections of variables and are of product form. The general idea is to order the weights in decreasing order and add the sets that correspond to the largest weights until it is guaranteed that the active set error is below a specified tolerance. Ordering weights requires significant computation, so we also outline a simplified construction of "quasi-optimal" active sets, which instead of ordering all weights, only orders "blocks" of weights. Here a block of weights is all the weights in a given interval. We will present numerical results which illustrate that our new active sets are smaller than the previous construction and also that the quasi-optimal active sets are only slightly larger than optimal ones.

This is joint work with Greg W. Wasilkowski.

Integration Over $\mathbb{R}^{\mathbf{N}}$ Using the Multivariate Decomposition Method and Higher-Order QMC Rules

Dirk Nuyens

NUMA, Department of Computer Science, KU Leuven, Belgium

dirk.nuyens@cs.kuleuven.be

We approximate an integral of an infinite-variate function $g(x_1, x_2, \ldots)$ over $\mathbf{R}^{\mathbf{N}}$

$$\begin{aligned} I_\rho(g) = I(f) &:= \lim_{s\to\infty} \int_{\mathbf{R}^s} g(x_1, \ldots, x_s, 0, 0, \ldots)\, \rho_1(x_1) \cdots \rho_s(x_s)\, \mathrm{d}x_1 \cdots \mathrm{d}x_s \\ &= \lim_{s\to\infty} \int_{\mathbf{R}^s} f(x_1, \ldots, x_s, 0, 0, \ldots)\, \mathrm{d}x_1 \cdots \mathrm{d}x_s, \end{aligned}$$

where f is evaluated by the anchored decomposition (with anchor $a = 0$)

$$f(\mathbf{x}) = \sum_{|u|<\infty} f_u(\mathbf{x}_u), \qquad f_u(\mathbf{x}_u) = g([\mathbf{x}_u; 0])\, \rho_u(\mathbf{x}_u) - \sum_{v\subset u} f_v(\mathbf{x}_v), \qquad f_\emptyset = g(0, 0, \ldots),$$

using the multivariate decomposition method

$$I(f) = \sum_{|u|<\infty} \int_{\mathbf{R}^{|u|}} f_u(\mathbf{x}_u)\, \mathrm{d}\mathbf{x}_u \quad \approx \quad \sum_{u\in U_\epsilon} \frac{1}{n_u} \sum_{k=1}^{n_u} f_u(\mathbf{x}_u^{(k)}) = \sum_{u\in U_\epsilon} Q_u^r(f_u; n_u) =: Q_\epsilon(f).$$

We achieve bounds of the form

$$|I(f) - Q_\epsilon(f)| \le c \, \|f\| \, \mathrm{Cost}(Q_\epsilon)^{-r+o(1)},$$

assuming product and order-dependent (POD) bounds on the derivates of f.

This is joint work with Dong T. P. Nguyen.

Truncation Dimension for Linear Problems on Weighted Anchored and ANOVA Spaces

Peter Kritzer

RICAM, Austrian Academy of Sciences

peter.kritzer@oeaw.ac.at

We discuss linear problems on weighted anchored and ANOVA spaces of high-dimensional functions. The main questions addressed are: When is it possible to approximate a continuous linear operator applied to multivariate functions with all but the first k variables set to zero, so that the corresponding error is small? What is the truncation dimension, i.e., the smallest number $k = k(\varepsilon)$ such that the corresponding error is bounded by a given error demand ε? As it turns out, $k(\varepsilon)$ could be very small for sufficiently fast decaying weights.

This is joint work with Aicke Hinrichs, Friedrich Pillichshammer, and Greg W. Wasilkowski.

Discrete Maximal Regularity and Discrete Error Estimate of a Non-Uniform Implicit Euler–Maruyama Scheme for a Class of Stochastic Evolution Equations

Yoshihito Kazashi

School of Mathematics and Statistics, UNSW Sydney

z5038125@unsw.edu.au

We study two discrete properties of an approximation algorithm for a class of stochastic evolution equations proposed by Müller-Gronbach and Ritter [1, 2]. Their algorithm treats the spatial and stochastic discretisation with a Galerkin method, which introduces the errors of the spatial and Q-Wiener process truncation. The temporal discretisation is treated with an implicit Euler–Maruyama scheme with a non-uniform time discretisation. The temporal discretisation error, too, can be related to the truncation degree of the Q-Wiener process, as shown by Müller-Gronbach and Ritter.

In this talk, we study this scheme in more detail. Our interest is in a discrete analogue of the maximal regularity estimate and a temporally discrete error estimate.

[1] T. Müller-Gronbach and K. Ritter. An implicit Euler scheme with non-uniform time discretization for heat equations with multiplicative noise. *BIT Numerical Mathematics* 47 (2007), 393–418.

[2] T. Müller-Gronbach and K. Ritter. Lower bounds and nonuniform time discretization for approximation of stochastic heat equations. *Foundations of Computational Mathematics* 7 (2007), 135–181.

Wednesday 10:30–12:30, **Room: Banque Scotia**

Stochastic Optimization

Chair: Fabian Bastin

Variable Sample-Size Stochastic Approximation with Finite Sampling Budget

Uday V. Shanbhag

Pennslvania State University, USA

udaybag@psu.edu

Traditional stochastic approximation (SA) schemes for constrained stochastic convex optimization employ a single gradient or a fixed batch of noisy gradients in computing a new iterate and require a projection on a possibly complex convex set for every update. As a consequence, practical implementations may require the solution of a large number of projection problems, and may render this scheme impractical. We present a variable sample-size stochastic approximation (VSSA) scheme where the batch of noisy gradients may change in size across iterations (batch size is denoted by N_k) and the scheme terminates when the prescribed sampling budget (given by M) is consumed. In this setting, we derive error bounds in strongly convex and convex differentiable regimes and focus on quantifying the rate of convergence in terms of both the number of projection steps and the simulation budget and comment on the optimality of the obtained rates. In addition, we present amongst the first stochastic approximation variant of the (Nesterov-)accelerated gradient scheme which displays the optimal accelerated rate of convergence in terms of projection steps (denoted by K) if the sample size grows sufficiently fast. Preliminary numerics suggest that VSSA schemes outperform their traditional counterparts in terms of computational time and compete well in terms of theoretical rates. More specifically, our key contributions include the following:

(I) *Strongly convex stochastic optimization:* We present an increasing sample-size SA scheme under constant ($\gamma_k = \gamma$) and diminishing ($\gamma_k = \theta/k$) steplength regimes. Specifically, given a simulation budget M, we prove that the exponential rates of convergence are recovered in terms of projection steps when the sample-size grows at a geometric rate. In fact, we show that the expected error in solution iterates diminishes at the optimal rate of $\mathcal{O}(1/\sqrt{M})$;

(II) *Convex differentiable stochastic optimization:* Next, when the strong convexity assumption is weakened, we show that by increasing N_k at suitably defined rates, based on M and problem parameters, we note that the scheme admits the optimal deterministic rate of convergence in terms of projection steps, namely $\mathcal{O}(1/K)$. In addition, we observe that the empirical error is $\mathcal{O}(1/\sqrt{M})$ and displays a constant factor improvement over the error bound obtained by Nemirovksi, Juditsky, Lan, and Shapiro in 2009 through their robust stochastic approximation procedure.

(III) *Accelerated schemes for stochastic approximation:* Next, we propose what we believe is a new scheme for stochastic approximation. This scheme utilizes a constant steplength but relies on an increasing sample size sequence to obtain a rate of convergence of $\mathcal{O}(1/K^2)$ in terms of projection steps while the expected sub-optimality diminishes at the rate of $\mathcal{O}(1/M^{1/3})$.

(IV) *Numerical studies:* Preliminary numerical studies are promising and the schemes provide comparable empirical error but do so at a fraction of the effort. For instance, accelerated stochastic approximation schemes are seen to obtain comparable results in less than a hundredth of the effort in projection steps.

This is joint work with Afrooz Jalilzadeh, Jose Blanchet, and Peter W. Glynn.

Tractable Models for Satisficing under Uncertainty

Sanjay Dominik Jena

École des Sciences de la Gestion (ESG UQAM), CIRRELT

jena.sanjay-dominik@uqam.ca

Satisficing, as an approach to decision-making under uncertainty, aims at achieving solutions that satisfy the problem's constraints as well as possible. Mathematical optimization problems that are related to this form of decision-making include the P-model of [1], where satisficing is the objective, as well as chance-constrained and robust optimization problems, where satisficing is articulated in the constraints. In this paper, we first propose the R-model, where satisficing is the objective, and where the problem consists in finding the most "robust" solution, feasible in the problem's constraints when uncertain outcomes arise over a maximally sized uncertainty set. We then study the key features of satisficing decision making that are associated with these problems and provide the complete functional characterization of a satisficing decision criterion. As a consequence, we are able to provide the most general framework of a satisficing model, which seeks to maximize a satisficing decision criterion in its objective, and the corresponding satisficing-constrained optimization problem that generalizes robust optimization and chance-constrained optimization problems. Next, we focus on a tractable probabilistic satisficing model, whose objective is a lower bound of the P-model.

We show that when probability densities of the uncertainties are log-concave, the T-model can admit a tractable concave objective function. In the case of discrete probability distributions, such as those based on samples, the model simplifies to a linear mixed integer program of moderate dimensions. We also show how the model can be extended to multi-stage decision-making and present the conditions under which the problem is computationally tractable. Computational experiments on a stochastic maximum coverage problem strongly suggest that the obtained solutions can be highly effective, thus allaying misconceptions of having to pay a high price for the satisficing models in terms of solution conservativeness. These observations are consistent with preliminary results for an optimization problem in Project Management. Furthermore, to solve the P-model exactly, the aforementioned techniques are used to sample and directly operate on the underlying conditional probabilities.

This is joint work with Patrick Jaillet, Adam Ng, and Melvyn Sim.

[1] Charnes, A., and W. Cooper (1963). Deterministic Equivalents for Optimizing and Satisficing under Chance Constraints. Operations Research 11(1):18–39.

Scenario Generation Methods that Replicate Crossing Times in Spatially Distributed Stochastic Systems

Joseph L. Durante

Department of Electrical Engineering, Princeton University

jdurante@princeton.edu

When developing policies for controlling a system under uncertainty, it is essential they perform well across a realistic population of scenarios. To generate scenarios of an exogenous stochastic process given a forecast of the process over a time horizon, standard time series models can be used to produce sample paths that replicate certain characteristics of observed forecast error series, such as autocorrelation, cross correlations, and the distribution of errors. However, we have observed that most models fail to capture one important, yet overlooked, characteristic of stochastic processes involved in sequential decision making problems - the *crossing times* of the process. A crossing time is a contiguous block of time for which

the observed value of the process is above or below its forecasted value. The crossing times of stochastic processes are important in applications ranging from energy storage and power system control to portfolio management. We present models for time series simulation, which we call *crossing state models*, that outperform standard time series models in their ability to replicate these crossing times. Both univariate and multivariate crossing state models are developed. In the multivariate case, the crossing times of the stochastic process, which may be spatially distributed, are replicated at both the disaggregate and aggregate levels. The models utilize two-level simulation techniques to simultaneously control crossing time and error distributions in the generated scenarios. The first level of simulation is concerned with the evolution of the crossing state, while error generation occurs on the second level.

If $X_{t,j}$ represents the forecast error of subprocess j and X_t^{agg} represents the aggregate error at lead time t, the crossing state is defined as $S_t^C = (\mathbf{1}_{\{X_{t,1}>0\}}, \mathbf{1}_{\{X_{t,2}>0\}}, ..., \mathbf{1}_{\{X_{t,k}>0\}}, \mathbf{1}_{\{X_t^{agg}>0\}})$. In low dimensions the evolution of the crossing state is modeled using a semi-Markov model in which crossing state sojourn times and transitions probabilities are drawn directly from historical empirical distributions. In higher dimensions we cannot maintain this form of the transition function. Instead, we approximate the crossing state transition probability for each subprocess using a logistic model: $P(S_{t+1,j}^C = 1 | S_t^C, S_{t-1}^C, ..., S_{t-p+1}^C) = 1/1 + exp\left[c_j + \beta_{0,j} S_t^C + \beta_{1,j} S_{t-1}^C + ... + \beta_{p-1,j} S_{t-p+1}^C\right]$. The second level error generation model is conditioned on the crossing state and chosen so that it is appropriate for the stochastic process of interest. Our choices include discrete state Markov chains, VAR models, and general linear models. Note if errors generated on the second level remain consistent with the simulated crossing states, then simulated crossing time distributions will closely match observed distributions. Consistency is ensured by rejection of inconsistent error vectors followed by resampling.

This is joint work with Raj Patel and Warren B. Powell.

A New Framework for Generating Scenario Trees Using Quasi-Monte Carlo Methods

Julien Keutchayan

Department of Mathematics and Industrial Engineering, Polytechnique Montréal.

Interuniversity Research Centre on Enterprise Networks, Logistics and Transportation (CIRRELT).

Julien.Keutchayan@cirrelt.ca

Stochastic programming models decision-making problems under uncertainty, for which the distribution of the random parameters is known or can be estimated. In multistage stochastic programming problems, decisions are made at several stages and the random information is revealed progressively between each stage. The goal is to find a feasible policy which maximizes or minimizes an objective function, typically the expectation of some function of the decisions and the random parameters. In general, random parameters have a large number of possible values (possibly infinitely many), which makes stochastic programming problems hard to solve. However, approximate solutions can be obtained by generating scenario trees, i.e., by discretizing the random parameters distribution with a finite number of scenarios. quasi-Monte Carlo (QMC) methods provide a way to perform such discretization, with the goal to outperform the other discretization methods, including the standard Monte Carlo sampling. While many advances have been made recently in the use of QMC for two-stage problems, a generalization to problems with more than two stages remains difficult because scenario trees tend to grow exponentially with the number of stages. In this presentation, we introduce a new framework for generating scenario trees for two-stage and multistage problems. In this framework, scenario trees are built by taking account of not only the probability distribution of parameters but also the objective function and the problem constraints, with the goal to obtain scenario trees more suitable to the problem and therefore more efficient.

This is joint work with David Munger, Michel Gendreau, and Fabian Bastin.

Wednesday 10:30–12:30, **Room: Ernst & Young (EY)**

Goodness-of-Fit Tests

Chair: Paula Whitlock

A Data-Dependent Choice of the Tuning Parameter for Goodness-of-Fit Tests Employing Bootstrapped Critical Values

Leonard Santana

Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa

leonard.santana@nwu.ac.za

In [1], a data-dependent choice of a tuning parameter that appears in many goodness-of-fit tests is discussed. However, this method is only applicable to a class of tests for which the null distribution is unknown and independent of unknown parameters, but can be approximated through simple Monte Carlo methods. This data-dependent choice of the tuning parameter is obtained by maximising the bootstrap power of a test when calculated from the Monte Carlo approximation of the appropriate critical value. We now consider a slightly broader class of tests, that is, where the null distribution is unknown and cannot be approximated via Monte Carlo methods. Unlike the approach followed in [1] the approach followed for these tests is to approximate the null distribution (and resulting critical values) using the bootstrap. Typical tests that fall in this class include testing for symmetry, and testing goodness-of-fit for the gamma, generalised exponential, skewed normal, and normal mixture distributions, to name but a few. The new method to obtain the data-dependent choice of the tuning parameter in these tests does not rely on maximising bootstrap power, but rather relies on attempting to find the value of the tuning parameter that allows the test to come as close as possible to the specified nominal significance level. An iterative bootstrap algorithm which employs the warp speed method of [2] is provided and the results of the performance of our new method is investigated in two testing scenarios: testing symmetry of the error distribution of a regression model, and testing goodness-of-fit for the gamma distribution. Various test statistics containing a tuning parameter are employed for these scenarios.

