Applied Probability Frontiers: Computational and Modeling Challenges
A Workshop in Honor of Peter W. Glynn's 60th Birthday
May 31 – June 5, 2015

MEALS

*Breakfast (Buffet): 7:00–9:30 am, Sally Borden Building, Monday–Friday
*Lunch (Buffet): 11:30 am–1:30 pm, Sally Borden Building, Monday–Friday
*Dinner (Buffet): 5:30–7:30 pm, Sally Borden Building, Sunday–Thursday
Coffee Breaks: As per daily schedule, in the foyer of the TransCanada Pipeline Pavilion (TCPL)
***Please remember to scan your meal card at the host/hostess station in the dining room for each meal.**

MEETING ROOMS

All lectures will be held in the lecture theater in the TransCanada Pipelines Pavilion (TCPL). An LCD projector, a laptop, a document camera, and blackboards are available for presentations.

SCHEDULE

Sunday

16:00	Check-in begins (Front Desk - Professional Development Centre - open 24 hours)
17:30–19:30	Buffet Dinner, Sally Borden Building

Monday

7:00–8:45	Breakfast
9:00–9:15	Introduction and Welcome by BIRS Station Manager, TCPL
9:15–10:15	Karl Sigman
10:15–11:00	Coffee Break
11:00–12:00	Open Problems
12:00–13:30	Lunch
13:30–14:30	Henry Lam
14:30–14:45	Group Photo; meet in foyer of TCPL (photograph will be taken outdoors so a jacket might be required).
14:45–15:30	Coffee Break
15:30–16:30	Sigrun Andradottir
16:30–17:30	Sandeep Juneja
17:30–19:30	Dinner

Tuesday

7:00–9:00	Breakfast
9:00–10:00	Soren Asmussen
10:00–11:00	Coffee Break
11:00–12:00	Amy Ward
12:00–13:30	Lunch
13:30–14:30	Eric Moulines
14:30–15:30	Coffee Break
15:30–16:30	Gareth Roberts
16:30–18:00	Poster Session
18:00–19:30	Dinner

Wednesday

7:00–8:30	Breakfast
8:30–9:30	Kavita Ramanan
9:30–10:30	Ton Dieker
10:30–11:00	Coffee Break
11:00–12:00	Hermann Thorisson
12:00–13:30	Lunch
	Free Afternoon
17:30–19:30	Dinner

Thursday

7:00–9:00	Breakfast
9:00–10:00	Mike Giles
10:00–11:00	Coffee Break
11:00–12:00	Mariana Olvera-Cravioto
12:00–13:30	Lunch
13:30–14:30	Onno Boxma
14:30–15:30	Coffee Break
15:30–16:30	Jim Dai
18:00–19:30	Dinner

Friday

7:00–9:00	Breakfast
9:00–10:00	Open Problems Reprise
10:00	Coffee Break
11:30–13:30	Lunch

Checkout by 12 noon.

** 5-day workshop participants are welcome to use BIRS facilities (BIRS Coffee Lounge, TCPL and Reading Room) until 3 pm on Friday, although participants are still required to checkout of the guest rooms by 12 noon. **

Applied Probability Frontiers: Computational and Modeling Challenges
A Workshop in Honor of Peter W. Glynn's 60th Birthday
May 31 – June 5, 2015

ABSTRACTS
(in alphabetical order by speaker surname)

Speaker: **Sigrun Andradottir** (Georgia Tech)
Title: *Simulation-based continuous optimization with stochastic constraints*
Joint work with Liujia Hu.

Abstract: We present an Adaptive Search with Discarding and Penalization (ASDP) method for solving continuous simulation optimization problems involving stochastic constraints. Rather than addressing feasibility separately, ASDP utilizes the penalty function method from deterministic optimization to convert the original problem into a series of simulation optimization problems without stochastic constraints. We present conditions under which the ASDP approach converges almost surely from inside the feasibility region, and under which it converges to the optimal solution but without feasibility guarantee. We also conduct numerical studies aimed at assessing the efficiency of ASDP under the two different convergence modes.

