ETH zürich

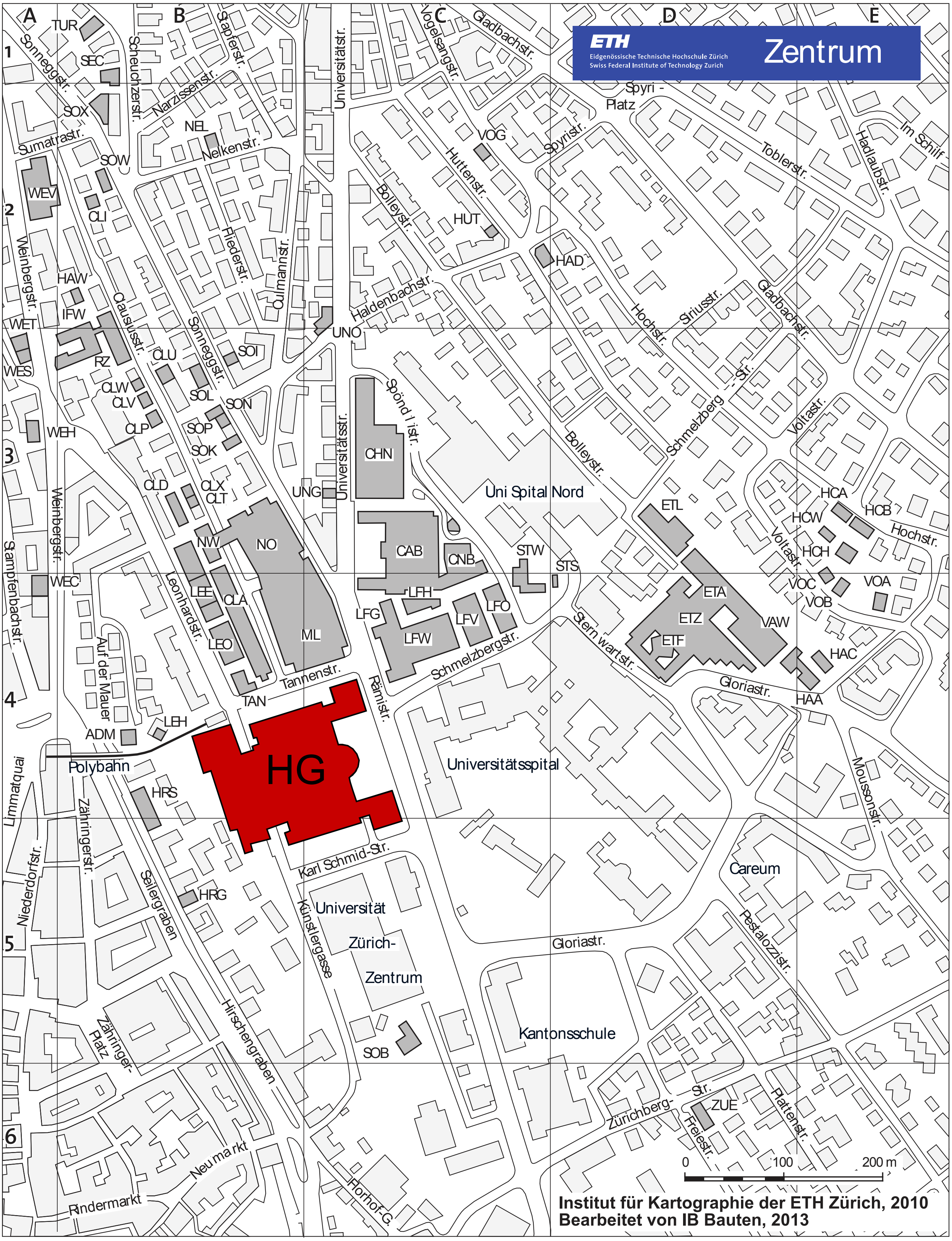
MascotNum 2014

Computer Experiments and Metamodels for Uncertainty Quantification

April 23 - 25, 2014
ETH Zürich

Conference Program

D BAUG


ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich
Zentrum
A
B
C
D
E
1
2
3
4
5
6
HG
Universitätsspital
Uni Spital Nord
Universität Zürich-Zentrum
Kantonsschule
Careum
Polybahn
CHN
CAB
CNB
NO
ML
LFW
LFV
LFO
LFH
LFG
ETA
ETZ
ETF
ETL
VAW
HAA
HAC
HCA
HCB
HCW
HCH
VOC
VOA
VOB
STW
STS
UNO
UNG
TAN
LEH
ADM
HRS
HRG
SOB
ZUE
HAD
HUT
VOG
TUR
SEC
SOX
NEL
SOW
WEV
CLI
HAW
IFW
WET
RZ
WES
CLU
SOI
CLW
CLV
SOL
SON
CLP
SOP
SOK
WEH
CLD
CLX
CLT
NW
LEE
CLA
LEO
WEC
Sonneggstr.
Sumatrastr.
Scheuchzerstr.
Narzissenstr.
Nelkenstr.
Stapferstr.
Universitätstr.
Vogelsangstr.
Gladbachstr.
Spyri-Platz
Spyristr.
Toblerstr.
Im Schilf
Hadlaubstr.
Huttenstr.
Bolleystr.
Fliederstr.
Culmannstr.
Weinbergstr.
Clausiusstr.
Haldenbachstr.
Hochstr.
Siriusstr.
Schmelzbergstr.
Voltastr.
Spöndlistr.
Stampfenbachstr.
Leonhardstr.
Auf der Mauer
Tannenstr.
Rämistr.
Sternwartstr.
Gloriastr.
Moussonstr.
Limmatquai
Zähringerstr.
Seilergraben
Niederdorfstr.
Karl Schmid-Str.
Künstlergasse
Pestalozzistr.
Hirschengraben
Zähringer-Platz
Zürichberg-Str.
Freiestr.
Plattenstr.
Neumarkt
Rindermarkt
Hofstr.
0
100
200 m
Institut für Kartographie der ETH Zürich, 2010
Bearbeitet von IB Bauten, 2013

Contents

Organisers and Sponsors

Organised by

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

UNIVERSITÄT BERN

Sponsored by

Groupement de Recherche MASCOT-NUM
Méthodes d'Analyse Stochastique pour les COdes et Traitements NUMériques

MascotNum Research Network

ReDICE

ReDice Consortium

Risk, Safety & Uncertainty Quantification

Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich

D BAUG

Department of Civil, Environmental and Geomatic Engineering, ETH Zürich

Kanton Zürich

Kanton Zürich, Staatskanzlei

Stadt Zürich

Stadt Zürich, Präsidialdepartement

Foreword

Dear Colleagues

It is our pleasure to welcome you to the 2014 edition of the MascotNum annual workshop. After eight meetings held in France (Toulouse (2006), Lyon (2007), Cadarache (2008), Paris (2009), Avignon (2010), Villard-de-Lans (2011,) Bruyères-le-Châtel (2012), Nice (2013)), this year's workshop is jointly organized in Zürich by the Chair of Risk, Safety and Uncertainty Quantification of the Swiss Federal Institute of Technology (ETH Zürich) and by the Institute of Mathematical Statistics and Actuarial Science of the University of Bern.

Originating from a French initiative, MascotNum had already started to get international recognition with the participation of numerous invited speakers from overseas in previous editions, and through the joint organization of the annual meeting together with the 7th International Conference on Sensitivity Analysis of Model Output (SAMO'2013) in Nice last year. The 2014 edition is for the first time organized outside France and we are happy to have contributing PhD students and invited lecturers from several neighbouring countries. We hope that this event will foster a number of scientific exchanges between research communities and new opportunities for international collaborations.

The theme of MascotNum 2014 is "Computer Experiments and Meta-models for Uncertainty Quantification". Computer experiments can be considered as one of the core interests of MascotNum. With the increase in accessible computational power, many real-world problems ranging from engineering to natural sciences are tackled today using numerical simulations. Because some of the parameters necessary to describe such problems (state variables, boundary conditions, etc.) may be unknown or only partly known to the modeller, computer experiments are needed to get a global picture of how variations in those parameters may affect the modelling results. This problem setting is now common to a variety of applications, including the optimal design of cars, aircrafts, etc., the risk analysis of civil structures against natural hazard, the study of climate change, or the search for efficient molecules and dosing in therapeutic strategies.

Meta-models are a key tool for getting this global picture and guiding new simulations when the computational budget is limited, as it often is the case when dealing with high-fidelity simulations involving a large number of input parameters. Meta-models are a wide class of methods including, amongst others, polynomial chaos expansions, reduced-basis approaches, splines and Gaussian Process modelling techniques. In this workshop, we aim at giving an overview of how meta-models are successfully used in several research communities, with a special focus on Uncertainty Quantification.

Uncertainty quantification itself has emerged as a well-identified, multi-disciplinary research field in the last few years, which lies at the boundary of statistics, applied mathematics and engineering applications. Its increasing popularity in the scientific community has led to the organization of new series of conferences devoted to this topic, and to the launch of not less than four new journals since 2011 in this area. Among others, scientific groups have emerged in the United Kingdom (Uncertainty in Computer Models (UCM) conferences), in Germany (GAMM Uncertainty Quantification group) and in the USA (SIAM Conferences on Uncertainty Quantification). In parallel, joint academic/industrial projects such the Dice and ReDice consortiums in France have fostered the research in this field.