This is joint work with Willem D. Schutte and James S. Allison.

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

[2] Raffaella Giacomini, Dimitris N. Politis, and Halbert White. A warp-speed method for conducting monte carlo experiments involving bootstrap estimators. *Econometric Theory*, 29:567–589, 2013.

A Monte Carlo Evaluation of the Performance of Two New Tests for Symmetry Based on the Empirical Characteristic Function

James S. Allison

Unit for Business Mathematics and Informatics, North-West University, Potchefstroom, South Africa

james.allison@nwu.ac.za

We propose two new tests for symmetry based on well-known characterisations of symmetric distributions. Both of these tests are based on the empirical characteristic function. Because the limit null

distribution of the new test statistics is complicated and depends on unknown parameters, we propose a Mote Carlo based wild bootstrap algorithm in order to obtain the critical values. The performance of the new tests is evaluated and compared to that of other existing tests by means of a Monte Carlo study. All tests are carried out in a regression setup where we test whether the error distribution in a linear regression model is symmetric. It is found that the newly proposed tests perform favourably compared to the other tests. The talk concludes with a real world example where we assess symmetry of the distribution of the error terms in a GARCH(1,1) model.

This is joint work with Marius Smuts and Charl Pretorius.

Support Points—A New Way to Compact Distributions

Simon Mak

Georgia Institute of Technology

smak6@gatech.edu

This talk first introduces a new way to compact a continuous probability distribution F into a set of representative points called *support points*. Support points are obtained by minimizing the energy distance, a statistical potential measure initially proposed by Székely and Rizzo (2004) for testing goodness-of-fit. The energy distance has two appealing features. First, its distance-based structure allows us to exploit the duality between powers of the Euclidean distance and its Fourier transform for theoretical analysis. Using this duality, we show that support points converge in distribution to F, and enjoy an improved error rate to Monte Carlo for integrating a large class of functions. Second, the minimization of the energy distance can be formulated as a difference-of-convex program, which we manipulate using two algorithms to efficiently generate representative point sets. In simulation studies, support points provide improved integration performance to both Monte Carlo and a specific Quasi-Monte Carlo method.

We then extend the notion of support points to a new type of representative point set called *projected support points*. The primary appeal of projected support points is that it provides an optimal representation of not only the full distribution F, but its marginal distributions as well. We present a unifying theoretical framework for projected support points, connecting the desired goodness-of-fit on marginal distributions with important principles in experimental design and Quasi-Monte Carlo. Two algorithms are then proposed for efficient optimization of projected support points, with simulation studies confirming the effectiveness of the proposed point set in integrating high-dimensional functions with low-dimensional structure. An important application of projected support points – as a way to optimally compact Markov-chain Monte Carlo (MCMC) chains – is then highlighted using a Bayesian age-cohort model for breast cancer.

This is joint work with V. Roshan Joseph.

[1] Székely, Gábor J., and Maria L. Rizzoe. Testing for equal distributions in high dimension. *InterStat*, 5:1-6 (2004).

[2] Mak, Simon, and V. Roshan Joseph. Support points. *Annals of Statistics*, under review. arXiv preprint arXiv:1609.01811 (2017).

[3] Mak, Simon, and V. Roshan Joseph. Projected support points, with application to optimal MCMC reduction. *Journal of the American Statistical Association*, in preparation.

Testing Soundness of Statistical Tests for Random Number Generators by Using a Three-Level Test

Hiroshi Haramoto

Faculty of Education, Ehime University

haramoto@ehime-u.ac.jp

Statistical tests are indispensable for evaluating pseudorandom number generators (PRNGs). There are many statistical test suites, but implementers of statistical test suites are faced with a difficult problem that weather each statistical test has flaws such as wrong mathematical analysis and poor implementation.

Okutomi and Nakamura proposed a three-level test which adopts an error-free function at the second level in order not to accumulate computational errors of the approximation formula for p-value.

In MCM 2015, we proposed a method to check flaws in statistical tests in NIST SP800-22 by using the above three-level test. This approach is experimentally so that it is suitable for tests which are difficult to analyze theoretically.

In this talk, we will show experimental results of this checking for statistical tests implemented in TestU01. Furthermore, we will examine the reliability of Discrete Fourier Transform test (Spectral test) for PRNGs proposed by several researchers.

Wednesday 15:30–17:00, **Room: Banque de Développement du Canada**

Sobol' Indices and sobol' Sequences

Chair: Christiane Lemieux

Sobol' Indices for Constrained Global Sensitivity Analysis

Sergei Kucherenko

Imperial College London, London, UK

s.kucherenko@imperial.ac.uk

Global sensitivity analysis (GSA) is used to identify key parameters whose uncertainty most affects the output. This information can be used to rank variables, fix or eliminate unessential variables and thus decrease problem dimensionality. Among different approaches to GSA variance-based Sobol sensitivity indices (SI) are most frequently used in practice owing to their efficiency and ease of interpretation. Most existing techniques for GSA were designed under the hypothesis that model inputs are independent. However, in many cases there are dependences among inputs, which may have significant impact on the results. Such dependences in a form of correlations have been considered in the generalised Sobol GSA framework developed by Kucherenko et al. [1]. There is an even wider class of models involving inequality constraints (hence the term constrained GSA or cGSA) which impose structural dependences between model variables. In this case the parameter space may assume any shape depending on the number and nature of constraints. This class of problems encompasses a wide range of situations encountered in the natural sciences, engineering, design, economics and finances where model variables are subject to certain limitations imposed e.g. by conservation laws, geometry, costs, quality constraints etc. An important particular case within this approach corresponds to imposing a minimum (maximum) threshold for the model output, i.e., $f(x_1, \cdots, x_d) > f_{min}$, in which case the constraint function can be defined as $g(x_1, \cdots, x_d) = f(x_1, \cdots, x_d) - f_{min}$ and the corresponding constraint as $g(x_1, \cdots, x_d) > 0$.

There are two distinct approaches for estimating Sobol' SI. The classical one is based on using direct integral formulas which in practice are substituted by MC estimates. This approach typically requires prohibitively large number of function evaluations. In the second approach, a metamodel of the original full model is constructed first and then this metamodel is used for estimating Sobol' SI. Typically, this approach significantly reduces the cost of evaluation Sobol' SI. In this work we compare two approaches for estimating Sobol' SI on several test functions related to cGSA.

This is joint work with Oleksiy V. Klymenko and Nilay Shah.

[1] Sobol', I.M. and Kucherenko, S. (2005) Global Sensitivity Indices for Nonlinear Mathematical Models. Review. *Wilmott Magazine*, 1,56–61.

[2] Kucherenko, S., Tarantola, S., and Annoni, P. (2012) Estimation of global sensitivity indices for models with dependent variables. *Comp. Physics Comm.*, 183,937–946.

Automatic Estimation of First-Order Sobol' Indices Using the Replication Procedure

Lluís Antoni Jiménez Rugama

Illinois Institute of Technology

ljimene1@hawk.iit.edu

We consider models of the form $f(\mathbf{X})$ where $\mathbf{X} \sim \mathcal{U}[0,1]^d$. The normalized Sobol' indices $S_{\mathbf{u}}$ in [4] measure what part of the variance $\mathrm{Var}[f(\mathbf{X})]$ is explained by a subset of inputs indexed by $\mathbf{u} \subseteq \{1,\dots,d\}$.

In practice, these indices are usually unknown and need to be estimated using model evaluations that can be expensive to obtain. For first-order indices, i.e. $\mathbf{u} = j \in \{1,\dots,d\}$, the main disadvantage is that we require a total of $(d+1)n$ model evaluations to estimate each index with n evaluations. The replication procedure introduced in [5] allows to estimate all first-order indices using orthogonal arrays with $2n$ evaluations instead.

In this talk we present an extension of our adaptive integration Sobol' rules [3] to estimate first-order indices with the replication procedure. These Sobol' rules choose n automatically to ensure that $|S_j - \widehat{S}_j| \leq \varepsilon$, for all $j \in \{1,\dots,d\}$ and a user-specified error tolerance ε.

This is joint work with Laurent Gilquin, Élise Arnaud, Fred J. Hickernell, Hervé Monod, and Clémentine Prieur.

[1] L. Gilquin, and L. A. Jiménez Rugama, "Reliable error estimation for Sobol' indices," *Statistics and Computing*, 2017+. Under review.

[2] L. Gilquin, L. A. Jiménez Rugama, E. Arnaud, F. J. Hickernell, H. Monod, and C. Prieur: "Iterative construction of replicated designs based on Sobol' sequences," *Comptes Rendus Mathématique*, 355, 10–14, 2017.

[3] F. J. Hickernell, and L. A. Jiménez Rugama, "Reliable adaptive cubature using digital sequences," in *Monte Carlo and Quasi-Monte Carlo Methods 2014* (R. Cools and D. Nuyens, eds.), 367–383, Springer International Publishing, 2016.

[4] I. M. Sobol', "Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates," *Mathematics and Computers in Simulation (MATCOM)*, 55(1), 271–280, 2001.

[5] J.-Y. Tissot, and C. Prieur, "A randomized orthogonal array-based procedure for the estimation of first- and second-order Sobol' indices," *Journal of Statistical Computation and Simulation*, 85(7), 1358–1381, 2015.

Importance Sampling Techniques for Semi-Parametric Single Index Models

Christiane Lemieux

University of Waterloo

clemieux@uwaterloo.ca

Importance sampling (IS) is a popular variance reduction technique for rare-event simulations. The crucial part of IS is the design of an efficient proposal distribution. In principle, a proposal distribution is constructed in such a way that it samples more heavily from the important region. The problem is that it becomes hard to characterize which part of the domain corresponds to the important region when many variables are involved, unless the problem has a low-dimensional structure. Problems based on semi-parametric single-index models often have such a structure, as they are characterized by having a

one-dimensional projection (single index) of input that captures a large majority of the overall variance. For such models, it becomes trivial to identify the important region through the projected variable. In this talk we propose an IS technique that exploits this feature and show how to construct an optimal proposal distribution for this type of models. Our method can also be viewed as a dimension reduction technique, as shown by a study of the corresponding Sobol' sensitivity indices. Hence it works well with quasi-Monte Carlo methods, an advantage we exploit by using Sobol' sequences as the sampling mechanism for our IS method. Numerical examples showing the efficiency of this method are presented.

This is joint work with Yoshihiro Taniguchi.

Wednesday 15:30–17:00, **Room: Banque CIBC**

Variance Reduction for Rare-Event Simulation

Chair: Zdravko I. Botev

Rare-Event Simulation for Products of Random Variables

Hui (Alice) Yao

University of Queensland

h.yao@uq.net.au

We explore variance reduction techniques for estimating rare–event probabilities associated to products of random variables. While the associated random variables might be light–tailed, their products can be heavy–tailed so their estimation is often difficult. Moreover, since the distribution of a product of random variables is seldom available in explicit form, then it is not always possible to implement traditional methods in a straightforward way. In this talk we show how to adapt existing techniques for estimating the probabilities of interest and we further analyse their asymptotic performance. We complement our results with applications in risk.

This is joint work with Leonardo Rojas-Nandayapa and Thomas Taimre.

On the Efficient Simulation of the Left-Tail of the Sum of Correlated Log-Normal Variates

Nadhir Ben Rached

Computer, Electrical, and Mathematical Science and Engineering (CEMSE) Division King Abdullah, University of Science and Technology (KAUST), Thuwal, Makkah Province, Saudi Arabia.

nadhir.benrached@kaust.edu.sa

The sum of Log-normal variates is encountered in many challenging applications such as in performance analysis of wireless communication systems and in financial engineering. Several approximation methods have been developed in the literature, the accuracy of which is not ensured in the tail regions. These regions are of primordial interest wherein small probability values have to be evaluated with high precision. Apart from the works of [1] and [2], most of the existing simulation approaches have considered the problem of estimating the right-tail of the sum of Log-normal random variables. In the present work, we consider instead the estimation of the left-tail region. We propose a logarithmic efficient estimator combining the mean-shifting importance sampling approach that was proposed in the work of [2] with a control variate technique. We prove via various simulation results that the proposed approach achieves a considerable amount of variance reduction compared to that of [2]. More precisely, the correlation coefficient between the estimator of [2] and the introduced control variable converges to 1 as we decrease the probability of interest which leads to a substantial amount of variance reduction.

This is joint work with Abla Kammoun, Mohamed-Slim Alouini, and Raul Tempone.

[1] Søren Asmussen, Jens Ledet Jensen, and Leonardo Rojas-Nandayapa. Exponential family techniques for the lognormal left tail. *Scandinavian Journal of Statistics*, 43(3):774–787, Sep. 2016.

[2] Archil Gulisashvili and Peter Tankov. Tail behavior of sums and differences of log-normal random variables. *Bernoulli*, 22(1):444–493, Feb. 2016.

Practical Estimators for the Distribution of the Sum of Correlated Log-Normal Random Variables

Zdravko I. Botev

UNSW Sydney

botev@unsw.edu.au

This talk is a complementary talk to the one given by Nadhir Ben Rached, and presents two novel Monte Carlo estimators for the left and the right tail of the sum of dependent log-normal random variables. Both estimators are logarithmically efficient, and the left-tail estimator enjoys the additional advantage that it is continuously differentiable. As a result of this, the left-tail estimator permits not only additional variance reduction via randomized Quasi Monte Carlo, but also a simple and reliable estimator of the corresponding probability density function. We briefly explain why, despite many proposed estimators in the literature, the proposed estimator for the right tail may be the only one of practical value.

This is joint work with D. Mackinlay, R. Salomone, and Pierre L'Ecuyer.

Wednesday 15:30–17:00, **Room: Banque Scotia**

Markov Chain Monte Carlo in Bayesian Statistics

Chair: Radu V. Craiu

Fast, Approximate MCMC for Bayesian Analysis of Large Data Sets: A Design Based Approach

Matthias Kaeding

University of Duisburg-Essen and RWI - Leibniz Institute for Economic Research

matthias.kaeding@rwi-essen.de

We propose a fast approximate Metropolis-Hastings algorithm for large data sets, embedded in a design based approach. Here, the loglikelihood ratios involved in the Metropolis-Hastings acceptance step are considered as data. The building block is one single subsample from the complete data set, so that the necessity to store the complete data set is bypassed. The subsample is taken via the cube method, a balanced sampling design, which is defined by the property that the sample mean of some auxiliary variables is close to the sample mean of the complete data set. We develop several computationally and statistically efficient design based estimators for the Metropolis-Hastings acceptance probability. Our simulation studies show that the approach works well and can lead to results which are close to the use of the complete data set, for considerably smaller computational cost. The methods are applied on a large data set consisting of all German diesel prices for the first quarter of 2015.