Speaker: **Soren Asmussen** (Aarhus University)
Title: *Markov processes and simulation. Snapshots from three decades*
Abstract: I will give an overview of some of Peter Glynn's work, with particular emphasis on the interactions I have had with him myself over the years.

Speaker: **Onno Boxma** (Eindhoven University of Technology)
Title: *Applied probability meets Bessel, Hermite, Kummer, Tricomi,*
Wiener & Hopf (and also Ornstein & Uhlenbeck)
Note: Based on collaborations with Hansjoerg Albrecher, Shaul Bar-Lev, Rim Essifi, Guido Janssen, Richard Kuijstermans, Britt Mathijsen and David Perry.

Abstract: We discuss a few applied probability models: a blood bank model, an insurance risk model and a queueing/inventory model. In each case, the process under study $\{X(t),\ t \geq 0\}$ has state space $(-\infty, \infty)$.
In the blood bank case, $X(t)$ indicates the amount of blood present at time t, if there is a positive amount present; otherwise, $-X(t)$ denotes the total amount that is being demanded.
In the insurance risk case, $X(t)$ indicates the capital at time t, if there is a positive amount present; otherwise, $-X(t)$ denotes the shortage. It is agreed that ruin does not immediately occur when $X(t)$ becomes negative, but the process ends (bankruptcy) according to a bankruptcy rate function $\omega(-X(t))$ when $X(t) < 0$.
In the queueing/inventory model, $X(t)$ indicates the amount of work present at time t, if there is a positive amount present; if no work is present, the tireless server keeps working, building up an inventory, and $-X(t)$ then denotes the inventory level. The inventory is removed at a rate $\omega(-X(t))$.
Under Poisson assumptions for the various arrival processes (of blood donations and blood requirements; of claims; of service requirements) and ergodicity conditions we try to determine the steady-state distributions of the $X(t)$ processes; in the insurance risk model, we determine the bankruptcy probability. This appears to involve various special functions, like those of Bessel, Hermite, Kummer and Tricomi.

Speaker: **Jim Dai** (Cornell University)
Title: *Steins method for steady-state diffusion approximations*
Joint work with Anton Braverman

Abstract: We consider $M/Ph/n+M$ queueing systems. We prove the rate of convergence for approximating the stationary distribution of the normalized system size process by that of a piecewise Ornstein-Uhlenbeck (OU) process. We prove that for a large class of functions, the difference of the expectation under the stationary measure of the piecewise OU process and the expectation under the stationary measure of the system size is at most $C/\lambda^{1/4}$, where the constant C is independent of the arrival rate λ and the number of servers n as long as they are in the Halfin-Whitt parameter regime. For the proof, we develop a modular framework that is based on Steins method. The framework has three components: gradient estimate for Poisson equation solution, generator coupling, and state space collapse.

Speaker: **Ton Dieker** (Columbia University)
Title: *Exact simulation of stationary max-stable random fields*
Joint work with Thomas Mikosch.

Abstract: Max-stable random fields can be viewed as the analogs of Gaussian processes in the world of extreme value analysis. We propose the first exact simulation method for a class of such random fields. The main idea is to apply suitable changes of measure, which yield new representations that can be exploited for simulation.

Speaker: **Mike Giles** (University of Oxford)
Title: *Multilevel Monte Carlo methods*

Abstract: With Monte Carlo methods, to achieve improved accuracy one often requires more expensive sampling (such as a finer timestep discretisation of a stochastic differential equation) in addition to more samples. Multilevel Monte Carlo methods aim to avoid this by combining simulations with different levels of accuracy. In the best cases, the average cost of each sample is independent of the overall target accuracy, leading to very large computational savings.