The organization of MascotNum2014 in Zürich is a step towards our goal of establishing bridges between the MascotNum network and these various initiatives, especially through invited lectures which highlight the active contribution of Swiss researchers and institutions in this area.

Such a workshop could not have been possible without the support of the Office for Events & Location Development of ETH Zürich. The financial support of the MascotNum CNRS research group, the ReDice Consortium, the Department of Civil, Environmental and Geomatic Engineering of ETH Zürich, the City and the Canton of Zürich are also gratefully acknowledged.

On behalf of the Organizing Committee and the MascotNum board we welcome you warmly in Zürich. We wish you a very fruitful conference and we hope that you will also enjoy the charms of the Lake and of the Old City Centre.

Prof. Dr. Bruno Sudret
Chair of MascotNum 2014

Dr. David Ginsbourger
Co-Chair of MascotNum 2014

Committees

Organising Committee

Chair Bruno Sudret, *Chair of Risk, Safety and Uncertainty Quantification, ETH Zürich, Switzerland*

Co-Chair David Ginsbourger, *Institute of Mathematical Statistics and Actuarial Science, University of Bern, Switzerland*

MascotNum Board

Mark Asch, *Université de Picardie, France*
Sébastien Da Veiga , *Safran, France*
Bertrand Iooss, *EDF R&D, France*
Fabien Mangeant, *EADS IW, France*
Amandine Marrel , *CEA Cadarache, France*
Hervé Monod, *INRA Jouy-en-Josas, Département MIA, France*
Roland Masson, *Université de Nice, France*
Anthony Nouy, *Ecole Centrale de Nantes, France*
Clémentine Prieur, *Université Joseph Fourier, France*
Luc Pronzato, *I3S, CNRS Sophia Antipolis, France*
Olivier Roustant, *École des Mines de Saint Étienne, France*
Nathalie Saint Geours, *Irstea, France*
Nathalie Villa-Vialaneix, *Université de Perpignan, France*
François Wahl, *IFPEnergiesNouvelles, France*

MascotNum Scientific Committee

Jean-Marc Azaïs, *Université Paul Sabatier, France*
Gérard Biau, *Université Paris VI, France*
Pierre Del Moral, *INRIA Bordeaux-Sud Ouest, France*
Fabrice Gamboa, *Université Paul Sabatier, France*
Josselin Garnier, *Université Paris VII, France*
Olivier Le Maître, *CNRS, France*
Clémentine Prieur, *Université Joseph Fourier, France*
Luc Pronzato, *I3S, CNRS Sophia Antipolis, France*
Christophe Prud'hommme, *Université de Strasbourg, France*
Emmanuel Vazquez, *SUPELEC, Gif sur Yvette, France*

Local Organizing Committee & Secretariat

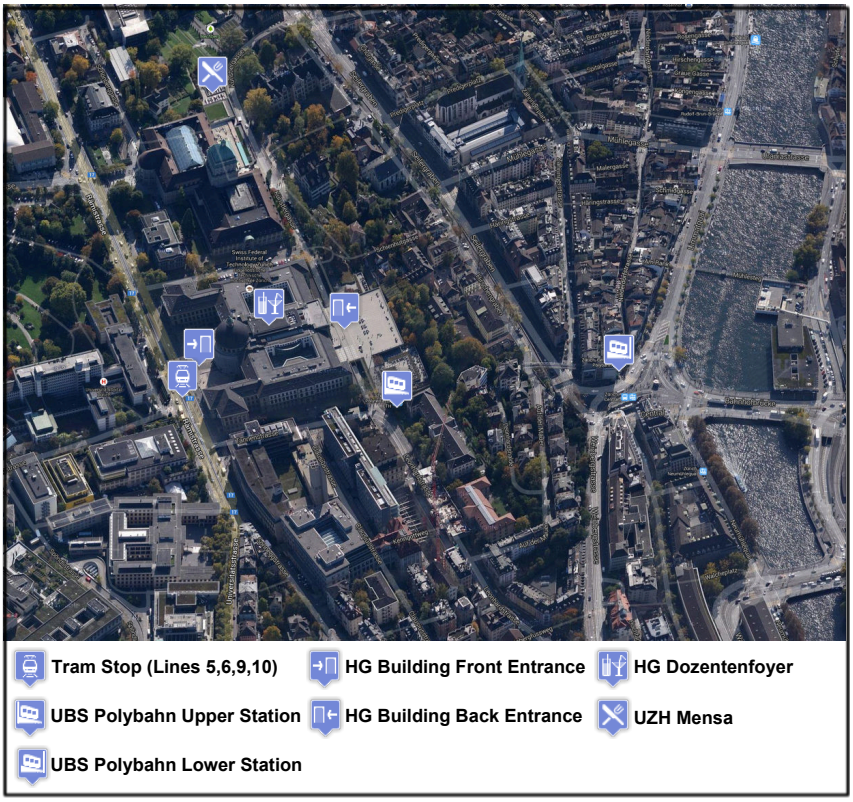
Christos Lataniotis
Stefano Marelli
Margaretha Neuhaus

General Information

Conference Venue

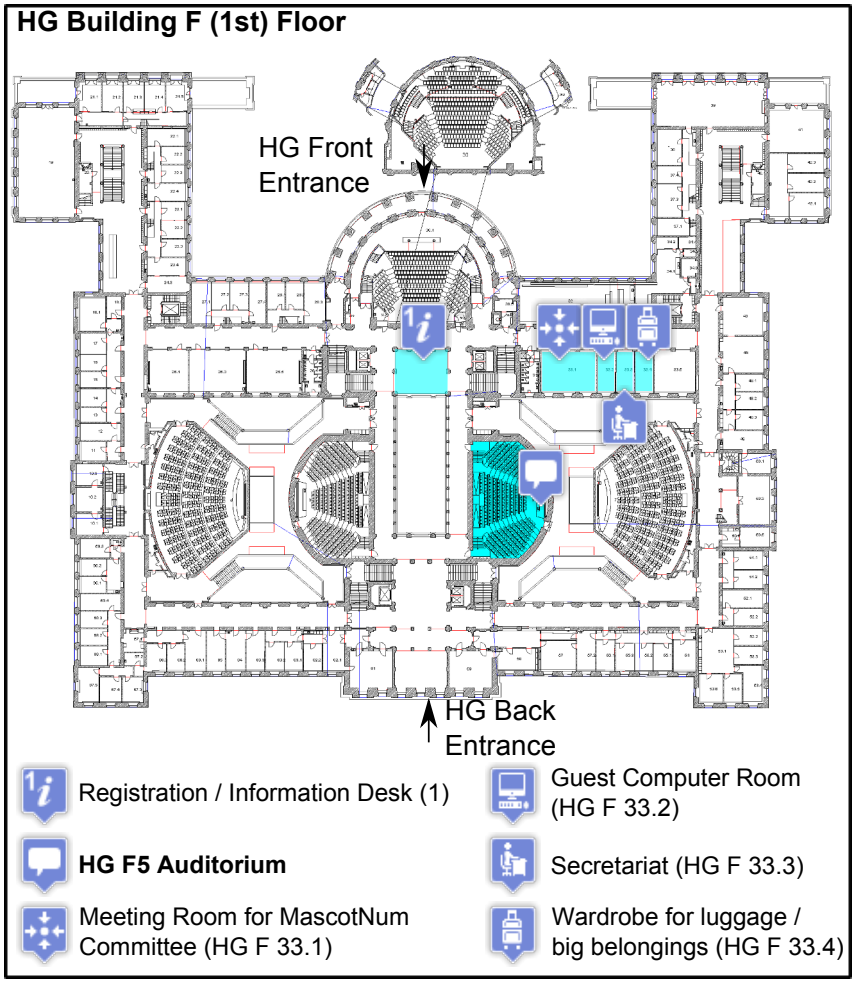
The conference will take place in the historical main building of ETH Zürich (Hauptgebaüde - HG Building).

Address: ETH Zürich Hauptgebäude (HG)
Rämistrasse 101
8092 Zürich
Tel: +41 44 632 11 11



Pre-registration

There will be a pre-registration on Tuesday April 22nd 2014 from 18:00 to 19:00 followed by a welcome apéro at *Dozentenfoyer*, located in the *HG Building, J-Floor*.