Variational Sequential Monte Carlo

Christian A. Naesseth

Linköping University
and
Columbia University

christian.a.naesseth@liu.se

Variational inference (VI) casts Bayesian inference as optimization that minimizes a divergence from an approximating family to the posterior distribution. The optimization problem in VI has been simplified and made more tractable by the use of stochastic gradients. These stochastic gradients come from both subsampling the data and sampling from the variational approximation, i.e., Monte Carlo approximations of expectations. The move away from the conjugacy constraints required by classical variational inference without stochastic gradients allows for more complex models and more flexible variational approximating families, and enables the practitioner to focus on modeling rather than inference.

Despite these advances in variational inference, the need for variational inference algorithms that are consistent and which control the variance of the Monte Carlo estimates remains. We propose variational sequential Monte Carlo (VSMC) to address this need. VSMC is a new approximating family of distributions that melds variational inference and sequential Monte Carlo (SMC). VSMC provides practitioners with a flexible, accurate, and powerful approximate Bayesian inference algorithm. The combination inherits two desirable properties of both VI and SMC: fast optimization of a well-defined objective as well as asymptotic consistency.

VSMC learns model parameters using variational expectation-maximization and scales to large data using inference networks. Additionally, VSMC provides a new motivation for importance weighted autoencoders (IWAE): the IWAE lower bound is a special case of the VSMC bound with sequence length

one. This provides us with new insights into how to use the IWAE variational approximation for inference. From the view of Monte Carlo algorithms, VSMC learns accurate proposal distributions for importance sampling and SMC-based methods such as particle Markov chain Monte Carlo.

Studying both real and synthetic data, we show that the VSMC achieves better performance than other state-of-the-art methods. Not only does VSMC achieve better lower bounds on the marginal likelihood, it also shows significantly better posterior samples.

This is joint work with Scott W. Linderman, Rajesh Ranganath, and David M. Blei.

Weak Convergence and Optimisation of the Reversible Jump Algorithm

Philippe Gagnon

Université de Montréal

gagnonp@dms.umontreal.ca

In the context of Bayesian statistical inference, model selection and parameter estimation are commonly based on the joint posterior distribution of models and parameters. Elaborated models can however lead to complex posterior distributions, requiring help from numerical approximation methods to actually compute the posterior model probabilities (which are used to discrimate among models) and parameter estimates. Markov chain Monte Carlo (MCMC) methods represent the most commonly used solution for obtaining approximations of these probabilities and estimates. The idea behind these methods is to construct a Markov chain with an invariant measure that corresponds to the posterior distribution that we are interested in (usually called the target distribution). The simulation of such chains, combined with the Law of Large Numbers, allow to obtain the approximations that we are looking for.

The reversible jump algorithm is a useful MCMC method introduced by [1] that allows switches between subspaces of differing dimensionality, and therefore, selecting models and estimating parameters from a single output. This method has a tremendous potential because of this capability to deliver information on both the "good" models and their parameters, simultaneously. For instance, it can be used in parametrical clustering to quickly estimate the number of components and the parameters of mixtures (see [2]). There is however a price to pay: many functions have to be specified in order to implement the reversible jump algorithm. This is why, although this method is now increasingly used in key areas of human activity (e.g. finance and biology), it still remains a challenge to practically and efficiently implement it. In our paper, we focus on a simple sampling context in order to obtain theoretical results that are used to propose an optimal design for the sampler. This allows users to easily and efficiently implement the reversible jump algorithm. The key result is the weak convergence of the sequence of stochastic processes engendered by the algorithm. This represents the main contribution of our paper as this is, to our knowledge, the first weak convergence result for the reversible jump algorithm.

This is joint work with Mylène Bédard and Alain Desgagné.

[1] Peter J Green. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrika*, 82(4):711–732, 1995.

[2] Sylvia Richardson and Peter J Green. On Bayesian analysis of mixtures with an unknown number of components (with discussion). *J. R. Stat. Soc. Ser. B. Stat. Methodol.*, 59(4):731–792, 1997.

Wednesday 15:30–17:00, **Room: Ernst & Young (EY)**

Monte Carlo in Particle and Quantum Physics

Chair: Erich Novak

Convergence Probabilistic Analysis of the Monte Carlo Method for Quantum Physics Problems

Ivan T. Dimov

Institute of Information and Communication Technologies & Institute of Mathematics and informatics, Bulgarian academy of sciences, Sofia, Bulgaria

ivdimov@bas.bg
http://parallel.bas.bg/dpa/BG/dimov/

Convergence of class of Monte Carlo methods dealing with observables in Quantum Physics is analyzed. We deal with the numerical approximation of observables related to the Wigner equation using probabilistic techniques based on branching particle systems.

We answer several questions about the behavior of the algorithm and demonstrate theoretically why almost always is not stable and how to deal with this instability. Our work relies exclusively on probabilistic techniques and the estimates related to the proposed algorithms can be seen as sharpening of the more general study of stochastic algorithms for the Wigner equation. The work also summarizes the formulation of the Wigner equation as an operator equation in suitable L_2 spaces.

This is joint work with Mladen Savov.

A New Monte Carlo Method for Estimation of Time Asymptotic Parameters of Polarized Radiation

Natalya Tracheva

Institute of Computational Mathematics and Mathematical Geophysics SB RAS
and
Novosibirsk State University, Novosibirsk, Russia

tnv@osmf.sscc.ru

There exists a number of problems in radiation transfer theory, where an asymptotic behavior of radiation beams for large time periods in light-scattering media is in focus. It is known that for nonpolarized radiation under some rather general conditions such time asymptotics is exponential. A parameter of this exponential asymptotics is the principal eigenvalue of the homogeneous stationary transfer equation with standard boundary conditions (see [1, 2]).The challenge is to extend and improve this assertion to the case of polarized radiation, on the basis of large-scale Monte Carlo simulation.

In this work we consider the problem of estimation of time asymptotic parameters of polarized radiation flux, outgoing from a semi-infinite layer of scattering and absorbing media with a light source on its boundary. We construct a distinctive Monte Carlo weighted algorithm for evaluation of time asymptotic parameters of polarized radiation flux. This algorithm is based on the randomized projective evaluation of the functionals via the orthonormal polynomial expansion. Using this method and precise computer simulation we investigate how significant polarization impact is.

This work was supported by Russian Foundation for Basic Research (project number 17-01-00823).

This is joint work with Sergey Ukhinov.

[1] B. Davison, *Neutron transport theory*. Oxford: Clarendon Press, 1957.

[2] G.A. Mikhailov, N.V. Tracheva, S.A. Ukhinov, Monte Carlo Estimate of Backscattering Noise Asymptotics Parameters with Allowance for Polarization, *Atmospheric and Oceanic Optics*, 2011, 24(2), 109–118. Pleiades Publishing, Ltd., 2011.

A Shannon Entropy-Based Strategy for Adjusting History Number of Time-Dependent Transport Problem Automatically

Danhua Shangguan

Institute of applied physics and computational mathematics, Beijing, China

sgdh@iapcm.ac.cn

This paper proposes a Shannon entropy-based strategy for adjusting history number of time-dependent transport problem automatically. By dividing the total history number of each step into many batches and simulating all batches one by one, we calculate the Shannon entropy of the survival particle distribution after each batch. If the on-the-fly diagnostic of convergence of entropy shows the survival particle distribution has converged, the calculation of current step will be stopped in advance and the next step will be activated immediately. Test for an one-dimensional model shows this strategy has decreased the calculation time greatly and keep the results almost unchanged simultaneously.

This is joint work with Li Deng, Gang Li, and Baoyin Zhang.

Thursday 10:30–12:30, **Room: Banque de Développement du Canada**

Stochastic Computation and Complexity IV

Chair: Pawel Przybylowicz

Deterministic, Randomized, and Bayesian Ways to Stop a Simulation

Fred J. Hickernell

Illinois Institute of Technology

hickernell@iit.edu
www.iit.edu/~hickernell

When performing a stochastic simulation, one needs to decide how many observations to make. The sample size, n, may be determined by a time budget. Alternatively, one may wish to choose n based on an error criterion involving an absolute tolerance, ε_{a}, and/or a relative error tolerance, ε_{r}. We present adaptive rules for choosing n to ensure that $|v(\boldsymbol{\mu}) - \hat{v}| \le \max(\varepsilon_{\mathrm{a}}, \varepsilon_{\mathrm{r}}|v(\boldsymbol{\mu})|)$. Our desired answer, $v(\boldsymbol{\mu})$, is a function of one or more means of random variables. Our approximate answer, $\hat{v}$, is based on an IID or low discrepancy sample of size n. The rules for choosing n may be deterministic, random, or Bayesian. We explain the theory and provide numerical examples.

This is joint work with Lluís Antoni Jiménez Rugama and Jagadeeswaran Rathinavel.

Adaptive Multilevel Monte Carlo for Ergodic SDEs with Non-Globally Lipschitz Drift in Infinite Time Interval

Wei Fang

University of Oxford

wei.fang@maths.ox.ac.uk
https://www.maths.ox.ac.uk/people/wei.fang

In this talk, based on [1, 2], we introduce an adaptive timestep construction for an Euler-Maruyama approximation of the ergodic SDEs with a drift which is not globally Lipschitz over an infinite time interval. If the timestep is bounded appropriately, we show not only the stability of the numerical solution and the standard strong convergence order, but also that the bound for moments and strong error of the numerical solution are uniform in T. Following the ideas in [3] and [4], we extend it to the adaptive multilevel Monte Carlo and unbiased schemes for the expectations with respect to the invariant measure. Numerical experiments support our analysis.

This is joint work with Mike Giles.

[1] W Fang and M.B. Giles. Adaptive Euler-Maruyama method for SDEs with non-globally Lipschitz drift: Part I, finite time interval. *arXiv preprint arXiv:1609.08101*, 2016.

[2] W Fang and M.B. Giles. Adaptive Euler-Maruyama method for SDEs with non-globally Lipschitz drift: Part II, infinite time interval. *Working paper in preparation*, 2017.

[3] M.B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.

[4] P. Glynn and C. Rhee. Exact estimation for Markov chain equilibrium expectations. *Journal of Applied Probability*, 51:377–389, 2014.

Strong Approximation of Stochastic Mechanical Systems with Nonlinear Holonomic Constraints

Holger Stroot

TU Kaiserslautern

holger.stroot@itwm.fraunhofer.de

In this talk we present a strong approximation result for a class of high-dimensional stochastic mechanical systems with nonlinear holonomic constraints. Such systems are described by higher-index stochastic differential-algebraic equations, involving an implicitly given Lagrange multiplier process. The explicit representation of the Lagrange multiplier leads to an underlying stochastic differential equation, whose coefficients are in general not one-sided Lipschitz continuous and of super-linear growth. We show strong convergence of a halfexplicit drift-truncated scheme which fulfills the constraint exactly. A concrete example for the considered system are spatially discretized models for the dynamics of inextensible fibers in turbulent flows as occurring, e.g., in the spunbond production process of non-woven textiles.

This is joint work with Felix Lindner and Raimund Wegener.

Optimal Strong Approximation of Cox-Ingersoll-Ross and Squared Bessel Processes

André Herzwurm

TU Kaiserslautern

herzwurm@mathematik.uni-kl.de

We study strong (pathwise) approximation of Cox-Ingersoll-Ross processes. The error criterion is given by the maximal L_1-distance of the solution and its approximation on a compact interval. We propose a Milstein-type scheme that is suitably truncated close to zero, where the diffusion coefficient of the corresponding stochastic differential equation (SDE) fails to be locally Lipschitz continuous. For this scheme we prove polynomial convergence rates for the full parameter range of the corresponding SDE including the accessible boundary regime. In the particular case of a squared Bessel process of dimension $\delta > 0$ the polynomial convergence rate is given by $\min(1/2, \delta/2)$, see [1]. Moreover, we present lower error bounds from [2, 3], which prove the upper bound to be sharp with respect to the rate of convergence.

This is joint work with Mario Hefter and Thomas Müller-Gronbach.

[1] Mario Hefter and André Herzwurm. Strong convergence rates for Cox-Ingersoll-Ross processes – full parameter range. *ArXiv e-prints*, aug 2016.

[2] Mario Hefter, André Herzwurm, and Thomas Müller-Gronbach. Lower error bounds for strong approximation of scalar SDEs with non-Lipschitzian coefficients. *in preparation*, 2017.

[3] Mario Hefter and Arnulf Jentzen. On arbitrarily slow convergence rates for strong numerical approximations of Cox-Ingersoll-Ross processes and squared Bessel processes. *ArXiv e-prints*, feb 2017.

Thursday 10:30–12:30, **Room: Banque CIBC**

Monte Carlo Methods for Molecular Evolution and Phylogenetics

Chair: Liangliang Wang

Phylogenetic Modeling of CpG Hypermutability Using Exact and Approximate Bayesian Computation

Simon Laurin-Lemay

Robert-Cedergren Center for Bioinformatics and Genomics, Department of Biochemistry and Molecular Medicine, Faculty of Medicine, Universite de Montreal, Montreal, Quebec, Canada

evol.simon@gmail.com

Approximate Bayesian Computation (ABC) has increasingly been employed for exploring novel models. In the area of molecular evolution and phylogenetics, however, ABC methods have received little attention, in spite of the pressing interest in richer models of substitution, that would better reflect the true processes governing long-term evolutionary change. Indeed, most of these models have no closed-form likelihood function, rendering classical implementation approaches technically challenging. Here, we explore the use of an ABC method to study a model that accounts for CpG hypermutability, a phenomenon whereby a propensity for deamination of some nucleotides constituting DNA leads to a context-specific, highly elevated mutation rate. We show the reliability of the method using simulations, and apply it to real data to show that CpC hypermutatbility is a major feature in the evolution of protein-coding DNA.

This is joint work with Nicolas Lartillot, Nicolas Rodrigue, and Herve Philippe.

Online Bayesian Phylogenetic Inference via Sequential Monte Carlo

Vu Dinh

Program in Computational Biology, Fred Hutchinson Cancer Research Center, Seattle

vdinh@fredhutch.org

Phylogenetics, the inference of evolutionary trees from molecular sequence data such as DNA, is an enterprise that yields valuable evolutionary understanding of many biological systems. Bayesian phylogenetic algorithms enjoy the flexibility to incorporate a wide range of ancillary model features such as geographical information or trait data which are essential for some applications. However, Bayesian tree inference with current implementations is a computationally intensive task, often requiring days or weeks of CPU time to analyze modest datasets with 100 or so sequences. On the other hand, modern data collection technologies are quickly adding new sequences to already substantial databases. With all current techniques for Bayesian phylogenetics, computation must start anew each time a sequence becomes available, making it costly to maintain an up-to-date estimate of a phylogenetic posterior. These considerations highlight the need for an *online* Bayesian phylogenetic method which can update an existing posterior with new sequences.