The talk will emphasise the simplicity of the approach, give an overview of the range of applications being worked on by various researchers, and mention some recent extensions including work by Peter Glynn and Chang-han Rhee. Applications to be discussed will include financial modelling, engineering uncertainty quantification, stochastic chemical reactions, and the Feynman-Kac formula for high-dimensional parabolic PDEs.

Further information can be obtained from http://people.maths.ox.ac.uk/gilesm/acta/

Speaker: **Sandeep Juneja** (Tata Institute of Fundamental Research)
Title: *Ordinal optimization - Empirical large deviations rate estimators, and multi-armed bandit methods*
Joint work with Peter W. Glynn

Abstract: Consider the ordinal optimization problem of finding a population amongst many with the largest mean when these means are unknown but population samples can be generated via simulation. Typically, by selecting a population with the largest sample mean, it can be shown that the false selection probability decays at an exponential rate. Lately researchers have sought algorithms that guarantee that this probability is restricted to a small δ in order $\log(1/\delta)$ computational time by estimating the associated large deviations rate function via simulation. We show that such guarantees are misleading. Enroute, we identify the large deviations principle followed by the empirically estimated rate function that may be of independent interest. Further, we show a negative result that when populations have unbounded support, any policy that asymptotically identifies the correct population with probability at least $1-\delta$ for each problem instance requires more than $O(\log(1/\delta))$ samples in making such a determination in any problem instance. This suggests that some restrictions are essential on populations to devise $O(\log(1/\delta))$ algorithms with $1-\delta$ correctness guarantees. We note that under restriction on population moments, methods from multi-armed bandit literature can be adapted to devise such algorithms.

Speaker: **Henry Lam** (University of Michigan)
Title: *Model Uncertainty and Robust Stochastic Modeling*
Abstract: Virtually any performance analysis in stochastic modeling relies on input model assumptions that, to some extent, deviate from the truth. This talk will investigate a worst-case framework to quantify these model errors and correspondingly robustify the stochastic outputs. It entails posting optimization programs over the input probability distributions, with constraints representing the modeler's partial, nonparametric knowledge about them. We illustrate these optimization formulations in several contexts in stochastic modeling, describe their computational challenges, and present some machinery in approximating their solutions.

Speaker: **Eric Moulines** (Télécom ParisTech)
Title: *Efficient sampling log-concave distribution over high-dimensional space*
Abstract: Sampling over high-dimensional space has become a prerequisite in the applications of Bayesian statistics to machine learning problem. In many situations of interest, the log-posterior distribution is concave. The likelihood part is generally smooth and gradient Lipshitz while the prior is concave but typically not smooth (the archetypical problem is the LASSO or the elastic-net penalty, but many other problems can be cast into this framework). We will describe methods to sample such distributions, which are adapted from the state-of-the-art optimization procedures which have been developed in this context. We will also provide convergence in Wasserstein distance to the equilibrium, showing explicitly the dependence in the dimension of the parameter space and the sparsity (effective dimension of the model). Several examples will be presented to illustrate our results.

Speaker: **Mariana Olvera-Cravioto** (Columbia University)
Title: *Efficient simulation for weighted branching trees*
Abstract: Motivated by recent results for the analysis of information ranking algorithms on complex networks and parallel queueing networks with synchronization, we propose an efficient algorithm for simulating the endogenous solution to max-plus branching stochastic fixed-point equations. The aforementioned solutions can be constructed on a weighted branching process, but closed-form expressions for their distributions are in general unavailable. Unlike for the Galton-Watson process, Laplace transform methods in this context are also problematic. Hence, a stochastic simulation approach seems to be the most natural way of numerically approximating the distributions and moments of these solutions. Naive Monte Carlo techniques, however, are extremely inefficient due to the geometric growth of the underlying trees. We describe in this talk an algorithm based on iterative bootstrap whose complexity grows linearly (as opposed to exponentially) in the number of generations of the weighted branching process being simulated. We also show the consistency of a wide class of estimators based on our algorithm.