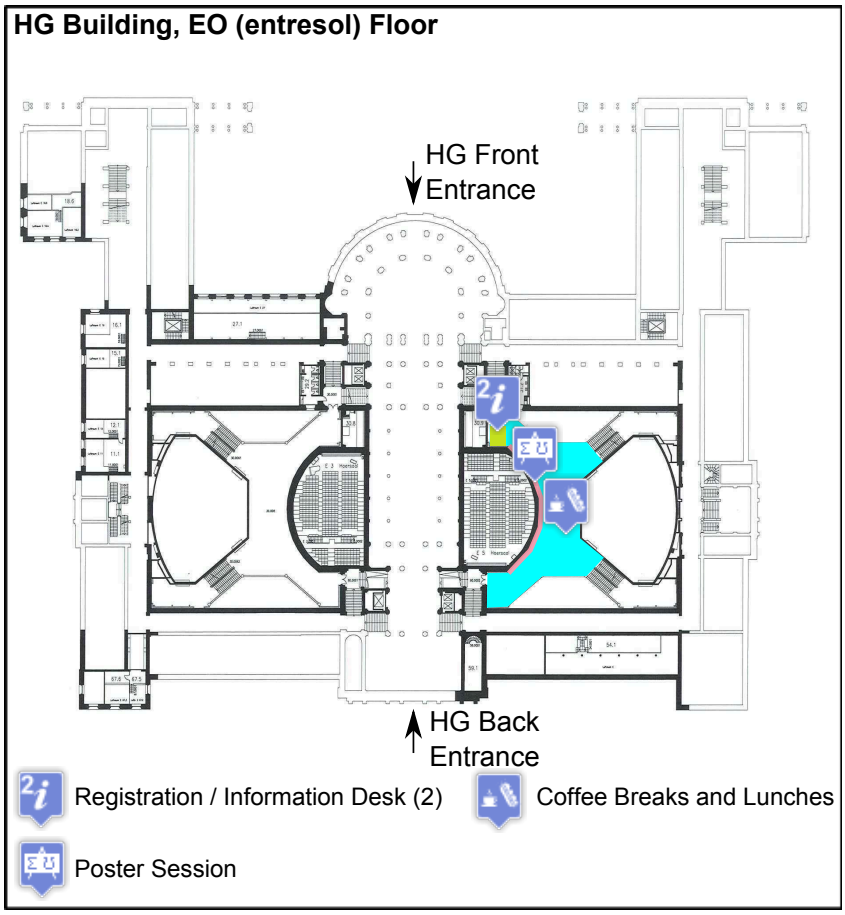


Talks/Lectures Room

The talks and lectures will take place in *Auditorium F5*. There is a wardrobe available inside the Auditorium with limited space. For luggage or other large-sized belongings please use the *wardrobe in HG F 33.4*.

Registration and Information Desk

The registration and information desk is located at the *Foyer Audi Max* (Floor F of HG building) on the first day (Wednesday, April 23rd) of the conference and at *HG EO Süd* (Floor EO of HG building) on the second and third day (Thursday-Friday, April 24-25th).



Poster Session

The Poster Session will take place in HG EO Süd. The posters are located in the area denoted by in the map above.

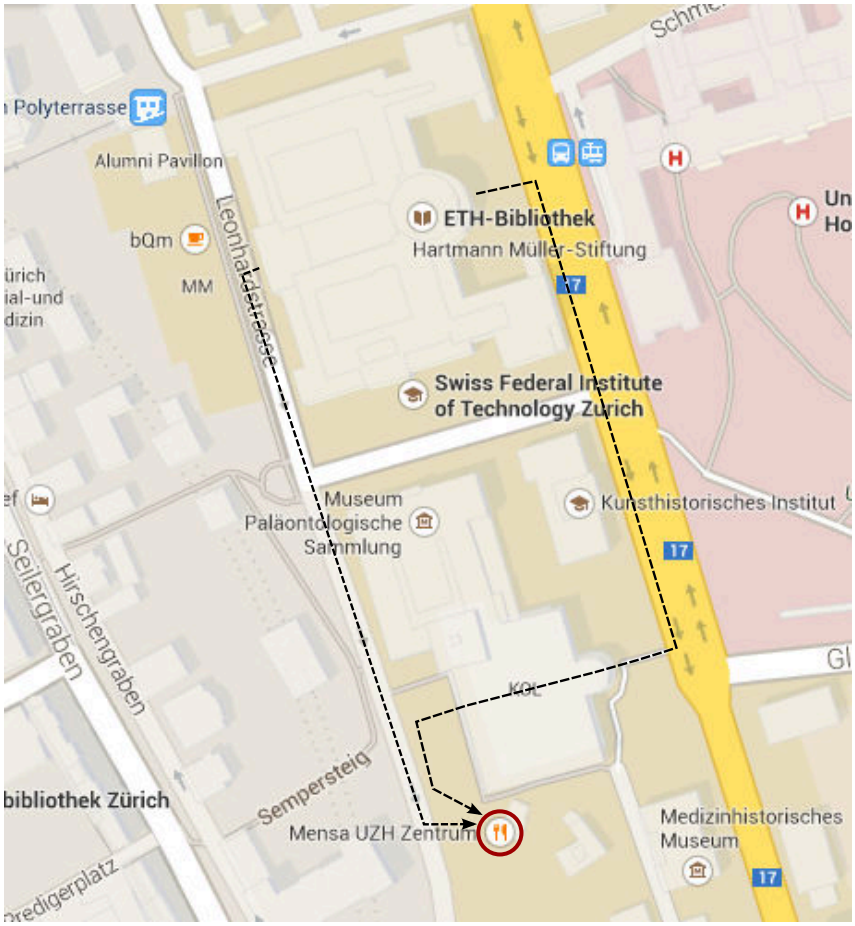
Coffee Breaks and Lunches

Coffee breaks and lunches will take place in HG EO Süd, in the area denoted by in the map above.

Conference Dinner

The conference dinner will take place in the Mensa of the University of Zürich, on Thursday, April 24th, 2014, at 19:30. The route to get there from the HG building is shown below.

Address: Mensa UZH Zentrum (1st Floor)
Künstlergasse 10
8001, Zürich
Tel: +41 44 634 23 70/71



Technical Program

Day 0 - Tuesday April 22nd, 2014

18:00 - 19:00	Registration	Dozentenfoyer
18:30 - 21:00	Ice Breaker	

Day 1 - Wednesday April 23rd, 2014

08:30 - 09:15	Registration	Foyer Audi Max
09:15 - 09:30	Opening Talk, B. Sudret / D. Ginsbourger	HG F5
09:30 - 10:30	Opening Lecture by C. Schillings (ETH Zürich)	**Chair: C. Schwab** HG F5
10:30 - 10:50	Coffee Break	HG EO Süd
10:50 - 11:20	S. Sukys (ETH Zürich). *Multi-level Monte Carlo finite volume methods for stochastic systems of hyperbolic conservation laws.*	**Chair: L. Pronzato** HG F5
11:20 - 11:50	D. Jacquemart (University Rennes I). *Adaptive particle methods for rare event simulation in a Markovian framework.*	
11:50 - 12:20	J. Nagel (ETH Zürich). *Bayesian multilevel model calibration for inversion of "perfect" data in the presence of uncertainty.*	
12:20 - 13:30	Lunch Break	HG EO Süd
13:30 - 15:00	Poster Session	HG EO Süd
15:00 - 15:30	A. Ahidar (Université Paul Sabatier). *Multivariate quantile surfaces and application to an aircraft problem.*	**Chair: F. Gamboa** HG F5
15:30 - 16:00	R. Fischer (Université Paris-Est). *Modeling dependence under constraint.*	
16:00 - 16:20	Coffee Break	HG EO Süd
16:20 - 16:50	F. Zertuche (Université Joseph Fourier). *Multi-fidelity regression using a non-parametric relationship.*	**Chair: F. Wahl** HG F5
16:50 - 17:20	S. Nanty (Université Joseph Fourier). *Uncertainty quantification and visualization for functional random variables.*	

Day 2 - Thursday April 24th, 2014

08:30 - 09:00	Registration	HG EO Süd
09:00 - 09:15	Conference opening - W. Kröger (ETH Risk Center)	HG F5
09:15 - 09:30	MascotNum news - C. Prieur (Université Joseph Fourier)	
09:30 - 10:30	F. Nobile (EPF Lausanne). *Discrete least-square polynomial approximations for high dimensional uncertainty propagation.*	**Chair: B. Sudret** HG F5
10:30 - 10:50	Coffee Break	HG EO Süd
10:50 - 11:50	J.P. Vert (Institut Curie). *Machine learning for personalised genomics.*	**Chair: O. Roustant** HG F5
11:50 - 12:50	A. Krause (ETH Zürich). *Focusing exploration with confidence.*	
12:50 - 14:15	Lunch Break	HG EO Süd
14:15 - 15:15	F.S. Koutsourelakis (TU Munich). *Simulation-based, high-dimensional stochastic optimization: application in robust topology optimization under large material uncertainties.*	**Chair: P. Koumoutsakos** HG F5
15:15 - 16:00	S. Marelli (ETH Zürich). *UQLab: A framework for Uncertainty Quantification in Matlab.*	
16:00 - 16:20	Coffee Break	HG EO Süd
16:20 - 17:20	D. Ginsbourger (University of Bern). *Incorporating structural priors in Gaussian random field models.*	**Chair: F. Nobile** HG F5
19:30 - 22:30	Conference Dinner	Mensa UZH Zentrum

Day 3 - Friday April 25th, 2014

		Chair: C. Prieur
09:00 - 10:00	P. Koumoutsakos (ETH Zürich). *Data driven, molecular dynamics for nanoscale fluid mechanics.*	HG F5
10:00 - 10:30	C. Chevalier (University of Bern), Y. Richet (IRSN), O. Roustant (EMS). *ReDICE Consortium*	
10:30 - 10:50	Coffee Break	HG EO Süd
		Chair: H. Monod
10:50 - 11:50	P. Renard (Université de Neuchâtel). *Multiple-point statistics as a tool to assess complex spatial uncertainty.*	HG F5
11:50 - 12:50	J. Irving (University of Lausanne). *Stochastic inverse methods for near-surface geophysical problems.*	
12:50 - 14:15	Lunch Break	HG EO Süd
		Chair: B. Iooss
14:15 - 15:15	J. Wiart (Orange Labs). *Contribution of statistics to the numerical assessment of the electromagnetic fields human exposure.*	HG F5
15:15 - 15:30	Closure	HG F5