In this work, we propose a framework for online Bayesian phylogenetic inference based on Sequential Monte Carlo (SMC) and Markov chain Monte Carlo (MCMC). We first show a consistency result, demonstrating that the method samples from the correct distribution in the limit of a large number of particles. Next we derive the first reported set of bounds on how phylogenetic likelihood surfaces change when new

sequences are added. These bounds enable us to characterize the theoretical performance of sampling algorithms by bounding the effective sample size (ESS) with a given number of particles from below. We show that the ESS is guaranteed to grow linearly as the number of particles in an SMC sampler grows. Surprisingly, this result holds even though the dimensions of the phylogenetic model grow with each new added sequence.

This is joint work with Aaron E. Darling and Frederick A. Matsen.

Parallelizable Monte Carlo Algorithms for Infinite Mixtures in the Detection of Molecular Adaptation

Nicolas Rodrigue

Department of Biology, Institute of Biochemistry, and School of Mathematics and Statistics, Carleton University, Ottawa, Canada

nicolas.rodrigue@carleton.ca

Most implementations of the infinite mixture device known as the *Dirichlet process* rely on a Chinese restaurant representation of the model. However, a Chinese restaurant Markov chain Monte Carlo (MCMC) implementation can only perform updates of the Dirichlet process in a serial manner across a data set; an update for any given observation must be done conditional on the current model configuration for all other observations. Here, we discuss the use of a stick-breaking representation of the Dirichlet process, and how it enables a paralellizable MCMC implementation. We present our application of this technique in the modeling of protein-coding DNA sequence evolution, with the aim of detecting adaptive substitution regimes. Finally, we report results from simulations and real-data analyses highlighting the potential of our approach.

A Sequential Monte Carlo Algorithm for Bayesian Phylogenetics

Liangliang Wang

Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby, BC, Canada

lwa68@sfu.ca

Bayesian phylogenetics, which approximates a posterior distribution of phylogenetic trees, has become more and more popular with the development of Monte Carlo methods. Standard Bayesian estimation of phylogenetic trees can handle rich evolutionary models but requires expensive Markov chain Monte Carlo (MCMC) simulations, which may suffer from two difficulties, the curse of dimensionality and the local-trap problem. Our previous work in [1] has shown that sequential Monte Carlo (SMC) methods can serve as a good alternative to MCMC in posterior inference over phylogenetic trees. In this talk, I will present our recent work on an SMC sampler for general non-clock trees that can incorporate the MCMC kernels from the rich literature of Bayesian phylogenetics. We illustrate our method using simulation studies and real data analysis.

[1] Liangliang Wang, Alexandre Bouchard-Côté, and Arnaud Doucet. Bayesian phylogenetic inference using a combinatorial sequential monte carlo method. *Journal of the American Statistical Association*, 110(512):1362–1374, 2015.

Thursday 10:30–12:30, **Room: Banque Scotia**

Construction of QMC Point sets

Chair: Friedrich Pillichshammer

A Reduced Fast Construction of Polynomial Lattice Point Sets with Low Weighted Star Discrepancy

Mario Neumüller

Johannes Kepler Universität Linz, Austria

mario.neumueller@jku.at

The weighted star discrepancy is a quantitative measure for the performance of point sets in quasi-Monte Carlo algorithms for numerical integration. Polynomial lattice point sets are often a good choice of point sets in this context. This point sets are constructed with the help of a generating vector $(g_1, \dots, g_s)$ which is obtained by a computer search algorithm. The standard way to find generating vectors and ensure a small weighted star discrepancy is the so called component-by-component construction. The computational cost of this algorithm depends linearly on the dimension s.
Our aim is to significantly speed up this procedure and reduce the construction cost of such generating vectors by restricting the size of the sets from which we select the components of the generating vector. To gain this reduction and still ensure a low weighted star discrepancy of the resulting point set we exploit the fact that the weights of the spaces we consider decay very fast. Due to this decay we could prove that the computational cost of our improved algorithm is independent of the dimension eventually.

All three authors are supported by the Austrian Sience Fund (FWF) and are part of the Special Research Program "Quasi-Monte Carlo Methods: Theory and Applications".

This is joint work with Ralph Kritzinger and Helene Laimer.

Enumeration of the Chebyshev-Frolov Lattice Points in Axis-Parallel Boxes

Kosuke Suzuki

Hiroshima University

kosuke-suzuki@hiroshima-u.ac.jp

For a positive integer d, the d-dimensional Chebyshev-Frolov lattice is the lattice in d-dimensional Euclidian space generated by the Vandermonde matrix associated to the roots of the d-dimensional Chebyshev polynomial. It is important to enumerate the points from the Chebyshev-Frolov lattices in axis-parallel boxes when $d = 2^n$ for a non-negative integer n, since the points are used for the nodes of Frolov's cubature formula, which achieves the optimal rate of convergence for many spaces of functions with bounded mixed derivatives and compact support. The existing enumeration algorithm for such points by Kacwin, Oettershagen and Ullrich [1] is efficient up to dimension $d = 16$. In this paper we suggest a new enumeration algorithm of such points for $d = 2^n$, efficient up to $d = 32$.

This is joint work with Takehito Yoshiki.

[1] Christopher Kacwin, Jens Oettershagen, and Tino Ullrich. On the orthogonality of the Chebyshev-Frolov lattice and applications, 2016. arXiv preprint arXiv:1606.00492 [math.NA].

Optimal Discrepancy Rate of Point Sets in Besov Spaces with Negative Smoothness

Ralph Kritzinger

Johannes Kepler University Linz, Austria

ralph.kritzinger@jku.at

We consider the local discrepancy of point sets in the unit square. As a measure for the irregularity of distribution we study the norm of the local discrepancy in Besov spaces with dominating mixed smoothness. Hinrichs proved in [1] that for Hammersley type points this norm has the best possible rate provided that the smoothness parameter of the Besov space is nonnegative. While these point sets fail to achieve the same for negative smoothness, we will demonstrate that the symmetrized versions overcome this drawback. We conclude with several consequences on discrepancy in further function spaces with dominating mixed smoothness and on numerical integration based on quasi-Monte Carlo rules.

[1] Aicke Hinrichs. Discrepancy of Hammersley points in Besov spaces of dominating mixed smoothness. *Math. Nachr.*, 283:478–488, 2010.

Richardson Extrapolation of Polynomial Lattice Rules

Takehito Yoshiki

Osaka City University

tttyoshiki@gmail.com

We study multivariate numerical integration of smooth functions in weighted Sobolev spaces with dominating mixed smoothness $\alpha \geq 2$ defined over the unit cube. We propose a new quasi-Monte Carlo (QMC)-based quadrature rule, named *extrapolated polynomial lattice rule*, which achieves the almost optimal rate of convergence. Extrapolated polynomial lattice rule consists of two steps: i) construction of classical polynomial lattice rules over $\mathbb{F}_b$ with α consecutive sizes of nodes, $N = b^{m-\alpha+1}, \dots, b^m$, and ii) recursive application of Richardson extrapolation to a chain of α approximate values of the integral obtained by consecutive polynomial lattice rules. We prove the existence of good extrapolated polynomial lattice rules achieving the almost optimal order of convergence of the worst-case error in Sobolev spaces with general weights. Then, by restricting to product weights, we show that such good extrapolated polynomial lattice rules can be constructed by the fast component-by-component algorithm under a computable quality criterion, and that the required total construction cost is of the same order as that of interlaced polynomial lattice rule. We also study a dependence of the worst-case error bound on the dimension. In contrast to interlaced polynomial lattice rule, extrapolated polynomial lattice rule has a straightforward application to so-called fast QMC matrix-vector multiplication while still achieving arbitrarily high order of convergence. Numerical experiments for test integrands support our theoretical result.

This is joint work with Josef Dick and Takashi Goda.

Thursday 10:30–12:00, **Room: Ernst & Young (EY)**

Applications in Workforce Planning and Scheduling

Chair: Vassil Alexandrov

Monte Carlo Method for Workforce Planning

Stefka Fidanova

Institute of Information and Communication Technologies, Bulgarian Academy of Science

stefka@parallel.bas.bg

The workforce planning is a difficult optimization problem. It is important real life problem which helps organizations to determine workforce which they need. A workforce planning problem is very complex and needs special algorithms to be solved using reasonable computational resources. The problem consists to select set of employers from a set of available workers and to assign this staff to the tasks to be performed. The objective is to minimize the costs associated to the human resources needed to fulfill the work requirements. A good workforce planing is important for an organization to accomplish its objectives. The mathematical description of the problem is as follows:

A set of jobs $J = \{1, \ldots, m\}$ must be completed during the next planning period. Each job j requires d_j hours during the planing period. There is a set $I = \{1, \ldots, n\}$ of available workers. The availability of the worker i is s_i hours. For reasons of efficiency a worker must perform a minimum number of hours (h_{min}) of any job to which he is assigned and no worker capable assigned to more than j_{max} jobs. A_i is the set of jobs that worker i is qualified to perform. No more than t workers can be assigned during the planed period and the set of selected workers can be capable to complete all the jobs. The goal is to find a feasible solution which minimizes the assignment cost.

The complexity of this problem does not allow the utilization of exact methods for instances of realistic size. Therefore we will apply Ant Colony Optimization (ACO) method which is a kind of Monte Carlo method for solving combinatorial optimization problems. The ACO algorithm was inspired by real ants behavior. An important is how the ants can find the shortest path between food sources and their nest. ACO algorithm is population based approach, which has been successfully applied to solve hard combinatorial optimization problems. One of its main ideas is the indirect communication among the individuals of a colony of agents with distributed numerical information called pheromone. The problem is represented by graph and the ants walk on the graph to construct solutions. The solutions are represented by paths in the graph. After the initialization of the pheromone trails, the ants construct feasible solutions, starting from random nodes, and then the pheromone trails are updated. At each step the ants compute a set of feasible moves and select the best one (according to some probabilistic rules) to continue the rest of the tour.

On this work we prepare variant of ACO algorithm to solve workforce optimization problem. The algorithm is tested on set of test problem. Achieved solutions are compared with other methods.

Evaluation of Aircraft Landing Scheduling Policies Using Monte Carlo Simulation

Wael Sboui

Department of Computer Science and Operations Research, University of Montreal

waelsboui@gmail.com

With the increase in air traffic, airport runway management is becoming a critical issue both at the operational and safety levels. Runways have limited capacity and a certain separation distance must be maintained between two aircrafts during both takeoff and landing phases due to wake turbulence. The minimum separation between two aircrafts depends on their size, influencing their sensitivity to turbulence. As a result, the ordering of takeoffs and landings affects the minimum total waiting time for a sequence of takeoffs or landings, as well as the delays imposed on passengers. Various scheduling policies have been proposed to reduce the delays, usually based on heuristics as the sequencing problem is NP hard. In addition, little attention is usually paid to the random factors influencing take-off and landing times.

We focus on landings as they are more critical than takeoffs and propose a discrete event simulation tool to evaluate different landing scheduling policies, taking account of the operational constraints of separation as well as the uncertainties of the arrival times in the airport area. We then compare several popular policies in the literature, based on various measures of performance. Contrary to the usual assumption of deterministic flight durations, we consider the noise between the expected arrival time and the realized arrival time and examine the impacts on scheduling performance.

This is joint work with Fabian Bastin.

On the Sample Average Approximation of a Two-Stage Staffing Problem with Chance Constraints and Recourse in Call Centers

Thuy Anh Ta

Department of Computer Science and Operations Research (DIRO), University of Montreal
and
GERAD & CIRRELT research centers, Montréal, Canada

tathuyan@iro.umontreal.ca

We consider a two-stage stochastic problem with chance constraints and recourse actions for the staffing of multi-skill call centers in which arrival rates are uncertain. The aim is to minimize the total cost of agents under some chance constraints, defined over the randomness of the service level in a given time period. In the first stage, the initial staffing is determined in advance based on a noisy forecast of the arrival rates. Later (perhaps at the last minute), a more accurate forecast becomes available and the staffing decided earlier can be corrected by recourse actions, by adding or removing agents, with penalty costs. The problem is to determine the initial staffing.

We solve this problem by a Monte Carlo method that generates N scenarios of arrival rates, and performs M simulation runs for each scenario, to estimate the probability that the service level is satisfied, for each call type. This provides a sample average approximation (SAA) of the original problem, which can then be solved by deterministic methods. In this talk, we study the convergence of the SAA problem to the original problem. We show that when N and M are large enough, both problems eventually have the exactly same optimal solutions, and this occurs exponentially fast as a function of N and M. We also report on numerical experiments for solving the SAA problem using a cutting plane methodology.

This is joint work with Wyean Chan, Pierre L'Ecuyer, and Fabian Bastin.

Thursday 15:30–17:30, **Room: Banque de Développement du Canada**

Stochastic computation and Complexity V

Chair: Andreas Neuenkirch

Multilevel Algorithms for Banach Space Valued and Parametric Stochastic Differential Equations

Stefan Heinrich

University of Kaiserslautern

http://www.uni-kl.de/AG-Heinrich/Stefan.html

We study the pathwise approximation of solutions of Banach space valued and of scalar valued parameter dependent stochastic differential equations. The analysis of the Banach space valued case is connected with certain Banach space geometry.

The Banach space case can be related to the parametric case by considering a scale of embedded Banach spaces. This in turn is used to develop and analyze a multilevel method for the solution of the parametric problem. We obtain convergence rates for various smoothness classes of input functions. Furthermore, the optimality of these rates is established by proving matching lower bounds.

Finally we discuss applications to the approximation of functionals of the solution of Banach space valued and scalar valued parametric stochastic differential equations.

Continuous Time Tug-of-War, p-Harmonic Functions, and an Approximation Problem

Stefan Geiss

Department of Mathematics and Statistics, University of Jyväskylä, Jyväskylä, Finland

stefan.geiss@jyu.fi

We consider the p-Laplacian

$$\Delta_p^N u := \frac{p-2}{|Du|^2}\Big(\sum_{i,j=1}^n u_{x_i x_j} u_{x_i} u_{x_j}\Big) + \Big(\sum_{i=1}^n u_{x_i x_i}\Big),$$

where $Du := (u_{x_1}, ..., u_{x_n})'$, and the boundary value problem

$$\Delta_p^N u = 0 \text{ in } U \text{ and } u(x) = g(x) \text{ on } \partial U$$

where $U \subseteq \mathrm{R}^n$, $n \geq 2$, is an open set satisfying certain regularity conditions and $2 < p < \infty$. We relate this problem to a two-player stochastic game - called tug-of-war game with noise - consider space-time discretized controls, and a corresponding approximation problem. The talk is based on joint work in progress.

This is joint work with Christel Geiss and Mikko Parviainen.

[1] K. Nyström and M. Parviainen: Tug-og-War, market manipulation, and option pricing. *Mathematical Finance.*

[2] A. Swiech: Another approach to the existence of value functions of stochastic differential games. *J. Math. Anal. Appl.*, 204(3):884–897, 1996.