Speaker: **Kavita Ramanan** (Brown University)
Title: *Sensitivity analysis of reflected Brownian motions*
Joint work with David Lipshutz.

Abstract: Semimartingale reflected Brownian motions (SRBMs) arise as diffusion approximations to queueing networks in heavy traffic. Since the parameters of the networks, and therefore the approximating diffusions, can be uncertain, it is of interest to compute the sensitivity of expectations of functionals of the SRBMs with respect to system parameters. With this as motivation, we consider the pathwise differentiability of a large class of SRBMs in convex polyhedra. In particular, we characterize derivatives of stochastic flows of SRBMs (which captures sensitivity with respect to initial conditions) and also obtain results on the sensitivity of SRBMs to small perturbations of the drift. We demonstrate how our results may be used to develop efficient computational algorithms for estimation of these sensitivities.

Speaker: **Gareth Roberts** (University of Warwick)
Title: *Towards not being afraid of the big bad data set*
Joint work with Paul Fearnhead, Murray Pollock and Adam Johansen.

Abstract: This talk will present the foundations behind a new algorithm for systematic error-free Monte Carlo simulation from intractable target distributions. The main motivation behind the work is to construct a method for exploring posterior distributions for Bayesian analyses of extremely large datasets where computation of the likelihood function at each iteration of an algorithm is prohibitively expensive. The algorithm is a continuous time sequential Monte Carlo procedure which extends many of the ideas used in exact simulation from diffusion sample paths.

Speaker: **Karl Sigman** (Columbia University)
Title: *Exact simulation for some multi-dimensional queueing models with renewal input*
Joint work with Jose Blanchet and Yanan Pei

Abstract: By using some new results on the exact simulation of FIFO GI/GI/1 queues introduced by J. Blanchet and A. Wallwater, we show how to simulate exactly the stationary distribution of the FIFO GI/GI/c queue, $c \geq 2$. This method uses dominated coupling from the past (DCFP) as well as the Random Assignment (RA) discipline, and complements the earlier work of K. Sigman in which Poisson arrivals were assumed. In addition, we show how to also exactly simulate from some other multidimensional models, including some simple multi-class networks.

Speaker: **Hermann Thorisson** (University of Iceland)
Title: *Mass-Stationarity, Shift-Coupling, and Brownian Motion*

Abstract: After considering mass-stationarity and shift-coupling briefly in an abstract setting, we focus on the special case of stochastic processes on the line associated with diffuse random measures. The main examples are Brownian motion and the Brownian bridge.

Speaker: **Amy Ward** (University of Southern California)
Title: *On the Control of Fork-Join Networks*
Joint work with Erhun Ozkan

Abstract: Networks in which the processing of jobs occurs both sequentially and in parallel are prevalent in many application domains, such as computer systems, healthcare, manufacturing, and project management. The parallel processing of jobs gives rise to synchronization constraints that can be a main reason for job delay, which results in holding costs. In comparison with feedforward queueing networks that have only sequential processing of jobs, the approximation and control of networks that have synchronization constraints is less understood. One well-known modeling framework in which synchronization constraints are prominent is the fork-join processing network.

Our objective is to find scheduling rules for fork-join processing networks that minimize holding costs. To do this, we focus on a prototypical network with two job classes (a and b), two fork operations, one shared server, and two join operations. The fork operations are first, followed by the simultaneous processing of type a (b) jobs by a dedicated server and a shared server, and, finally the join operations. We solve the scheduling problem for the shared server (that is, which type of job to prioritize each time the server becomes available) when that server is in heavy traffic. We show that a $c\mu$-type static priority policy is asymptotically optimal when the shared server is in some sense slow at processing the more expensive type a jobs. Otherwise, an asymptotically optimal control is a state-dependent slow departure pacing control in which the shared server slows its processing of type a jobs to match the departure process of those jobs from the dedicated type-a server. Finally, by considering a broader class of fork-join networks, we see that the departure pacing idea is to some extent robust.