Invited Speakers

List of Abstracts

Exploiting Sparsity in Bayesian Inverse Problems of Parametric Operator Equations

Dr. Claudia Schillings and Prof. Dr. Christoph Schwab
ETH Zürich

Affiliation: *Seminar für Angewandte Mathematik*, ETH Zürich, Rämistrasse 101, CH-8092 Zürich, Switzerland

Email: christoph.schwab@sam.math.ethz.ch – **URL**: http://www.sam.math.ethz.ch/~schwabp

Abstract:

In this talk, we will discuss a parametric deterministic formulation of Bayesian inverse problems with distributed parameter uncertainty from infinite dimensional, separable Banach spaces, with uniform prior probability measure on the uncertain parameter. The underlying forward problems are parametric, deterministic operator equations, and computational Bayesian inversion is to evaluate expectations of quantities of interest under the Bayesian posterior, conditional on given noisy observational data. For forward problems belonging to a certain sparsity class, we quantify analytic regularity of the Bayesian posterior and prove that the parametric, deterministic density of the Bayesian posterior belongs to the same sparsity class. These results suggest in particular dimension independent convergence rates for data-adaptive Smolyak integration algorithms. The proposed approach is applicable for instance for definite or indefinite elliptic and for parabolic evolution problems, with scalar or tensorial unknowns and also with uncertainty in domains, and to highdimensional initial value problems with uncertain coefficients.

This work is supported by the European Research Council under FP7 Grant AdG247277.

Discrete least square polynomial approximations for high dimensional uncertainty propagation

Prof. Dr. Fabio Nobile
EPFL Lausanne

Affiliation: *CADMOS Chair of Scientific Computing and Uncertainty Quantification*, EPFL Lausanne, Station 8, CH-1015 Lausanne, Switzerland

Email: fabio.nobile@epfl.ch – **URL**: http://csqi.epfl.ch/

Abstract:

We consider a general problem $F(u, y) = 0$ where u is the solution of the problem and y a set of uncertain parameters. We specifically address the situation in which the parameter-to-solution map $u(y)$ is smooth, however y could be very high (or even infinite) dimensional. In particular, we are interested in cases in which F is a differential operator and y a distributed, space and/or time varying, random field. We aim at effectively propagate the input uncertainty on y onto the solution u of the problem.

In this talk, we consider polynomial approximations of the parameter-to-solution map $u(y)$, obtained by discrete least square approximation starting from random evaluations.

We discuss the stability and convergence properties of the method as well as possible strategies to select, either a-priori or by adaptive algorithms, sequences of approximating polynomial spaces that allow to reduce, and in some cases break, the curse of dimensionality.

Machine learning for personalized genomics

Prof. Dr. Jean-Philippe Vert
Mines ParisTech

Affiliation: *Centre for Computational Biology*, Mines ParisTech, 35 rue Saint-Honore, 77305 Fontainebleau cedex, France

Email: Jean-Philippe.Vert@mines.org – **URL**: http://cbio.ensmp.fr/~jvert/

Abstract:

The development of DNA sequencing technologies allow us to collect large amounts of molecular data about the genome of each individual, and opens the possibility to predict drug response or evaluate the risk of various diseases from one's molecular identity. In this talk I will discuss some regularization-based approaches we have developed to estimate complex, high-dimensional predictive models from relatively few samples, in particular in cancer prognosis and toxicogenetics.

Focusing exploration with confidence

Prof. Dr. Andreas Krause
ETH Zürich

Affiliation: *Learning & Adaptive Systems Group*, ETH Zürich, Universitätstrasse 6, CH-8093 Zürich, Switzerland

Email: krausea@ethz.ch – **URL**: http://las.ethz.ch/krausea.html

Abstract:

In many applications, ranging from autonomous experimental design to robotic monitoring to system tuning, we wish to gather information about some unknown function. Often, acquiring samples is noisy and expensive. In this talk, I will discuss how Bayesian confidence bounds can play a natural role in focusing exploration: Reducing uncertainty in a structured way to reliably estimate properties of interest such as extremal values, location of critical regions, Pareto-frontiers etc. First, I will show how a simple confidence-guided sampling rule attains near-minimal regret for bandit problems involving objectives modeled via Gaussian process priors. I will further demonstrate how the approach allows to scale up through parallelization, effectively localize level-sets, and address multi-objective tradeoffs. I will illustrate the approach in several real-world applications. Applied to experimental design for protein structure optimization, our approach enabled engineering of active P450 enzymes that are more thermostable than any previously made by chimeragenesis, rational design, or directed evolution.

Simulation-based, high-dimensional stochastic optimization: application in robust topology optimization under large material uncertainties

Prof. Dr. Faidon-Stelios Koutsourelakis
Technische Universität München

Affiliation: *Continuum Mechanics Group*, Technische Universität München, Boltzmannstr. 15, DE-85748 Garching, Germany

Email: p.s.koutsourelakis@tum.de – **URL**: http://www.contmech.mw.tum.de/

Abstract:

This talk is concerned with the optimization/design/control of complex systems characterized by high-dimensional uncertainties and design variables. While analogous problems in a deterministic setting, and particularly in the context of PDE-based models, have been extensively studied and several algorithmic tools have been developed, their extension to stochastic settings poses several challenges. We discuss two alternative strategies. The first is based on stochastic approximation tools [1]. We discuss Variational Bayesian approximations that enable the estimation of gradients in a manner that reduces the sampling noise and the computational effort. The second approach reformulates the problems as one of probabilistic inference [2] and employs sampling tools suitable for high-dimensions [3, 4]. We are especially concerned with problems relating to random heterogeneous materials where uncertainties arise from the stochastic variability of their properties.

References:

[1] H. Robbins and S. Monro (1951). "A stochastic approximation method." *The Annals of Mathematical Statistics*, 22(3), 400–407.

[2] P. Müller (1998). "Simulation based optimal design." *Proceedings of the Sixth Valencia International Meeting*, pp. 323–341.

[3] H. Kück and N. de Freitas and A. Doucet (2006). "SMC Samplers for Bayesian Optimal Nonlinear Design." *Nonlinear Statistical Signal Processing Workshop (NSSPW).*

[4] R. Sternfels and P. S. Koutsourelakis (2011). "Stochastic design and control in random heterogeneous materials." *Journal of Multiscale Computational Engineering*, 9(4), 425–443.

UQLab: a framework for Uncertainty Quantification in Matlab

Dr. Stefano Marelli and Prof. Dr. Bruno Sudret
ETH Zürich

Affiliation: *Chair of Risk, Safety and Uncertainty Quantification*, ETH Zürich, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland

Email: marelli@ibk.baug.ethz.ch – **URL**: http://www.ibk.ethz.ch/su/people/marellis

Abstract:

The general formulation of Uncertainty Quantification covers a vast field of applications including, among others, structural reliability, sensitivity analysis, reliability-based design optimization and Bayesian techniques for calibration and validation of computer models. Many computational tools have been developed by the main scientific communities involved (i.e., engineering, statistics and applied mathematics) but only a few are readily available (*e.g.* free software like FERUM[2], OpenTURNS[1] and Dakota[3], or commercial ones like COSSAN and Nessus) while none of them covers the broad scope mentioned above.

Adopting the global theoretical framework developed in [4] and [5], the UQLab project aims at developing a complete Matlab-based software framework for uncertainty quantification. Such software would allow researchers and field engineers to easily use and further develop uncertainty quantification algorithms on a variety of distributed high-performance computing (HPC) environments.

The platform comprises a highly modular, optimized core probabilistic modelling engine (schematized in Figure 1) and a content management system that allows users to develop additional custom modules to suit their needs.

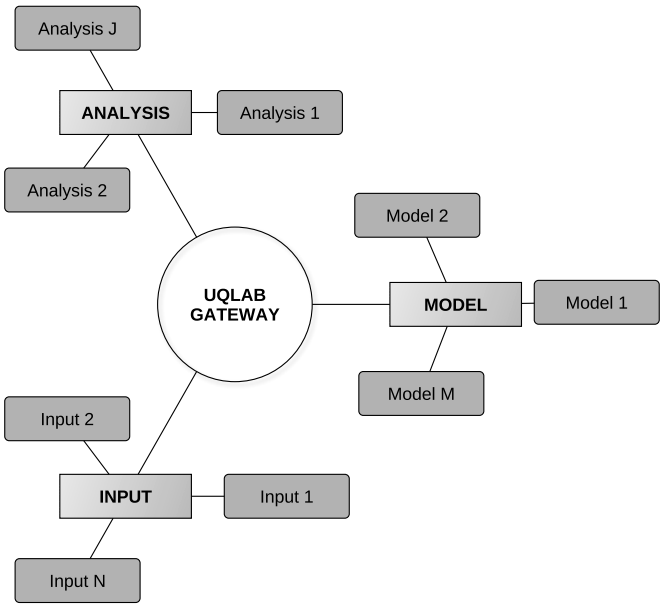


Figure 1: The modular core at the heart of the UQLab framework architecture. An arbitrary number of elements can be connected at any stage of the uncertainty quantification problem.