Optimal Liquidation Under Partial Information with Price Impact

Michaela Szölgyenyi

Vienna University of Economics and Business

michaela.szoelgyenyi@wu.ac.at

We study the problem of a trader who wants to maximize the expected reward from liquidating a given stock position. We propose a model for the stock price with local characteristics driven by an unobservable finite-state Markov chain and the liquidation rate. This reflects uncertainty about the state of the market and feedback effects from trading.

In this model we solve the optimal liquidation problem and characterise the value function as the unique viscosity solution of the associated HJB equation by applying techniques for piecewise deterministic Markov processes (PDMPs). This allows for a numerical study of the problem. We present numerical results illustrating the impact of partial information and feedback effects on the value function and on the optimal liquidation rate.

This motivates the study of the simulation of PDMPs in general.

This is joint work with Katia Colaneri, Zehra Eksi-Altay, and Rüdiger Frey.

On the Approximation of Tensor Product Operators

David Krieg

Friedrich Schiller University of Jena

david.krieg@uni-jena.de

Let T be a bounded linear operator between two Hilbert spaces. We want to approximate T by an algorithm that evaluates less than a given number n of linear functionals. The minimal worst case error of such an algorithm is given by the nth approximation number of T. If T is a tensor product operator, this is the nth largest number in the tensor product of the sequences of approximation numbers of its factors. I will talk about the asymptotic and preasymptotic behavior of tensor products of sequences of polynomial decay. The results will be applied to the L_2-approximation of mixed order Sobolev functions on the d-cube. It turns out that this problem is much harder for nonperiodic functions than for periodic functions, if $n \leq 2^d$. Asymptotically, however, there is no difference at all. This investigation is inspired by [1] and can be found in [2].

[1] T. Kühn, W. Sickel, T. Ullrich: *Approximation of mixed order Sobolev functions on the d-torus – asymptotics, preasymptotics and d-dependence*. Constructive Approximation 42, 353–398, 2015.

[2] D. Krieg: *Tensor power sequences and the approximation of tensor product operators*. ArXiv e-prints, 2016. arXiv:1612.07680 [math.NA]

Thursday 15:30–17:30, **Room: Banque CIBC**

Statistical Applications of Monte Carlo Methods

Chair: Fred J. Hickernell

Faster Estimates with User-Specified Error for $[0,1]$ Random Variables

Mark Huber

Claremont McKenna College

http://autotomic.wixsite.com/markhubernews

Consider a stream of independent, identically distributed random variables $X_1, X_2, \ldots$ where each X_i falls in $[0,1]$, has mean μ, and variance σ^2. The goal is to construct an estimate $\hat{\mu}$ for μ where the relative error is specified ahead of time by the user. That is, the user gives $\epsilon > 0$ and $\delta > 0$ so that $\hat{\mu} = \hat{\mu}(X_1, X_2, \ldots, X_T)$ satisfies $\mathbf{P}(|\hat{\mu} - \mu| > \epsilon\mu) \leq \delta$ where T is a stopping time that represents how many draws from the X_i were needed. Call such an estimate an (ϵ, δ) randomized approximation scheme (or (ϵ, δ)-ras for short.)

Dagum et. al [1] proved that for $[0,1]$ random variables, there exists a constant c_1 such that any (ϵ, δ)-ras has T with mean at least

$$c_1 \frac{\ln(3/\delta)}{\epsilon^2 \mu^2} \max\{\sigma^2, \epsilon\mu\}.$$

Moreover, they showed that there exists such an (ϵ, δ)-ras with $\mathbf{E}[T]$ at most a second constant c_2 times this lower bound. They did not explicitly state this constant in [1], however, a lower bound on the mean number of steps taken in their algorithm is:

$$\mathbf{E}[T] \geq \frac{\ln(3/\delta)}{\epsilon^2 \mu^2} \left[8(e-2)\max\{\sigma^2, \epsilon\mu\} + 20(e-2)\epsilon\mu\right].$$

In this talk a new method for building an (ϵ, δ)-ras will be given where

$$\mathbf{E}[T] \leq \frac{\ln(3/\delta)}{\epsilon^2 \mu^2} \left[2.86\sigma^2 + 6.72\epsilon\mu\right].$$

This is less than half of the number of samples needed for the previous algorithm, and in practice can be far less.

This is joint work with Bo Jones.

[1] P. Dagum, R. Karp, M. Luby, and S. Ross. An optimal algorithm for Monte Carlo estimation. *Siam. J. Comput.*, 29(5):1484–1496, 2000.

Projected Support Points, with Application to Optimal MCMC Reduction

Simon Mak

Georgia Institute of Technology

smak6@gatech.edu

This talk presents a new point set called *projected support points*, which compacts a continuous probability distribution F into representative points (the original idea of support points is presented at an earlier talk). The primary appeal of projected support points is that it provides an optimal representation not only of the full distribution F, but of its marginal distributions as well. These point sets have many important applications in statistics and engineering, because many practical, high-dimensional sampling or integration problems typically have low-dimensional structure which can be exploited. We present a unifying theoretical framework for projected support points, connecting the desired goodness-of-fit on marginal distributions with important principles in experimental design and Quasi-Monte Carlo. Two algorithms are then proposed for efficient optimization of projected support points, with simulation studies confirming its effectiveness in (a) representing marginal distributions, and (b) integrating high-dimensional functions with low-dimensional structure. An important application of projected support points – as a way to optimally compact Markov-chain Monte Carlo (MCMC) chains – is then highlighted using a Bayesian age-cohort model for breast cancer.

This is joint work with Roshan Joseph Vengazhiyil.

[1] Mak, Simon, and V. Roshan Joseph (2017+). Support points. *Annals of Statistics*, under review. arXiv preprint arXiv:1609.01811.

[2] Mak, Simon, and V. Roshan Joseph (2017+). Projected support points, with application to optimal MCMC reduction. *Journal of the American Statistical Association*, submitted.

An Adaptive Quasi-Monte Carlo Method for Bayesian Inference with User-Specified Error Tolerance

Kan Zhang

Illinois Institute of Technology

kzhang23@hawk.iit.edu

Bayesian inference is based on observed data plus prior beliefs. Computing the expected value of a parameter via Bayesian inference involves the numerical approximation of the quotient of two intractable integrals. In this paper, an adaptive quasi-Monte Carlo(QMC) method is proposed to evaluate this quotient to a user-specified error tolerance. The method is illustrated by a logistic regression model. The efficiency of the computation using two different sampling distributions, and their combination is studied.

This is joint work with Fred J. Hickernell.

Convergence Rates of Control Functional Estimators Based on Stein's Identity

Mark Girolami

Department of Mathematics, Imperial College London

http://www.imperial.ac.uk/people/m.girolami

Gradient information on the sampling distribution can be used to reduce Monte Carlo variance. An important application is that of estimating an expectation along the sample path of a Markov chain, where empirical results indicate that gradient information enables improvement on root-n convergence. This talk introduces Control Functionals, a class of estimator based on Stein's identity. Analysis will be presented describing convergence rates that account for the degree of smoothness of the sampling distribution and test function, and the dimension of the state space. These results provide much- needed insight into the rapid convergence of gradient-based estimators observed for low-dimensional problems, as well as clarifying a curse-of-dimensionality that appears inherent to such methods without further modification.

Thursday 15:30–17:30, **Room: Banque Scotia**

QMC in Special Settings and Applications

Chair: Mathieu Gerber

Quasi-Monte Carlo Simulation of Coagulation and Fragmentation

Christian Lécot

Université Savoie Mont Blanc, CNRS, LAMA, F-73000 Chambéry, France

http://www.lama.univ-savoie.fr/~lecot/

Coagulation and fragmentation models have applications in many domains of science, technology and engineering: aerosol dynamics, nanoparticle generation, crystallization, precipitation, granulation, polymerization, combustion processes, food processes, pollutant formation in flames, microbial systems.

The Monte Carlo method is a powerful tool for solving many problems in the applied sciences. This is a simple, versatile, and robust method but it may suffer from a lack of precision. We explore the quasi-Monte Carlo (QMC) way to improve the accuracy of Monte Carlo simulations of coagulation and fragmentation by replacing the pseudo-random numbers with low discrepancy point sets; the present work extends the method analyzed in [1].

The mass distribution is approximated by a finite number N of numerical particles. Time is discretized and quasi-random points are used at every time step to determine whether each particle is undergoing a coagulation or a fragmentation. In addition, the particles are relabeled according to their increasing mass at each time step. Convergence of the schemes is analyzed when N goes to infinity. Numerical tests show that the computed solutions are in good agreement with analytical ones, when available. And the QMC algorithm reduces the discrepancy of the standard Monte Carlo approach.

This is joint work with Ali Tarhini.

[1] C. Lécot and A. Tarhini. A quasi-Monte Carlo method for the coagulation equation. In G. Larcher, F. Pillichshammer, A. Winterhof, and C. Xing, editors, *Applied Algebra and Number Theory*, pages 216–234. Cambridge University Press, Cambridge, 2014.

Quasi-Monte Carlo for Calculating Multi-Dimensional Expectations Against Gaussian Distribution

Yuya Suzuki

KU Leuven

yuya.suzuki@kuleuven.be

We want to calculate the expectation:

$$\mathbb{E}[g(\boldsymbol{X})] = \int_{\mathbb{R}^d} g(\boldsymbol{x})\phi(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},$$

with a probability density function $\phi(\boldsymbol{x})$ and some interesting function $g(\boldsymbol{x})$ which is dominated by $\phi(\boldsymbol{x})$ when $\boldsymbol{x}$ is far enough from the origin. For instance, let $\boldsymbol{X}$ be a standard normal random variable (i.e., $\phi(\boldsymbol{x}) = \exp(-\|\boldsymbol{x}\|_2^2/2)/\sqrt{2\pi}^d$) and

$$g(\boldsymbol{x})\phi(\boldsymbol{x}) \lesssim \phi(\boldsymbol{x}) = \frac{\exp(-\|\boldsymbol{x}\|_2^2/2)}{\sqrt{2\pi}^d},$$

where $\lesssim$ means there exists a constant c satisfying $g(\boldsymbol{x})\phi(\boldsymbol{x}) \leq c\phi(\boldsymbol{x})$ for all $\boldsymbol{x}$ far enough from the origin.

Integration on $\mathbb{R}^d$ by Quasi-Monte Carlo (QMC) often truncates the domain to a hypercube but it can suffer from the curse of dimensionality. Another alternative is using the inverse transformation method but this is not always QMC-friendly. Here we propose to truncate $\mathbb{R}^d$ depending on the dominating term which is $\exp(-\|\boldsymbol{x}\|_2^2/2)$ in the above example. In that case, we truncate the domain to $\|\boldsymbol{x}\|_2 \leq R$, and then apply a QMC rule to the truncated integral. We illustrate this by some numerical examples such as option pricing.

This is joint work with Dirk Nuyens.

A New Rotation Invariant Sampling Design on the Sphere

Javier Gonzalez-Villa

Universidad de Cantabria

javier.gonzalezv@alumnos.unican.es

We address the problem of estimating the integral of a measurement function defined on the unit sphere. The nucleator [1, 2], based on an isotropic ray emanating from a fixed point, is an unbiased sampling tool used for instance in biology to estimate cell volume [3]. To predict the variance of the nucleator is a nontrivial problem. In [4], a systematic sampling design was proposed with corresponding variance prediction formulae. Their performance, however, has recently been shown to be poor [5].

Drawing some ideas from [6], we propose a different design which, in particular, allows genuine systematic sampling at the vertices of the Platonic solids. We also give prediction formulae whose performance is checked empirically with the aid of automatic Monte Carlo simulations on computer renderings of real objects.

This is joint work with Domingo Gómez-Pérez and Luis M. Cruz-Orive.

[1] Hans-Jørgen G Gundersen. The nucleator. *J. Microsc.*, 151:3–21, 1988.

[2] Eva B. Vedel Jensen. *Local Stereology.* World Scientific, 1998.

[3] Mark J. West. *Basic Stereology for Biologists and Neuroscientists.* Cold Spring Harbor Laboratory Press, 2012.

[4] Ximo Gual-Arnau and Luis M. Cruz-Orive. Variance prediction for pseudosystematic sampling on the sphere. *Adv. Appl. Probab.*, 34:469–483, 2002.

[5] Javier Gonzalez-Villa, Marcos Cruz, and Luis M. Cruz-Orive. On the precision of the nucleator. *Image Anal. Stereol.*, (in press), 2017.

[6] Aicke Hinrichs and Jens Oettershagen. Optimal point sets for quasi-Monte Carlo integration of bivariate periodic functions with bounded mixed derivatives. *Springer Proc. Math. Stat.*, 163:385–405, 2016.

Density Estimation by Randomized Quasi-Monte Carlo

Amal Ben Abdellah

DIRO, Université de Montréal, Canada

amalbenabdellah@gmail.com

Randomized quasi Monte Carlo (RQMC) is commonly used to estimate the mathematical expectation of some random variable Y, written as an integral over the s-dimensional unit cube $(0,1)^s$. Under certain conditions, the RQMC estimator converges at a faster rate than a crude Monte Carlo estimator of the integral.

In this talk, to extend the range of application of RQMC, we examine how it can improve the convergence rate when estimating the entire distribution of Y, not only its expectation. We consider estimating the cumulative density function by its empirical version, as well as density estimation by histograms, averaged shifted histograms, and kernel density estimators. We provide both theoretical and empirical results.

This is joint work with Pierre L'Ecuyer and Art B. Owen.

Thursday 15:30–17:30, **Room: Ernst & Young (EY)**

Applications in Business and Operations Management

Chair: Johan Van Kerckhoven

A Sequential Design for Gas Storage Optimization Using Kriging Metamodels

Stephan Prell

Chair for Energy Trading and Finance, University of Duisburg-Essen

Stephan.Prell@uni-due.de

Until lately, considerably less attention has been payed to optimal switching problems using regression Monte Carlo methods. An area where this kind of problems arise frequently, is the valuation of operational flexibility of exotic energy derivatives such as power plants, swing options or gas storages. We introduce a new regression Monte Carlo algorithm to value gas storages. We reformulate the optimal control problem as a generalization of a multiple stopping problem, which boils down to a classification problem of ranked value functions. Moreover, we take the cost of sampling and regression into consideration by adopting sequential space-filling designs. [1] exposes the benefits of stochastic kriging as a flexible, nonparametric regression approach in the light of sequential sampling, which will be used to estimate the value function. The benefit of that framework is the smart sampling of highly complex value functions, where a priori little is known about the exact decision boundaries, i.e. the intersections of the value functions, and the efficient updating of the regression.

[1] M. Ludkovski. Kriging metamodels and experimental design for bermudan option pricing. *Journal of Computational Finance (to Appear)*, 2016.

Solving the Territorial Design Problem for Business Sales Plan Using Monte Carlo Method

Laura Hervert-Escobar

Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, Nuevo León, México
and
ICREA, Barcelona Supercomputing Center, Barcelona, Spain

laura.hervert@itesm.mx

A well designed territory enhances customer coverage, increases sales, fosters fair performance and rewards systems and lower travel costs. This paper considers a real life case study to design the sales territory for a business sales plan. The business plan consists of assigning the optimal quantity of sellers to a given territory and includes also the scheduling and routing plans for each seller. The problem is formulated as a combination of assignment, scheduling and routing optimization problems. While these set of problems can be solved using deterministic algorithms, in this paper we consider and apply a Monte Carlo method which is able to tackle larger problems more efficiently. Also a comparison between the efficiency of Monte Carlo method and the deterministic approaches is made. Several real life instances of different sizes were tested with the methods on data that represent raise/fall in the customerÂ´s demand as well as the appearance/ loss of clients.

This is joint work with Vassil Alexandrov.

Better Together? Performance Dynamics in Retail Chain Expansion Before and After Mergers

Nathan Yang

McGill University, Montreal, Québec, Canada

nathan.yang3@mcgill.ca

This paper evaluates how mergers affect the performance efficiency of retail chains. We estimate a dynamic model of retail expansion using data on convenience-store chains in Japan before and after an actual merger event. Our estimation leverages recent advances in particle filtering and structural econometrics to allow for the presence of performance efficiency, in the form of serially correlated state variables that evolve both endogenously and stochastically. The estimates reveal that although the merged firm benefited from lower expansion costs, underlying performance efficiency for the merged entity did not improve following the merger, and such changes in performance varied across markets. Simulation analysis reveals the dampened performance is associated with the merged firm's diminished ability to retain efficiency gains from one year to the next. However, these negative effects can be mitigated if the merged firm inherits the primitives behind the performance efficiency of the more dominant merging party.

This is joint work with Mitsukuni Nishida.

Networking Simulation Results Across Organizational Boundaries

Eric Torkia

Executive Partner–Analytics Practice, Technology Partnerz Ltd., Montreal, Quebec, Canada

etorkia@technologypartnerz.com
www.technologypartnerz.com

One of the major challenges Monte Carlo simulation is that the applications tend to be focused on a specific problem or area in the organization. Of course, the true power of this analysis needs to be shared and used as inputs to other models otherwise there is the potential of missing out on key insights because important models are not being aggregated from all areas of the organization.

According to Savage, Sholtes and Zweidler (2006), by vectorizing the results of a Monte-Carlo Simulation, they can be aggregated, rolled-up or manipulated mathematically without requiring the re-run of large models. This approach allows for, through a structured approach, the ability create company-wide meta-models that combine all the insight from sub-models created in other parts of the organization. This notion of networking multiple models also extends outside the firm. Organizations could and can integrate probabilistic data from outside entities (such as government agencies, banks, open data repositories, etc.) to enrich their model with "real" data.

This presentation will cover the why and how of networking Monte-Carlo models using the SIP (Stochastic Information Packet) standard, including a practical example.

This is joint work with Sam Savage.

[1] Sam Savage, Stefan Scholtes, and Daniel Zweidler. Probability Management. *OR/MS Today*, 33(1), Feb. 2006. http://probabilitymanagement.org/library/Probability_Management_Part1s.pdf.

[2] Sam Savage and Marc Thibault. Towards a Simulation Network–or–The Medium is the Monte Carlo (with apologies to Marshall McLuhan). *Proceedings of the Winter Simulation Conference*, 2015. http://www.informs-sim.org/wsc15papers/467.pdf.

[3] Sam Savage, Farshad Miraftab, Melissa Kirmse, Christine Cowsert Chapman, and Jordan Allen. Probability Management: Rolling up operational risk at PG&E. *OR/MS Today*, December 2016.

Friday 9:00–10:30, **Room: Banque de Développement du Canada**

Accelerated Monte Carlo in Optimization statistics, and PDEs with Random Imput

Chair: Raghu Pasupathy

Adaptive MCMC For Everyone

Jeffrey Rosenthal

University of Toronto

http://probability.ca/jeff/

Markov chain Monte Carlo (MCMC) algorithms, such as the Metropolis Algorithm and the Gibbs Sampler, are an extremely useful and popular method of approximately sampling from complicated probability distributions. Adaptive MCMC attempts to automatically modify the algorithm while it runs, to improve its performance on the fly. However, such adaptation often destroys the ergodicity properties necessary for the algorithm to be valid. In this talk, we first illustrate MCMC algorithms using simple graphical Java applets. We then discuss adaptive MCMC, and present examples and theorems concerning its ergodicity and efficiency. We close with some recent ideas which make adaptive MCMC more widely applicable in broader contexts.

The Adaptive Sampling Gradient Method: Optimizing Smooth Functions with an Inexact Oracle

Raghu Pasupathy

Purdue Statistics, Virginia Tech ISE

pasupath@purdue.edu

Consider *stochastic optimization* settings where a smooth objective function f is unknown but can be estimated with an *inexact oracle* such as quasi-Monte Carlo (QMC) or numerical quadrature. The inexact oracle yields function estimates whose error decays with increasing oracle effort. For solving such problems, we present the derivative-free Adaptive Sampling Gradient Method (ASGM) in two flavors depending on whether the step size used within ASGM is constant or determined through a backtracking line search. ASGM's salient feature is the adaptive manner in which it constructs gradient estimates, by exerting just enough oracle effort at each iterate to keep the error (in the gradient approximate) within a constant factor of the norm of the gradient approximate. We show that both flavors of ASGM exhibit global convergence. We also prove two sets of results on ASGM's *work complexity* with respect to the gradient norm: (i) when f is smooth, ASGM's work complexity is arbitrarily close to $\mathcal{O}(\epsilon^{-2-\frac{1}{\mu(\alpha)}})$, where $\mu(\alpha)$ is the error decay rate of the gradient estimate expressed in terms of the error decay rate α of the objective function estimate; (ii) when f is smooth and strongly convex, ASGM's work complexity is arbitrarily close to $\mathcal{O}(\epsilon^{-\frac{1}{\mu(\alpha)}})$. We illustrate the calculation of α and $\mu(\alpha)$ for common choices, e.g., QMC with finite difference gradients.

This is joint work with F. S. Hashemi and M. R. Taaffe.

Rare Event Analysis and Efficient Simulation for Random Elliptic PDEs with Small Noise

Xiaoou Li

University of Minnesota

lixx1766@umn.edu

Partial differential equations with random inputs have become popular models to characterize physical systems with uncertainty coming from, e.g., imprecise measurement and intrinsic randomness. We perform asymptotic rare event analysis and importance sampling for such elliptic PDEs with random inputs. In particular, we consider the asymptotic regime that the noise level converges to zero suggesting that the system uncertainty is low, but does exist. We develop sharp approximations of the probability of a large class of rare events and propose an efficient importance sampling algorithm.

This is joint work with Jingchen Liu, Jianfeng Lu, and Xiang Zhou.

Friday 9:00–10:30, **Room: Banque CIBC**

Advanced Monte Carlo Methods in Non-Linear Finance

Chair: Emmanuel Gobet

MCMC Design-Based Non-Parametric Regression for Rare-Event: Application to Nested Risk Computations

Gersende Fort

IMT, CNRS & Université de Toulouse, in Toulouse, France

gersende.fort@math.univ-toulouse.fr

Motivated by financial risk computational problems, we propose a new algorithm for a Monte Carlo-based approximation of quantities of the form

$$\mathbb{E}\left[f(Y, \mathbb{E}\left[R|Y\right])|Y \in \mathcal{A}\right] \tag{1}$$

where $\{Y \in \mathcal{A}\}$ is a rare event and exact sampling from the conditional distribution of R given Y is possible: the outer expectation is approximated by the output $\{Y_1, \cdots, Y_M\}$ of a Markov chain Monte Carlo sampler which is designed to be robust to the rare event setting [2]; the inner expectation is computed through a non parametric least-square regression approach by using M draws $\{R_1, \cdots, R_M\}$ such that the conditional distribution of R_i given Y_i is the conditional distribution of R given Y.

We also provide convergence analyses of this new algorithm: we establish non asymptotic bounds for the L_2-empirical risk associated to the least-squares regression; this generalizes the error bounds usually obtained in the case of i.i.d. observations $\{Y_1, \cdots, Y_M\}$. Global error bounds are also derived for the approximation of the nested expectation problem (1).

Finally, through numerical applications, we will discuss the role of some design parameters of our algorithm on the efficiency of the approximation of (1).

This is joint work with Emmanuel Gobet and Eric Moulines. [1]

[1] G. Fort, E. Gobet, and E. Moulines. MCMC design-based non-parametric regression for rare event. Application to nested risk computation. *Monte Carlo Methods and Applications*, 23(1):21–42, 2017.

[2] E. Gobet and G. Liu. Rare event simulation using reversible shaking transformations. *SIAM Scientific Computing*, 37(5):A2295–A2316, 2015.

Sequential Design for Estimating Value-at-Risk

Mike Ludkovski

Department of Statistics and Applied Probability, UC Santa Barbara

ludkovski@pstat.ucsb.edu

The implementation of Solvency II necessitates an accurate way of estimating value-at-risk (VaR) of future losses. Loss distributions can rarely be modeled precisely, leading to simulation approaches. This requires nested simulation, i.e. allocation of simulation budget across the outer scenarios, with the inner simulations used to approximate the respective portfolio loss. Efficiency demands focusing the efforts on the high-loss scenarios, leading to adaptive, sequential approaches.

In this talk I will describe our investigation of using Gaussian Process (GP) metamodels to construct a statistical surrogate for expected portfolio losses. In turn the GP metamodel is used to construct greedy sequential design methods that allocate remaining simulation budget to scenarios that most improve the VaR estimate. Simultaneously we use the GP surrogate to construct low-bias probabilistic estimates/credible intervals for VaR and TVaR. We compare our fully sequential spatial-modeling approach to existing two-stage designs [2] and parametric least-squares Monte Carlo techniques [1]. Examples from option pricing portfolios and equity-linked annuity portfolios will be presented.

This is joint work with Jimmy Risk.

[1] Mark Broadie, Yiping Du, and Ciamac C Moallemi. Risk estimation via regression. *Operations Research*, 63(5):1077–1097, 2015.

[2] Ming Liu and Jeremy Staum. Stochastic kriging for efficient nested simulation of expected shortfall. *The Journal of Risk*, 12(3):3, 2010.

Hedging with Non-Quadratic Local Risk Minimization Using Least-Squares Monte Carlo

Isaque Pimentel

Ecole Polytechnique, EDF

pimentel.isaque@gmail.com

In this talk, we will present a hedging method via a local risk-minimization with an *asymmetric strongly convex risk function* in a complete market where the investor need to follow a sequence of intermediate targets given by the regulator. In discrete time, we characterize the optimal solution using *least-squares Monte-Carlo*. In a continuous time framework, we deal with a financial market in a Markovian structure, in which we characterize the optimal strategy in terms of a *PDE* with nonlinearity on the second derivative.

This is joint work with Emmanuel Gobet and Xavier Warin.

Friday 9:00–10:30, **Room: Banque Scotia**

Monte Carlo simulation Using SimJulia

Chair: Fabian Bastin

SimJulia: The Good, the Bad and the Ugly

Nicolas Andriessen

Royal Military Academy (Brussels)

ben.lauwens@rma.ac.be

SimJulia is a combined continuous time / discrete event process oriented simulation framework written in Julia[1] inspired by the Simula library DISCO[2] and the Python library SimPy[3]. In this presentation the internals of SimJulia are dissected.

Firstly, the interactions between de scheduler, the events and its callback functions are presented. Events are entities that can happen, are happening or have happened. They can be time bounded or they can be triggered by internal or external state changes. The scheduler select the next event to be happening based on trigger time, priority and event creation time. The happening of an event is nothing else than the execution of its callback functions.

SimJulia is however a process oriented framework freeing its users from the burden of interacting directly with events, so secondly processes are introduced. Processes are defined as semi-coroutines, i.e. calling a process function multiple time will resume the function where it left the previous time. This allows to model simulated entities by focusing on its logic: client goes to a shop, has to wait before being served, is getting served, pays and leave the shop. Other clients or the working of the shop will not interfere with the logic of the client but during the interaction these will have an impact on the shopping experience of our client.

Julia had some nice features allowing to implement in a seamless way processes: `Tasks` and the `consume/produce` functions. However in Julia v0.6 the `consume/produce` functions are deprecated. The overhead of the proposed solution based on `Channels` is huge and can therefor not be used as task switching method in SimJulia. So a macro is developed that allows a function to keep its state between function calls. It turns out that this is a lot faster than the previous implementation. The main disadvantage is that only the process function can do a `@yield return`. This has a minor impact on the existing API. Two less important issues remain: a `@yield return` in a `try/catch` clause or in a `do` clause are not yet possible.

This is joint work with Ben Lauwens.

[1] Jeff Bezanson, Stefan Karpinski, Viral B Shah, and Alan Edelman. Julia: A fast dynamic language for technical computing. *arXiv preprint arXiv:1209.5145*, 2012.

[2] Keld Helsgaun. Disco - a simula-based language for continuous combined and discrete simulation simulation software. *Simulation*, 35(1):1–12, 1980.

[3] Norm Matloff. Introduction to discrete-event simulation and the simpy language. *CA. Dept of Computer Science. University of California at Davis*, 2, 2008.

Simulation Based Manpower Planning: An Introduction Using SimJulia

Johan Van Kerckhoven

Koninklijke Militaire School

johan.vankerckhoven@rma.ac.be

Human resource management becomes more and more important. Not only do managers have to balance the needs of their organisation with the needs of their employees, but they also need to make well-informed manpower planning decisions. One part of manpower planning covers the short term, for example, setting up employee rosters such that the workload is appropriately distributed over the employees. The long term on the other hand concerns hiring and promotion strategies to meet the company's goals without alienating your employees, or to compensate for external policy changes such as an increase in the minimum retirement age.

Especially in the current day and age, these decisions are too important to be left to the manager's "gut feeling". Instead, he will rely on a mathematical model to provide the necessary insights to make appropriate decisions. Several types of models are possible here: Markov models, optimisation models using mathematical programming, simulation models, and system dynamics models. Each of these models can be used to model the transient state or the steady state of the system, and has its particular advantages and drawbacks for these tasks (see Jun Wang (2005)[1] for an overview).

Here we concern ourselves with a discrete event simulation model. In particular, we wish to illustrate, by means of a simple example, how we can use the programming language Julia[2], and in particular the SimJulia module[3], to develop such a simulation. We also wish to shed some light on the development goals, both in the general, as in what we want to achieve, and specific, to which purpose(s) we will apply these results, sense.

This is joint work with Ben Lauwens, Oussama Mazari-Abdessameud, and Filip Van Utterbeeck.