Ease of use and development for both academic researchers and industrial end-users without extensive information technology (IT) background are central to the design philosophy of UQLab, resulting in the choice of Matlab as the main programming language. Efficient re-use of existent modelling/analysis

codes is promoted by the possibility of easily adding interfaces (wrappers) to existing software (e.g., FEM modelling codes), and by extensive import/export facilities.

At the current state of development, UQLab includes state-of-the art modules for the representation of random vectors (many marginal types, copula-based dependence structure), for structural reliability analysis (FORM, SORM, Importance Sampling, etc.), for sensitivity analysis (Sobol' indices, principal effects, etc.) and a powerful HPC-ready modelling infrastructure (including metamodelling tools like Polynomial Chaos Expansions, Kriging, etc.).

References:

[1] G. Andrianov, Burriel, S., Cambier, S., Dutfoy, A., Dutka-Malen, I., de Rocquigny, E., Sudret, B., Benjamin, P., Lebrun, R., Mangeant, F. & Pendola, M., Open TURNS, an open source initiative to Treat Uncertainties, Risks'N Statistics in a structured industrial approach. *Proceedings of the ESREL'2007 Safety and Reliability Conference*, Stavenger: Norway, 2007

[2] J.-M. Bourinet, Mattrand, C. and Dubourg, V. A., Review of recent features and improvements added to FERUM software, Proceedings of the 10th International Conference on Structural Safety and Reliability, ICOSSAR'09, 2009.

[3] Giunta, A. A., van Bloemen Waanders, B. G., Wojtkiewicz, S. F., Hart, W. E., and Alleva, M. P. (2007). DAKOTA, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 4.1 reference manual. Albuquerque, NM: Sandia National Laboratories.

[4] Sudret, B., Uncertainty propagation and sensitivity analysis in mechanical models - Contributions to structural reliability and stochastic spectral methods. Habilitation à diriger des recherches, Université Blaise Pascal, Clermont-Ferrand,France, 2007, `http://www.ibk.ethz.ch/su/publications/Reports/HDRSudret.pdf`.

[5] E. De Rocquigny, Devictor, N. and Tarantola, S., Uncertainty in industrial practice – A guide to quantitative uncertainty management. John Wiley & Sons, 2008

Incorporating structural priors in Gaussian random field models

Dr. David Ginsbourger
University of Bern

Affiliation: *Institute of Mathematical Statistics and Actuarial Science*, University of Bern, Sidlerstrasse 5, CH-3012 Berne, Switzerland

Email: david.ginsbourger@stat.unibe.ch – **URL**: http://www.ginsbourger.ch

Abstract:

Gaussian random field models have become commonplace in the design and analysis of costly experiments. Thanks to convenient properties of associated conditional distributions, Gaussian field models not only allow approximating deterministic functions based on scarce evaluation results, but can also be used as a basis for evaluation strategies dedicated to optimization, inversion, uncertainty quantification, probability of failure estimation, etc.

In this talk, we will mainly focus on two recent contributions that concern the incorporation of so-called structural priors in Gaussian random field models.

First, results on covariance-driven pathwise invariances of random fields will be presented. Simulation and regression examples will illustrate how Gaussian random field models can incorporate a number of structural priors such as group invariances or harmonicity.

Second, these results will be extended and applied to global sensitivity analysis. In particular, we will present a functional ANOVA decomposition of covariance kernels, and discuss the interplay between sparsity properties of the covariance kernel and of corresponding Gaussian random field paths.

Based on scientific collaborations with Nicolas Durrande, Nicolas Lenz, Olivier Roustant, and Dominic Schuhmacher.

Data driven, molecular dynamics for nanoscale fluid mechanics

Prof. Dr. Petros Koumoutsakos
ETH Zürich

Affiliation: *Computational Science and Engineering Laboratory*, ETH Zürich, Clausiusstrasse 33, CH-8092 Zürich, Switzerland

Email: petros@ethz.ch – **URL**: http://www.cse-lab.ethz.ch/

Abstract:

For over five decades, molecular dynamics (MD) simulations have helped to elucidate critical mechanisms in a broad range of physiological systems and technological innovations. MD simulations are synergetic with experiments, relying on measurements to calibrate their parameters and probing "what if scenarios" for systems that are difficult to investigate experimentally. However, in certain systems, such as nanofluidics, the results of experiments and MD simulations differ by several orders of magnitude. This discrepancy may be attributed to the spatiotemporal scales and structural information accessible by experiments and simulations. Furthermore, MD simulations rely on parameters that are often calibrated semiempirically, while the effects of their computational implementation on their predictive capabilities have only been sporadically probed. In this work, we show that experimental and MD investigations can be consolidated through a rigorous uncertainty quantification framework. We employ a Bayesian probabilistic framework for large scale MD simulations of graphitic nanostructures in aqueous environments. We assess the uncertainties in the MD predictions for quantities of interest regarding wetting behavior and hydrophobicity. We focus on three representative systems: water wetting of graphene, the aggregation of fullerenes in aqueous solution, and the water transport across carbon nanotubes. We demonstrate that the dominant mode of calibrating MD potentials in nanoscale fluid mechanics, through single values of water contact angle on graphene, leads to large uncertainties and fallible quantitative predictions. We demonstrate that the use of additional experimental data reduces uncertainty, improves the predictive accuracy of MD models, and consolidates the results of experiments and simulations.

References:

[1] P. Angelikopoulos and C. Papadimitriou and P. Koumoutsakos (2013). "Data Driven, Predictive Molecular Dynamics for Nanoscale Flow Simulations under Uncertainty." *The Journal of Physical Chemistry B*, 117(47), 14808–14816.

[2] P. Angelikopoulos and C. Papadimitriou and P. Koumoutsakos (2012). "Bayesian uncertainty quantification and propagation in molecular dynamics simulations: A high performance computing framework." *The Journal of Chemical Physics*, 137(14), 144103.

Multiple-point statistics as a tool to assess complex spatial uncertainty

Prof. Dr. Philippe Renard
Université de Neuchâtel

Affiliation: *Stochastic Hydrogeology Group*, Université de Neuchâtel, Rue Emile-Argand 11, 2000 Neuchâtel, Switzerland

Email: philippe.renard@unine.ch – **URL**: http://www2.unine.ch/cms/lang/fr/pid/28514

Abstract:

In most engineering applications related to the underground, heterogeneity associated with lack of data implies uncertainties. When in addition one has to predict fluid flow (groundwater, oil, gaz, CO2) in those environments, a key controlling factor is the connectivity of the highly permeable or impermeable structures. Even if major progresses have been done over the last 50 years in the field of multi-Gaussian random processes to estimate uncertainty, a broader class of random functions is required to be able to better describe geological uncertainty. In this talk, I will introduce the multiple-points statistic framework and some of its recent advances. The basic principle is to start from a training data set from which non parametric statistics are derived on pattern occurrences. Those statistics control the generation of random fields which are able to mimic realistic structures present in the training set. While the algorithm is pretty obvious, it allows simulating a broad range of random processes. Examples will be shown for the simulation of porous medium, geomorphology of braided rivers, rainfall time series, or multivariate satellite images.

Stochastic inverse methods for near-surface geophysical problems

Dr. James Irving
University of Lausanne

Affiliation: *Applied and Environmental Geophysics Group*, Université de Lausanne, Géopolis Building, Quartier Mouline, CH-1015 Lausanne, Switzerland

Email: james.irving@unil.ch – **URL**: http://www.unil.ch/ig/page84045_fr.html

Abstract:

Geophysical methods have wide-reaching applications in near-surface environmental and hydrological research because of their ability to provide estimates of spatially distributed subsurface physical properties at a scale and resolution that are often highly relevant to corresponding modeling and prediction needs. Over the past two decades, significant advancements have been made in terms of obtaining hydrologically and/or environmentally relevant information from geophysical data, most notably addressing the challenge that geophysically derived properties (e.g., electrical conductivity, seismic wave velocity) are typically poorly linked with the subsurface properties of direct interest to the problem at hand (e.g., permeability, contaminant concentration). One key area of near-surface geophysical research that is still in its infancy, however, involves uncertainty quantification of the corresponding inverse problems. Knowledge regarding spatially distributed parameter uncertainty is absolutely essential for making predictions based on geophysical data, which in turn serves as the basis for effective decision-making.

Uncertainty quantification (UQ) for geophysical problems is notoriously challenging because of the typically high dimension of the model parameter space combined with the computational complexity of the associated numerical simulation processes. Although a wide variety of UQ approaches have been developed, one methodology that has become increasingly popular in recent years involves stochastic sampling from the Bayesian posterior parameter distribution using Markov-chain-Monte-Carlo (MCMC) methods. In this presentation, I will summarize work that we have conducted in this overall domain of research with applications to vadose and saturated zone hydrological problems. In addition, I will present recent developments with respect to the incorporation of geostatistical information into Bayesian inverse problems through sequential resampling, along with an evaluation of the numerical efficiency of this approach. Finally, the issue of model error in Bayesian-MCMC inverse approaches will be investigated, with a view towards potential strategies for accounting for the bias of simplified, computationally efficient forward models in the calculated posterior statistics.