Simulation of Medical Response to Disasters Using SimJulia

Selma Koghee

Department of Mathematics, Royal Military Academy, Brussels, Belgium

selma.koghee@mil.be

In this talk a new implementation of the SIMEDIS model will be presented. The SIMEDIS (simulation for the assessment and optimization of medical disaster management) simulates the pre-hospital medical response in disaster scenarios. The aim of the research is to develop a model that provides evidence for the optimal medical response in various disaster scenarios. The simulation involves the creation of victims using a database of victim profiles, search and rescue, arrival ans possible relocation of medical teams, arrival of ambulances, the triage, treatment, and transport of victims and possible deterioration of the victims health state. The victims health states provide the main indicator of the effectiveness of the medical response. In the previous studies [1, 2], two scenarios have been implemented in Arena, a commercially available software program. Here, the model has been implemented in Julia, using SimJulia. The scenario is very similar to that of the first model [1], although the procedures corresponding to the two possible rescue policies have been adapted such that they follow the second model [2] more closely. Furthermore, the current model allows for variation in the victim sets and provides additional output

[1] Jun Wang, *A Review of Operations Research Applications in Workforce Planning and Potential Modelling of Military Training*, 2005.

[2] `http://julialang.org/`

[3] `https://github.com/BenLauwens/SimJulia.jl`

regarding the use of resources. The current implementation forms a proof of principle and can be extended to more complicated disaster scenarios.

This is joint work with Johan Van Kerckhoven and Filip Van Utterbeeck.

[1] Ullrich, C., Van Utterbeeck, F., Dejardin, E., Dhondt, E, and Debacker, M. Pre-hospital simulation model for medical disaster management. *Proceedings of the 2013 Winter Simulation Conference*, Pasupathy, R., Kim, S.-H., Tolk, A., Hill, R., and M. E. Kuhl, M.E. (eds.), pp. 2432–2443 (2013).

[2] Debacker, M., Van Utterbeeck, F., Ullrich, C., Dhondt, E., and Hubloue, I.. SIMEDIS: a discrete event simulation model for testing responses to mass causality Incidents. *Journal of Medical Systems*, 40, 273 (2013).

Friday 9:00–10:30, **Room: Ernst & Young (EY)**

Parallel computation and codes

Chair: Ivan Dimov

On Efficient Parallel Monte Carlo and Quasi-Monte Carlo Hybrid Methods for Matrix Computations

Vassil Alexandrov

ICREA, Barcelona Supercomputing Center, Barcelona, Spain
and
Instituto Tecnológico y de Estudios Superiores de Monterrey, Monterrey, Nuevo León, México

vassil.alexandrov@bsc.es

This paper focuses on the latest advances in designing hybrid, e.g. stochastic/deterministic, Monte Carlo and quasi-Monte Carlo methods for Linear Algebra problems with minimum communication in the parallel implementation. It presents an enhanced version of a stochastic SParse Approximate Inverse (SPAI) preconditioner for general matrices. This method is used in contrast to the standard deterministic preconditioners computed by the Modified SParse Approximate Inverse Preconditioner (MSPAI). Thus we present a Monte Carlo and quasi-Monte Carlo methods to compute a rough approximate matrix inverse first, which can further be optimized by an iterative filter process and a parallel refinement, to enhance the accuracy of the inverse and the preconditioner respectively. The proposed approach allows efficient minimization of communications in the parallel implementation of Monte Carlo and quasi-Monte Carlo methods for Linear Algebra and thus improving the overall performance. Several approaches are implemented, including producing a set of small number of covering Markov chains which are much longer that the usually produced ones.

The overall advantage of the proposed approach is that finding the sparse Monte Carlo or quasi-Monte Carlo matrix inversion has a computational complexity linear of the size of the matrix, it is inherently parallel and thus can be obtained very efficiently for large matrices and can be used also as an efficient preconditioner while solving systems of linear algebraic equations. The proposed hybrid approach uses the so obtained preconditioner in combination with GMRES or any other efficient parallel iterative solver to solve the corresponding system of linear algebraic equations. A comparison of the efficiency of the new approach in case of Sparse Approximate Matrix Inversion and hybrid Monte Carlo and quasi-Monte Carlo methods for solving Systems of Linear Algebraic Equations is carried out.

The behaviour of the proposed algorithms is studied and their performance measured, evaluated and compared with MSPAI together with the efficiency of hybrid methods on a selection of test matrices from matrix market selection of matrices and some matrices from real life problems. The numerical experiments have been executed on the MareNostrum III supercomputer at the Barcelona Supercomputing Center (BSC).

This is joint work with Oscar A. Esquivel-Flores and Aneta Karaivanova.

Optimal Implementation of Quasi-Monte Carlo Methods for Matrix Computations on Intel MIC Architecture

Sofiya Ivanovska

Institute for Information and Communication Technologies,
Bulgarian Academy of Sciences, Sofia, Bulgaria

sofia@parallel.bas.bg

The tightened energy requirements when designing state-of-the-art high performance computing systems lead to the increased use of computational accelerators. Intel introduced the MIC architecture for their line of accelerators and successfully competes with NVIDIA on basis of price/performance and ease of development. Although some codes may be ported succesfully to Intel MIC architecture without significant modifications, in order to achieve optimal performance one has to make the best use of the vector processing capabilities of the architecture. When low-discrepancy sequences are used for Monte Carlo matrix computations, it is important to strike balance between accuracy and speed of generation. In this work we present our implementation of Quasi-Monte Carlo methods for matrix computations specifically optimised for the Intel Xeon Phi accelerators. To evaluate the performance of our implementaiton, we consider not only the problem of solving linear equations but also the problem of finding eigenvalues of the matrix. To achieve optimal parallel efficiency we make use of both MPI and OpenMP, since such kind of problems are usually solved using multiple servers. Using established test cases we compare the Sobol and Halton low-discrepancy sequences with pseudorandom number generators from point of view of speed and accuracy. The energy efficiency of the algorithms is also studied.

This is joint work with Emanouil Atanassov, Aneta Karaivanova, and Todor Gurov.

JMCT: A 3D Monte Carlo Particle Transport Code

Gang Li

Institute of applied physics and computational mathematics, Beijing, China

li_gang@iapcm.ac.cn

JMCT, a new Monte Carlo particle transport code is introduced in detail. First is the four special features of JMCT: high resolution numerical reactor simulations, high efficient use of computers and massive parallelization, visual integration of pre- and post-processing, realization based on field framework and subjects separation. Second is the wide verification and application of JMCT, including numerical simulations of commercial reactor models such as Daya Bay, BEAVRS, QS-I, QS-II and experimental facilities like the SG-III Laser facility. JMCT is designed to provide precise models and simulations of nuclear reactor physics and geometric materials in nuclear field, and meet the demand of high resolution numerical simulations of particle transport and high performance parallel processing, and seek for systematic methods for analysis of reactor physics and shielded designs. Besides, it can be used to provide referenced verifications for deterministic codes and algorithms.

This is joint work with Li Deng, Baoyin Zhang, Rui Li, Danhua Shangguan, Dunfu Shi, Yan Ma, Lingyu Zhang, and Yuanguang Fu

Friday 10:50–11:50, **Room: Banque de Développement du Canada**

Monte Carlo Simulation and its Applications in Stochastic Dynamic Programming

Chair: Fabian Bastin

A Primal-Dual Iterative Monte Carlo Method for Stochastic Dynamic Programs and Its Applications in Finance

Ma Xiang

Dept. of Systems Engineering and Engineering Management,
The Chinese University of Hong Kong, Hong Kong

xma@se.cuhk.edu.hk

In this paper we use the information relaxation technique to develop a value-and-policy iterative method to solve stochastic dynamic programming problems. Each iteration generates a confidence interval estimate for the true value function and a corresponding suboptimal policy so that we can use the gap between the upper and lower bounds to access the quality of the policy. We show that the resulted sequences of suboptimal policies will converge to the optimal one within finite number of iterations through our method.

A regression-based Monte Carlo algorithm is introduced to overcome the dimensionality curse in the implementation of this approach for high dimensional cases. Our formulation reduces the original problem to solving a sequence of open loop control problems. We can thereby rely on a variety of well-developed deterministic optimization algorithms to accelerate the computational speed. It is different from the traditional literature of approximate dynamic programs where a majority of methods need to solve stochastic optimization problems. As numerical illustrations, we apply the algorithm to the optimal order execution problem and the portfolio selection problems. Some new insights about optimal value and optimal policy are also discussed.

This is joint work with Nan Chen.

MCMC Methods for Dynamic Stochastic Optimization

John R. Birge

The University of Chicago Booth School of Business, Chicago, USA

john.birge@chicagobooth.edu

Dynamic stochastic optimization generally suffers from the curse of dimensionality as state spaces grow exponentially in dimension and in the number of periods. Particle methods for filtering and smoothing, however, maintain a fixed number of states in each period and can converge to a posterior distribution using Markov Chain Monte Carlo methods. This talk will discuss how this approach can be applied in the context of dynamic stochastic optimization and conditions for convergence to an optimal solution.

Friday 10:50–11:50, **Room: Banque CIBC**

Finance Applications

Chair: Gersende Fort

Calibration and Monte Carlo Pricing Under a Hybrid Local-Stochastic Volatility Model

Andrei Cozma

Mathematical Institute, University of Oxford

andrei.cozma@maths.ox.ac.uk

The efficient pricing and hedging of vanilla and exotic foreign exchange options requires an adequate model that takes into account both the local and the stochastic features of the volatility dynamics. We put forward a four-factor hybrid local-stochastic volatility (LSV) model that combines state-of-the-art dynamics for the exchange rate with stochastic dynamics for the domestic and foreign short rates, and provide a consistent and self-contained calibration and pricing framework.

For the calibration, we propose a novel and generic algorithm that builds on the particle method of Guyon and Labordere. We combine it with new variance reduction techniques to accelerate convergence and use control variates derived from a pure local volatility model, the stochastic short rates and the two-factor Heston-type LSV model with deterministic rates. Our numerical experiments show a dramatic variance reduction that allows us to calibrate the four-factor model at almost no extra computational cost.

For the pricing, we propose a simple and efficient Monte Carlo simulation scheme that combines the log-Euler scheme for the exchange rate with a modified full truncation Euler (FTE) scheme for the stochastic volatility and the stochastic short rates. We find a lower bound on the explosion time of exponential moments of FTE approximations and establish, up to a critical time, the uniform boundedness of moments of order higher than 1 of approximation schemes for Heston-type (LSV) models, a result which plays a key role in the convergence analysis and which has been a long-standing open problem until now. We prove the strong convergence (without a rate) of the exchange rate approximations and the convergence of Monte Carlo estimators for a number of vanilla and exotic options. We then carry out numerical tests to justify our choice of model and to demonstrate convergence. Finally, under slightly different model dynamics and stronger model assumptions, we establish an optimal strong convergence order of 1/2 (up to a logarithmic term) of the approximation scheme.

This is joint work with Matthieu Mariapragassam and Christoph Reisinger.

Extracting Latent States from High Frequency Option Prices

Geneviève Gauthier

HEC Montreal and GERAD

https://www.gerad.ca/fr/people/genevieve-gauthier

We propose the realized option variance as a new observable variable to integrate high frequency option prices in the inference of option pricing models. Using simulation and empirical studies, this presentation documents the incremental information offered by this realized measure. Our empirical results show that the information contained in the realized option variance improves the inference of model variables such as the instantaneous variance and variance jumps of the S&P 500 index. Parameter estimates indicate that the risk premium breakdown between jump and diffusive risks is affected by the omission of this information.

This is joint work with Diego Amaya and Jean-Francois Bégin.

List of participants

Participants registered by June 14, 2017:

Massil Achab
École Polytechnique
France
massil.achab@polytechnique.edu

Vassil Alexandrov
ICREA-BSC, ITESM
Spain
vassil.alexandrov@bsc.es

James Allison
North-West University
South Africa
james.allison@nwu.ac.za

Christian Andersson Naesseth
Linköping University
Sweden
christian.a.naesseth@liu.se

Nicolas Andriessen
Royal Military Academy, Belgium
Belgium
nicolas_andriessen@hotmail.com

David Barrera
École Polytechnique
France
jdbarrer@gmail.com

Fabian Bastin
Université de Montréal
Canada
bastin@iro.umontreal.ca

Mylène Bédard
Université de Montréal
Canada
bedard@dms.umontreal.ca

Amal Ben Abdellah
Université de Montréal
Canada
amalbenabdellah@gmail.com

Nadhir Ben Rached
Computer, Electrical, and Mathematical Science and Engineering (CEMSE) Division King Abdullah University of Science and Technology (KAUST)
Saudi Arabia
nadhir.benrached@kaust.edu.sa

Krzysztof Bisewski
Centrum Wiskunde & Informatica
Netherlands
bisewski@cwi.nl

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

Maksym Byshkin
Università della Svizzera italiana
Switzerland
maksym.byshkin@usi.ch

Antonio Cabeda
HEC Montréal
Canada
antonio.cabeda@hec.ca

James Calvin
New Jersey Institute of Technology
USA
calvin@njit.edu

Michael Chiu
University of Toronto
Canada
chiu@cs.toronto.edu

Andrei Cozma
University of Oxford
United Kingdom
andrei.cozma@maths.ox.ac.uk

Radu Craiu
University of Toronto
Canada
craiu@utstat.toronto.edu

Sen Deborshee
National University of Singapore
Singapore
deborshee.sen@u.nus.edu

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

Ivan Dimov
IICT, Bulgarian Academy of Sciences
Bulgaria
ivdimov@bas.bg

Dong Ding

United Kingdom
dong.ding10@imperial.ac.uk

Vu Dinh
Fred Hutchinson Cancer Research Center
USA
vdinh@fredhutch.org

Paul Dupuis
Brown University
USA
Paul_Dupuis@brown.edu

Joseph Durante
Princeton University
USA
jdurante@princeton.edu

Víctor Elvira
IMT Lille Douai
France
victor.elvira.arregui@gmail.com

Wei Fang
University of Oxford
United Kingdom
fangwei.pku10@gmail.com

Henri Faure
Aix Marseille Université
France
henri.faure@univ-amu.fr

Michael Feischl
UNSW Sydney
Australia
m.feischl@unsw.edu.au

Stefka Fidanova
Institute of information and communication technologies
Bulgaria
stefka@parallel.bas.bg

Gersende Fort
CNRS
France
gersende.fort@math.univ-toulouse.fr

Jordan Franks
University of Jyväskylä
Finland
jordan.j.franks@jyu.fi

Philippe Gagnon
Université de Montréal
Canada
gagnonp@dms.umontreal.ca

Geneviève Gauthier
HEC Montréal
Canada
genevieve.gauthier@hec.ca

Stefan Geiss
University of Jyväskylä, Finland
Finland
stefan.geiss@jyu.fi

Iliyan Georgiev
Solid Angle
United Kingdom
iliyan@solidangle.com

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

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

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

Mark Girolami
Imperial College
United Kingdom
m.girolami@imperial.ac.uk

Emmanuel Gobet
École Polytechnique, Paris
France
emmanuel.gobet@polytechnique.edu

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

Domingo Gomez
University of Cantabria
Spain
domingo.gomez@unican.es

Javier González-Villa
University of Cantabria
Spain
javiergvgv@gmail.com

Jackson Gorham
Stanford University
USA
jacksongorham@gmail.com

Abdul-Lateef Haji-Ali
University of Oxford, The Mathematical Institute
United Kingdom
abdullateef.hajiali@maths.ox.ac.uk