Contribution of Statistics to the Numerical Assessment of the Electromagnetic Fields Human Exposure

Dr. Joe Wiart,
Main Contributors: Pierric Kersaudy, Marjorie Jala, Amal Gahmni, Emmanuelle Conil, Abdelhamid Hadjem and Nadege Varsier
Orange Labs & WHIST Lab

Affiliation: *Orange Labs*, 38 Rue du Gnral Leclerc, 92130 Issy-les-Moulineaux, France

Email: joe.wiart@orange-ftgroup.com – **URL**: http://whist.institut-telecom.fr/

Abstract:

Exposure Assessment and Deterministic Methods

In spite of the existing international protection limits (e.g. those of the ICNIRP International Commission) and despite the fact that none of the research results have, today, proven any health effects below these limits, the public fears towards electromagnetic fields (EMF) and the radiofrequencies in particular, remains important even if in the same time there is an increasing and intensive use of wireless communication systems. Today billions of people are using wireless systems such as mobile phone, tablet or computer. In such context the EMF exposure, quantified using the Specific Absorption Rate (SAR), must be monitored carefully. The works that have been conducted for twenty years [1, 2], have shown that morphologies, postures, source location and Radio Frequency (RF) bands used are influencing the human exposure. In the 90's the first objective was to define methodologies able to check the compliance (to the recommended limits) of the EMF emitted by wireless communication equipment (e.g. mobile or base station). To achieve this objective and overcome the possible large variability of the exposure, "worst-case" scenarios (e.g. phantom head) have been designed by international bodies such as CENELEC, IEC or IEEE. These protocols and methods have been successfully used to define such compliance tests. Unfortunately such approach, if useful for compliance checking, is not suitable to assess the real personal exposure [3]. Such "worst case" approach has been first challenged with the epidemiological studies (e.g. the international INTERPHONE study) that request the real exposure and not a maximum. The need of a comprehensive exposure assessment has been reinforced with the increasing and versatile use of the wireless communication systems. Today these systems can emit in different RF bands and in different ways depending on the application, they can be used in versatile use (listening, watching, ...).

Contribution of Statistics to Human Exposure Assessment

In the 2000s new approaches, involving statistics, have been investigated to respond to these questions and be able to handle new usages such as those of children who, at earlier and earlier ages, use tremendously wireless communication systems. The numerical method, mainly used in bio-electromagnetism, to estimate the human exposure induced by EMF, is the finite difference in time domain (FDTD). This method is based on an explicit iterative calculation of the electric and magnetic field (linked by the well known Maxwell equations) over an orthogonal grid. The main advantage of the FDTD is to proceed without any matrix inversion that can be cumbersome. The main problem of the FDTD is the time computation that can be very large (i.e. few hours if calculation involved the whole body). Large efforts have been achieved toward high performance calculation using parallel computation and recently graphic processors unit. But even with these efforts, the time computation is still important and is not compatible with Monte Carlo Method. To overcome this issue, efforts have been dedicated to build simplified surrogate models. Works have be done to use the polynomial expansion with the coefficient of the polynomials estimated using projection methods and sparse grids[4]. The results of these studies showed that this approach is still not compatible with the FDTD even when reducing the number of computation. For instance a quite simple problem such as a phone close to the head, involving 4 input parameters, governing angles and location, can request a few hundred of simulations. To reduce the number of simulations and taking into account previous work carried out in mechanics, a sparse polynomial chaos (PC) expansion (using a least angle regression method) combined with a regression and an iterative planning of experiments has been

used to build a surrogate model with a minimum calculation [5]. This approach has been able to reduce significantly the number of simulations but has limited ability to manage the accuracy of quantile estimation. Due to the influence of the morphology [2] this quantile estimation is a key question to determine the threshold of the Whole Body SAR (WBSAR) at 95 % for a given population [6]. A similar problem is the fetus exposure induced by plane wave having arbitrary incidence [7]. In these cases the challenge is to build a parsimonious iterative planning experiments able to select additional calculations increasing the accuracy of the quantile estimation. To achieve this objective, Gaussian Process and sequential planning numerical experiments method [6, 7] have been studied. Gaussian Process Shrunk and existing Stepwise Uncertainty Reduction have been studied to select iteratively new calculations with a strong constraint on the minimum of calculations to be performed. The uncertainty estimation obtained with PC is often a global one (using bootstrap or leave one out methods) and since method such as Kriging has shown a good ability to address this problem, on going work are performed to combine Kriging and PC to build an efficient and parsimonious iterative planning able to assess quantiles and the associated uncertainties.

Conclusions

With the versatile use of wireless, the exposure assessment have to handle complex and variable configurations. The usual deterministic approaches are not able to handle such variability and dosimetric calculations can be cumbersome. As described in previous sections, statistical methods allow one to handle complex and variable configurations that cannot be managed using the regular deterministic approaches, These methods have shown a good capability, in bio-electromagnetism where the number of input can be important. During the presentation, examples will be given to illustrate the challenges and methods used to characterize statistical distribution of the local and whole body human exposure.

References:

[1] J. Wiart, A. Hadjem, N. Varsier, E. Conil. Numerical dosimetry dedicated to children RF exposure. Progress in Biophysics and Molecular Biology, 2013, vol 107, Issue 3

[2] E. Conil, A. Hadjem, F. Lacroux, M. F. Wong and J. Wiart. Variability analysis of SAR from 20 MHz to 2.4 GHz for different adult and child models using finite-difference time-domain. Phys. Med. Biol (2008). 53 1511-1525

[3] T. Kientega, E. Conil, A. Hadjem, E. Richalot, A. Gati, M.-F. Wong, O. Picon, J. Wiart. A surrogate model to assess the whole body SAR induced by multiple plane waves at 2.4 GHz. Ann. Telecommun. (2011) 66:419–428

[4] J. Silly-Carette, D. Lautru, M.-F. Wong, A. Gati, J. Wiart and V. Fouad Hanna. Variability on the Propagation of a Plane Wave Using Stochastic Collocation Methods in a Bio Electromagnetic Application. IEEE MWC letters (2009) Vol. 19, N. 4.

[5] A. Ghanmi, PHD Analyse de l'exposition aux ondes électromagnétiques des enfants dans le cadre des nouveaux usages et nouveaux réseaux. 2013

[6] A. El Habachi , E. Conil , A. Hadjem , E. Vazquez , M. F. Wong , A. Gati , G. Fleury and J. Wiart. Statistical analysis of whole-body absorption depending on anatomical human characteristics at a frequency of 2.1 GHz. Phys. Med. Biol. (2010) 55 pp 1875 -1887

[7] Marjorie Jala, Céline Lvy-Leduc, Eric Moulines, Emmanuelle Conil and Joe Wiart. "Sequential Design of Computer Experiments for the Assessment of Fetus Exposure to Electromagnetic Fields".

Presentations

List of Abstracts

Multi-level Monte Carlo Finite Volume methods for stochastic systems of hyperbolic conservation laws

J. Šukys, S. Mishra, Ch. Schwab
ETH Zürich

Affiliation: *Seminar for Applied Mathematics*, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Email: jonas.sukys@sam.math.ethz.ch – **URL**: http://www.sam.math.ethz.ch/~sukysj/

Master: ETH Zürich

Ph.D. (2011-2014): ETH Zürich

Supervisor(s): Prof. Siddhartha Mishra (ETH Zürich) and Prof. Christoph Schwab (ETH Zürich)

Abstract:

A large number of problems in physics and engineering such as global climate, propagation of tsunamis and avalanches, waves in the solar atmosphere, design of efficient aircraft, and the structural mechanics are modeled by systems of non-linear partial differential equations termed as *systems of balance laws*:

$$\begin{cases} \mathbf{U}_t(\mathbf{x},t) + \operatorname{div}(\mathbf{F}(\mathbf{x},\mathbf{U})) = \mathbf{S}(\mathbf{x},t,\mathbf{U}), \\ \mathbf{U}(\mathbf{x},0) = \mathbf{U}_0(\mathbf{x}), \end{cases} \quad \forall (\mathbf{x},t) \in \mathbf{D} \times \mathbb{R}_+. \tag{1}$$

Examples of conservation laws include the shallow water equations of oceanography, the Euler equations of aerodynamics, the Magnetohydrodynamics (MHD) equations of plasma physics and the equations of elasticity. In general, solutions of (1) develop *shock waves* in finite time even for smooth initial data. Hence, solutions are sought in the sense of distributions, additionally imposing *entropy conditions* to ensure uniqueness. For fluxes that are *non-linear* or with *varying coefficients*, analytical solution formulas are only available in very special cases. Apart from many other numerical methods, Finite Volume methods (FVM) emerged as the the most successful paradigm for practical computations in geophysics, aerodynamics and astrophysics. In FVM, the corresponding numerical fluxes are based on (approximate) solutions of Riemann problems at mesh cell interfaces. Higher order spatial accuracy is obtained by non-oscillatory reconstruction procedures such as TVD limiting, (W)ENO or by the Discontinuous Galerkin (DG) method; higher order temporal accuracy is obtained by SSP-RK method.