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

Shin Harase
Ritsumeikan University
Japan
harase@fc.ritsumei.ac.jp

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

Itsaso Hermosilla
Royal Military Academy Brussels
Belgium
itsaso.hermosilla@gmail.com

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

Lukas Herrmann
ETH Zürich
Switzerland
lukas.herrmann@sam.math.ethz.ch

Andre Herzwurm
TU Kaiserslautern
Germany
herzwurm@mathematik.uni-kl.de

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

Aicke Hinrichs
Johannes Kepler University Linz
Austria
aicke.hinrichs@jku.at

Mark Huber
Claremont McKenna College
USA
autotomic@gmail.com

Sofiya Ivanovska
IICT-BAS
Bulgaria
sofia@parallel.bas.bg

Ken Jackson
University of Toronto
Canada
krj@cs.toronto.edu

Sanjay Dominik Jena
ESG UQAM
Canada
sanjay.jena@cirrelt.ca

Matthias Kaeding

Germany
matthiaskaeding@posteo.de

Yoshihito Kazashi
School of Mathematics and Statistics, UNSW Sydney
Australia
z5038125@unsw.edu.au

Alexander Keller
NVIDIA
Germany
keller.alexander@gmail.com

Julien Keutchayan
Polytechnique Montréal
Canada
j.keutchayan@gmail.com

David Krieg
Friedrich Schiller University Jena
Germany
david.krieg@uni-jena.de

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

Ralph Kritzinger
Johannes Kepler University Linz
Austria
ralph.kritzinger@jku.at

Jaroslav Krivanek
Charles University, Prague
Czech Republic
jaroslav.krivanek@mff.cuni.cz

Raphael Kruse
Technische Universitat Berlin
Germany
kruse@math.tu-berlin.de

Sergei Kucherenko
Imperial College London
United Kingdom
s.kucherenko@imperial.ac.uk

Thomas Kühn
Mathematisches Institut, Universität Leipzig
Germany
kuehn@math.uni-leipzig.de

Robert Kunsch
Friedrich Schiller University Jena, Institute for Mathematics
Germany
Robert.Kunsch@uni-jena.de

Frances Kuo
School of Mathematics and Statistics, University of New South Wales
Australia
f.kuo@unsw.edu.au

Simon Lacoste-Julien
Université de Montréal
Canada
slacoste@iro.umontreal.ca

Patrick Laub
University of Queensland
Australia
patrick.laub@gmail.com

Simon Laurin-Lemay
Université de Montréal
Canada
evol.simon@gmail.com

Kody Law
Oak Ridge National Laboratory
USA
kodylaw@gmail.com

Christian Lecot
Universite Savoie Mont Blanc
France
Christian.Lecot@univ-smb.fr

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

Antoine Lejay
Inria
France
Antoine.Lejay@univ-lorraine.fr

Christiane Lemieux
University of Waterloo
Canada
clemieux@uwaterloo.ca

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

Josef Leydold
WU Wien
Austria
josef.leydold@wu.ac.at

Gang Li
Institute of applied physics and computational mathematics
China
lgseven@126.com

Xiaoou Li
University of Minnesota
USA
lixx1766@umn.edu

Fredrik Lindsten
Uppsala University
Sweden
fredrik.lindsten@it.uu.se

Supranee Lisawadi
Thammasat University
Thailand
lisawadi@gmail.com

Mike Ludkovski
UC Santa Barbara
USA
ludkovski@pstat.ucsb.edu

Florian Maire
University College Dublin
Ireland
florian.maire@ucd.ie

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

Oren Mangoubi
University of Ottawa
Canada
omangoubi@gmail.com

Ricardo Marques
Universitat Pompeu Fabra
Spain
ricardo.marques@upf.edu

Till Massing
Department of Economics, University of Duisburg-Essen
Germany
till.massing@uni-due.de

Thomas Muller-Gronbach
University of Passau
Germany
thomas.mueller-gronbach@uni-passau.de

Andreas Neuenkirch
University of Mannheim
Germany
andreas.neuenkirch@googlemail.com

Mario Neumüller
Johannes Kepler Universität
Austria
mario.neumueller@jku.at

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

Dirk Nuyens
KU Leuven
Belgium
dirk.nuyens@gmail.com

Art Owen
Stanford University
USA
owen@stanford.edu

Gilles Pagès
Université Paris 6
France
gilles.pages@upmc.fr

Raghu Pasupathy
Purdue University
USA
pasupath@purdue.edu

Florian Pausinger
TU München
Germany
florian.pausinger@gmx.at

Helene Perrier
LIRIS
France
helene.perrier@liris.cnrs.fr

Friedrich Pillichshammer
University of Linz
Austria
friedrich.pillichshammer@jku.at

Isaque Pimentel
École Polytechnique
France
pimentel.isaque@gmail.com

Warren Powell
Princeton University
USA
powell@princeton.edu

Stephan Prell
University of Duisburg-ESSEN
Germany
stephan.prell@uni-due.de

Pawel Przybylowicz
AGH University of Science and Technology, Faculty of Applied Mathematics
Poland
pprzybyl@agh.edu.pl

Aaditya Ramdas
University of California, Berkeley
USA
aaditya.ramdas@gmail.com

Rajesh Ranganath
Princeton University
USA
rajeshr@cs.princeton.edu

Lizanne Raubenheimer
Rhodes University
South Africa
L.Raubenheimer@ru.ac.za

Matthew Richey
St. Olaf College
USA
richeym@stolaf.edu

Pieterjan Robbe
KU Leuven
Belgium
pieterjan.robbe@kuleuven.be

Nicolas Rodrigue
Carleton University
Canada
nicolas.rodrigue@carleton.ca

Jeffrey Rosenthal
University of Toronto
Canada
jeff@math.toronto.edu

Sotirios Sabanis
University of Edinburgh
United Kingdom
s.sabanis@ed.ac.uk

Leonard Santana
North West University
South Africa
Leonard.Santana@nwu.ac.za

Wael Sboui
Université de Montréal
Canada
waelsboui@gmail.com

Wolfgang Ch. Schmid
Department of Mathematics, University of Salzburg
Austria
wolfgang.schmid@sbg.ac.at

Ingmar Schuster
FU Berlin
Germany
ingmar.schuster@fu-berlin.de

Taras Shalaiko
Uni Mannheim
Germany
tarasenya@gmail.com

DanHua ShangGuan
Institute of Applied Physics and Computational Mathematics
China
sgdh@iapcm.ac.cn

Alexander Shkolnik
UC Berkeley
USA
ads2@berkeley.edu

Ian Sloan
University of New South Wales
Australia
i.sloan@unsw.edu.au

Patchanok Srisuradetchai
Thammasat University
Thailand
s.patchanok@gmail.com

Andreas Stein
SimTech, University of Stuttgart
Germany
andreas.stein@mathematik.uni-stuttgart.de

Magdalena Strauss
MRC Biostatistics Unit, University of Cambridge
United Kingdom
magdalena.strauss@mrc-bsu.cam.ac.uk

Holger Stroot

Germany
stroot@mathematik.uni-kl.de

Yuya Suzuki
KU Leuven
Belgium
yuya.suzuki@kuleuven.be

Kosuke Suzuki
Hiroshima University
Japan
kosuke-suzuki@hiroshima-u.ac.jp

Michaela Szölgyenyi
Vienna University of Economics and Business
Austria
michaela.szoelgyenyi@wu.ac.at

Thuy Anh Ta
Université de Montréal
Canada
tathuyan@iro.umontreal.ca

Yoshihiro Taniguchi
University of Waterloo
Canada
ytaniguc@uwaterloo.ca

Andreas Thalhammer
Johannes Kepler University Linz
Austria
andreas.thalhammer@jku.at

Eric Torkia
Technology Partnerz Ltd
Canada
etorkia@technologypartnerz.com

Natalya Tracheva

Russia
natatracheva@gmail.com

Benjamas Tulyanitikul
Thammasat University
Thailand
benjamast@gmail.com

Ryan Turner
IVADO
Canada
turnerry@iro.umontreal.ca

Johan Van Kerckhoven
Koninklijke Militaire School
Belgium
johan.vankerckhoven@mil.be

Liangliang Wang
Simon Fraser University
Canada
lwa68@sfu.ca

Yiwei Wang
The Chinese University of Hong Kong
Hong Kong
ywwang@se.cuhk.edu.hk

Zhenru Wang
Mathematical Institute, University of Oxford
United Kingdom
zhenru.wang@maths.ox.ac.uk

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

Paula A Whitlock
Brooklyn College
USA
whitlock@brooklyn.cuny.edu

Cheryl Woodall
US Geological Survey
USA
cwoodall@usgs.gov

Henryk Wozniakowski
Columbia University
USA
henryk@cs.columbia.edu

Ma Xiang
Chinese University of Hong Kong
Hong Kong
xma@se.cuhk.edu.hk

Nathan Yang
McGill Desautels Faculty of Management
Canada
nathan.cc.yang@gmail.com

Hui Yao
The University of Queensland
Australia
h.yao@uq.edu.au

Takehito Yoshiki
Osaka City University
Japan
tttyoshiki@gmail.com

Fan Zhang
Columbia University
USA
fz2222@columbia.edu

Kan Zhang
Illinois Institute of Technology
USA
zhangkan.zhangkan@gmail.com

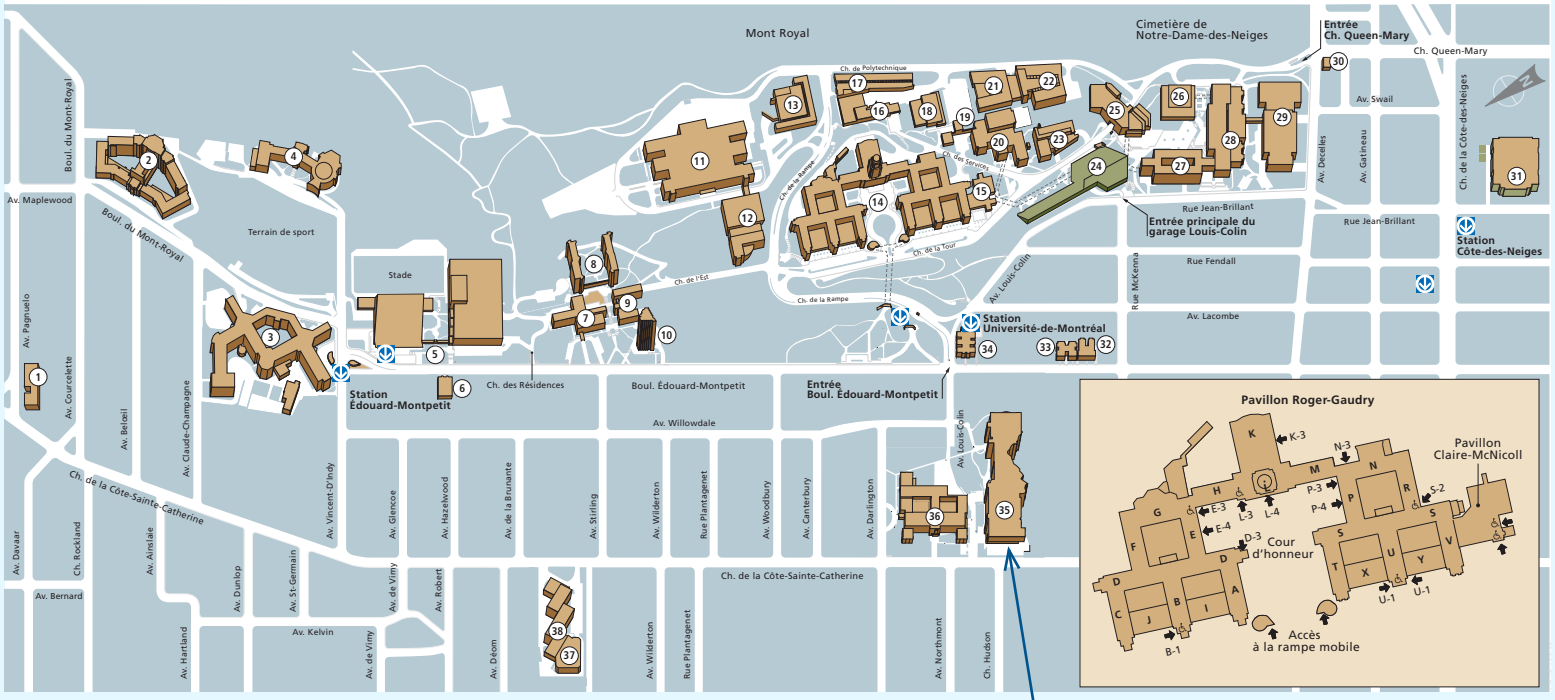
List of session chairs

Index

Campus Map

1 520, chemin de la Côte-Sainte-Catherine
2 1420, boulevard Mont-Royal
3 Pavillon Marie-Victorin
4 Pavillon de la Faculté de musique
5 Centre d'éducation physique et des sports (CEPSUM)
6 2101, boulevard Édouard-Montpetit
7 Pavillon J.-A.-DeSève (Centre étudiant)
8 Résidence C
9 Résidence A et annexe
10 Pavillon Thérèse-Casgrain
11 École Polytechnique
12 Pavillons Pierre-Lassonde et Claudette McKay-Lassonde
13 Pavillon J.-Armand-Bombardier
14 Pavillon Roger-Gaudry
15 Pavillon Claire-McNicoll
16 Pavillon de la Direction des immeubles
17 Centre des technologies de fabrication en aérospatiale
18 Centrale thermique
19 Pavillon René-J.-A.-Lévesque
20 Pavillon André-Aisenstadt
21 Pavillon Jean-Coutu
22 Pavillon Marcelle-Coutu
23 Pavillon Paul-G.-Desmarais
24 Garage Louis-Colin
25 Pavillon Samuel-Bronfman
26 Pavillon Maximilien-Caron
27 Pavillon Lionel-Groulx
28 3200, rue Jean-Brillant
29 HEC Montréal – 5255, av. Decelles
30 3333, chemin Queen-Mary
31 3744, rue Jean-Brillant
32 3050-3060, boulevard Édouard-Montpetit
33 3032-3034, boulevard Édouard-Montpetit
34 2910, boulevard Édouard-Montpetit
35 HEC Montréal – Pavillon principal
36 Pavillon de la Faculté de l'aménagement
37 Pavillon Liliane de Stewart
38 Pavillon Marguerite-d'Youville

Stationnement pour visiteurs



HEC MONTRÉAL

CENTRE DE RECHERCHES MATHÉMATIQUES

GERAD
GROUP FOR RESEARCH IN DECISION ANALYSIS

DIRO

IVADO
INSTITUT DE VALORISATION DES DONNÉES

Université de Montréal

Pierre L'Ecuyer
Département d'informatique et de recherche opérationnelle (DIRO)
Université de Montréal
C.P. 6128, Succ. Centre-Ville
Montréal (Québec)
Canada, H3C 3J7

www.iro.umontreal.ca/~lecuyer