Uncertainty quantification. A Finite Volume scheme requires the initial data, fluxes, and source terms as inputs. These inputs are, in general, uncertain, i.e., initial condition $\mathbf{U}_0 = \mathbf{U}_0(x,\omega)$, source term $\mathbf{S} = \mathbf{S}(x,\omega,\mathbf{U})$ and fluxes $\mathbf{F} = \mathbf{F}(\mathbf{x},\omega,\mathbf{U})$ are random fields where $\omega \in \Omega$ and $(\Omega,\Sigma,\mathbb{P})$ denotes a complete probability space. Consequently, the solution $\mathbf{U}$ is sought as the *random entropy solution* of the *random balance law*:

$$\begin{cases} \mathbf{U}(x,t,\omega)_t + \operatorname{div}(\mathbf{F}(\mathbf{x},\omega,\mathbf{U})) = \mathbf{S}(\mathbf{x},\omega,\mathbf{U}), \\ \mathbf{U}(x,t,\omega) = \mathbf{U}_0(x,\omega), \end{cases} \quad (\mathbf{x},t) \in \mathbf{D} \times \mathbb{R}_+, \quad \omega \in \Omega. \tag{2}$$

Under certain assumptions on input data $\mathbf{U}_0$, $\mathbf{S}$, $\mathbf{F}$, the *existence* of the *k-th statistical moments* $\mathcal{M}^k(\mathbf{U})$ of the *random entropy solution* is established. The next step is the design of efficient numerical methods for the approximation of the *random balance law* (2). These methods include the stochastic Galerkin, stochastic collocation and stochastic Finite Volume. Currently these methods are not able to handle large number of uncertainty sources, are *intrusive* (existing deterministic solvers need to be reconfigured) and hard to parallelize. Hence, we focus on the sampling-type Monte Carlo methods.

Multi-Level Monte Carlo Finite Volume Method

Due to the slow convergence of the conventional Monte Carlo FVM sampling methods, we propose the Multi-Level Monte Carlo method (MLMC-FVM). MLMC was introduced by Giles for Itô SPDE. The key idea is to simultaneously draw MC samples on a hierarchy of nested grids:

0. **Nested meshes:** Consider *nested* meshes $\{\mathcal{T}_\ell\}_{\ell=0}^{\infty}$ of the domain $\mathbf{D}$ with corresponding mesh widths $\Delta x_\ell = 2^{-\ell}\Delta x_0$, where Δx_0 is the mesh width of the coarsest resolution.
1. **Sample:** For each level $\ell \in \mathbb{N}_0$, we draw M_ℓ independent identically distributed (i.i.d) samples $\mathbf{I}_\ell^i$ with $i = 1, \ldots, M_\ell$ from the random input data $\mathbf{I}(\omega)$ and approximate these by cell averages.
2. **Solve:** For each level ℓ and each realization $\mathbf{I}_\ell^i$, the balance law (1) is solved for $\mathbf{U}^{i,n}_{\Delta x_\ell}$ and $\mathbf{U}^{i,n}_{\Delta x_{\ell-1}}$ by the FVM method on meshes $\mathcal{T}_\ell$ and $\mathcal{T}_{\ell-1}$ with mesh widths Δx_ℓ and $\Delta x_{\ell-1}$, respectively.
3. **Estimate Statistics:** Fix $L < \infty$ corresponding to the highest level. Denoting MC estimator with $M = M_\ell$ by E_{M_ℓ}, the expectation of the random solution field $\mathbf{U}$ is estimated by

$$E^L[\mathbf{U}^n_{\Delta x_L}] := \sum_{\ell=0}^{L} E_{M_\ell}[\mathbf{U}^n_{\Delta x_\ell} - \mathbf{U}^n_{\Delta x_{\ell-1}}]. \tag{5}$$

To equilibrate the statistical and the spatio-temporal errors, we require $M_\ell = \mathcal{O}(2^{2(L-\ell)s})$ for $0 \le \ell \le L$. Notice that that the largest number of MC samples is required on the coarsest mesh level $\ell = 0$, whereas only a few MC samples are needed for $\ell = L$. Using such M_ℓ, we obtained the error vs. work estimate

$$\|\mathbb{E}[\mathbf{U}(t^n)] - E^L[\mathbf{U}^n_{\Delta x_L}](\omega)\|_{L^2(\Omega,\cdot)} \lesssim \begin{cases} (\text{Work})^{\min\{-s/(d+1),1/2\}} \cdot \log(\text{Work}) & \text{if } s \neq (d+1)/2, \\ (\text{Work})^{-1/2} \cdot \log(\text{Work})^{3/2} & \text{if } s = (d+1)/2. \end{cases} \tag{7}$$

The above estimate (7) shows that the MLMC-FVM is superior to the MC-FVM. For $s < (d+1)/2$, estimate (7) is exactly of the same order (modulo a log term) as the estimate for the *deterministic* FVM.

Parallel implementation. We have developed a massively parallel code ALSVID-UQ, which implements the MLMC-FVM algorithm to solve the systems of stochastic balance laws (2). We designed novel *static* and *adaptive* load balancing procedures and achieved linear (strong and weak) scaling up to 40 000 cores.

Numerical example. We consider three-dimensional Euler equations in domain $\mathbf{D} = [0,1]^3$ and the so-called *cloud-shock* initial data with 11 sources of uncertainty, i.e. with *random* initial shock at *random* location (near $x = 0.1$) heading towards high density cloud with uncertain shape of its boundary and uncertain inner density. The mean and variance for the density of the solution at time $t = 0.06$ are shown in Figure 1. The results are from a MLMC-WENO run with 7 nested levels of resolution ($L = 6$) and the finest resolution is set to 1024^3 mesh. The flow in this case consists of the supersonic initial shock moving to the right, interacting with the high density bubble and leading to a complex flow pattern that consists of a leading bow shock, trailing tail shocks and a very complex center region possessing sharp gradients as well as turbulent like smooth features. Runtime: 5 hours on 21 844 cores.

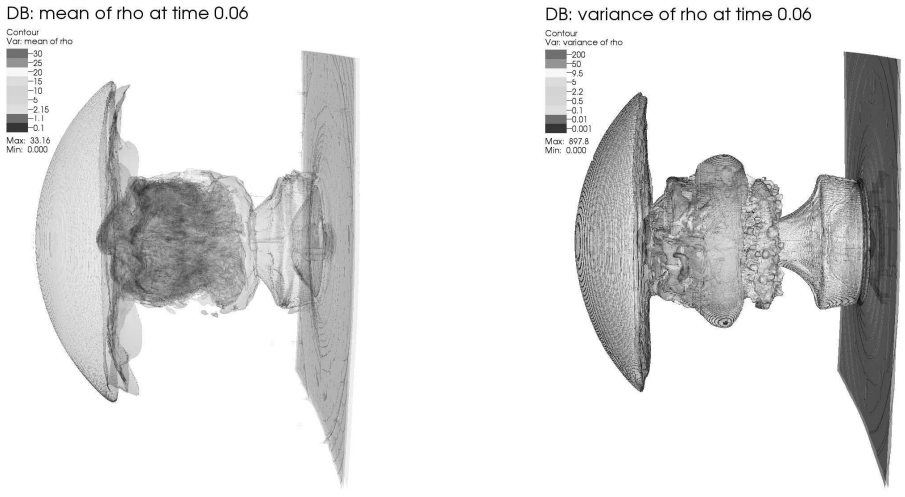


Figure 1: Mean and variance of density in the cloud-shock estimated with MLMC-FVM.

Short biography – I have obtained my BSc in Mathematics in Jacobs University Bremen, Germany, and my MSc in ETH Zürich. My scientific interests are: hyperbolic nonlinear stochastic partial differential equations, numerical analysis and simulations, massively parallel high performance computing, finite volume methods, multi-level Monte Carlo methods. More information under: `http://cv.sukys.lt`.

Sources of funding: ETH grant CH1-03 10-1, Swiss National Supercomputing Center project S366.

Adaptive particle methods for rare event simulation in a Markovian framework

D. Jacquemart
University Rennes I

Affiliation: *Onera, the French Aerospace Lab*, 91123 Palaiseau, France

Email: damien.jacquemart@onera.fr – **URL**: http://www.inria.fr/en/teams/aspi

Master: University Aix-Marseille III

Ph.D. (2011-2014): University Rennes I

Supervisor(s): Prof. François Le Gland (INRIA, Rennes) and Dr. Jérôme Morio (Onera, Toulouse)

Abstract:

In recent decades, standards of quality and safety requirements is increasingly demanding in numerous industrial and scientific areas. Thus, the estimation of probability of rare events is become of great interest. Two popular methods in this regards are the splitting algorithm [1] and a weighted importance resampling algorithm (WIR) [2]. Despite a deep theoretical study of these two algorithms, the practical side has not been handled completely. The choice of a good tuning often proves to be fastidious and may decrease their usefulness. This study aims at answering this point. More precisely, we propose some adaptive methods to make the implementation of the splitting algorithm and the WIR algorithm automatic. Furthermore, we enlarge the area of application of the WIR algorithm to make WIR and the splitting estimating the same kind of probability.

Firstly, we focus on the splitting algorithm. The main goal of this algorithm is to estimate the probability that a Markov process X_t enters a critical set B before some stopping time T. Namely, it gives some numerical approximations of

$$\mathbb{P}(T_B \leq T) = \mathbb{P}(X_t \in B, \text{ for some } t \in [0,T]), \tag{1}$$

where T_B is the first time X_t enters B. This kind of problems typically arises in air traffic management [3], telecommunication networks [4] and electrical grid reliability estimation [5].
The principle of the splitting is to consider a sequence of decreasing supersets of B, $B_1 \supset \cdots \supset B_{m-1} \supset B_m = B$, and to estimate each probability of reaching B_k starting from B_{k-1}. Consequently, we first have to determine the sequence of decreasing supersets B_k. To this aim, the existing algorithms [6, 4] are too restrictive. Indeed, they can only be used with time-homogeneous X_t process and T random. Moreover, the algorithm proposed in [6] is only made for a process X_t that takes its value in $\mathbb{R}$. That is why we propose a general adaptive splitting algorithm for T random or deterministic, X_t possibly time-homogeneous and multi-dimensional. For that purpose, we assume that the rare event is characterized by some exceedance over a given threshold S of a real valued function Φ, $B = \{x, \Phi(x) \geq S\}$. For all we know, all the problems addressed in the scientific literature ([4, 3, 5]) can be expressed in such a way. In that case, we derive an adaptive algorithm for the choice of the supersets B_k, as concisely detailed in [7]. To this end, the proposed algorithm is based on some quantile estimation of the random variable of the maxima of a trajectory of $\Phi(X_t)$ over final time T. Thus, the supersets B_k are implicitly defined with Φ. Finally, we give some numerical estimations of the conflict probability between aircraft.

Secondly, we proposed some improvement of the WIR algorithm [2]. The goal of the WIR algorithm is to estimate

$$\mathbb{P}(\Psi(Z_n) \in C), \tag{2}$$

where $(Z_k,\, k = 0, \ldots, n)$ is a Markov chain, n a fixed integer, Ψ a real valued function and C a subset of $\mathbb{R}$. Such problems appears in the estimation of credit portfolio losses [8] and in fiber optic [2].

To estimate the probability (2), WIR consists in a set of N random paths $(Z_{0:n}^{(i)},\, 1 \leq i \leq N)$. The construction of these paths is performed in two steps. First, at each iteration time, the trajectories which

are more likely to reach the rare set C are multiplied and the others are killed. This is made through a selection function. Secondly, the Markov transition kernel of $(Z_k,\, k = 0, \ldots, n)$ is applied for the trajectory evolution.

To our knowledge, all the selection functions used in the scientific literature for the WIR algorithm only depend on a real parameter, denoted here by α [2]. Indeed, the choice of α strongly influences the variance of the estimated probability. However, there is no formula neither rules on the choice of some good parameters. We thus propose an algorithm to compute some good parameter for the WIR algorithm. This procedure is motivated by an original characterisation of the parameters that should be used. This characterisation is made with the proportion of the trajectories that reach the rare set at final time. Moreover, we numerically show that this procedure achieves to determine the optimal parameters. Then we test the proposed algorithm with the estimation of outage probability in optic fiber and the estimation of large credit portfolio losses.

At first sight, the splitting and WIR algorithms cannot be used for the same estimation problem. Indeed, the first one estimates the probability (1) that a critical event occurs during the evolution of a Markov process whereas the second one estimates the probability (2) that a critical event occurs at the deterministic final evolution time of a Markov chain. However, if one considers the process of the maxima of $\Phi(X_t)$, namely $M_t = \max_{0 \leq s \leq t} \Phi(X_s)$ we show that the WIR algorithm can indeed be used for the estimation of (1). Finally, we compare the efficiency of our modified version of the WIR algorithm and the splitting algorithm on the estimation of the conflict probability. We conclude that the WIR algorithm gives better variance than the splitting algorithm.

References:

[1] P. L'Écuyer, V. Demers, and B. Tuffin. Rare events, splitting, and quasi–Monte Carlo. *ACM Transactions on Modeling and Computer Simulation*, 2007.

[2] P. Del Moral and J. Garnier. Genealogical particle analysis of rare events. *The Annals of Applied Probability*, 15(4):2496–2534, 2005.

[3] M. Prandini, H. Blom, and G.J. Bakker. Air traffic complexity and the interacting particle system method: An integrated approach for collision risk estimation. In *American Control Conference, San Francisco, CA*, 2011.

[4] M. Garvels, J. Van Ommeren, and D. Kroese. On the importance function in splitting simulation. *European Transactions on Telecommunications*, 13(4):363–371, 2002.

[5] W. Wadman, D. Crommelin, and J. Frank. Applying a splitting technique to estimate electrical grid reliability. In *Winter Simulation Conference, Washington, DC*, 2013.

[6] F. Cérou and A. Guyader. Adaptive multilevel splitting for rare event analysis. *Stochastic Analysis and Applications*, 25(2):417–443, 2007.

[7] D. Jacquemart, F. Le Gland, and J. Morio. A combined importance splitting and sampling algorithm for rare event estimation. In *Winter Simulation Conference, Washington, DC*, 2013.

[8] R. Carmona, J-P. Fouque, and D. Vestal. Interacting particle systems for the computation of rare credit portfolio losses. *Finance and Stochastics*, 13(4):613–633, 2009.

Short biography – Damien Jacquemart is a Ph.D student in applied mathematics at ONERA, the French aerospace lab and INRIA. He obtained a Master degree in high education teaching, in 2010. Then, he graduated from the University of Marseille, France with a research Masters degree in probability and statistics, in 2011. His research interests include rare event estimation with aerospace applications. He teaches at ENSAI, the national school of statistics and at the engineering school ENSTA.

Bayesian Multilevel Model Calibration for Inversion of "Perfect" Data in the Presence of Uncertainty

Joseph B. Nagel & Bruno Sudret
ETH Zürich

Affiliation: *Chair of Risk, Safety and Uncertainty Quantification*, ETH Zürich, D-BAUG, Institute of Structural Engineering, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland

Email: nagel@ibk.baug.ethz.ch – **URL**: http://www.ibk.ethz.ch/su/people/nagelj

Master: Rheinische Friedrich-Wilhelms-Universität Bonn

Ph.D. (2012-2015): ETH Zürich

Supervisor(s): Prof. Dr. Bruno Sudret (ETH)

Abstract:

Hierarchical or multilevel modeling establishes a convenient framework for solving complex inverse problems [1, 2] in the presence of uncertainty. In the last two decades it has been studied from a frequentist [3] and a Bayesian perspective [4]. We will adopt a Bayesian point of view to statistical inversion and uncertainty quantification and present a Bayesian multilevel framework that allows for inversion and optimal analysis of "perfect" or noise-free data in the presence of aleatory and epistemic types of uncertainty and in experimental situations when data is scarce or expensive to acquire. In this contribution to the annual MascotNum workshop we will discuss the abovementioned framework on the basis of an application example within the domain of aerospace engineering [5]. We will not only illustrate the very potential of Bayesian multilevel modeling as well as ways to overcome its immanent major challenges, but more importantly we will discuss the main observations, considerations and key questions that the practical problem solution [6] has given rise to.

A forward model $\mathcal{M}\colon (\boldsymbol{m}, \boldsymbol{x}, \boldsymbol{\zeta}, \boldsymbol{d}) \mapsto \tilde{\boldsymbol{y}}$ describes a system or phenomenon under consideration. Throughout a number of $i = 1, \dots, n$ experiments forward model inputs may be represented corresponding to a certain model of epistemic and aleatory uncertainty. There are fixed albeit insufficiently well-known model parameters $\boldsymbol{m}$, model inputs $\boldsymbol{x}$ and $\boldsymbol{\zeta}$ that are subject to imperfectly or perfectly known variability, respectively, and experimental conditions $\boldsymbol{d}$ that are known with certainty. Constant yet unknown model parameters are represented as random variables $\boldsymbol{M} \sim \pi_{\boldsymbol{M}}(\boldsymbol{m})$ where $\pi_{\boldsymbol{M}}(\boldsymbol{m})$ is a Bayesian prior belief about their true values. Model inputs with perfectly known variability are modeled as experiment-specific realizations $\boldsymbol{\zeta}_i$ of random variables $(\boldsymbol{Z}_i|\boldsymbol{\theta}_{\boldsymbol{Z}}) \sim f_{\boldsymbol{Z}|\boldsymbol{\Theta}_{\boldsymbol{Z}}}(\boldsymbol{\zeta}_i|\boldsymbol{\theta}_{\boldsymbol{Z}})$ with known hyperparameters $\boldsymbol{\theta}_{\boldsymbol{Z}}$ that prescribe the variability. Model inputs with imperfectly known variability are modeled as experiment-specific realizations $\boldsymbol{x}_i$ of exchangeable random variables $(\boldsymbol{X}_i|\boldsymbol{\theta}_{\boldsymbol{X}}) \sim f_{\boldsymbol{X}|\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{x}_i|\boldsymbol{\theta}_{\boldsymbol{X}})$ with hyperparameters $\boldsymbol{\theta}_{\boldsymbol{X}}$ about which only Bayesian prior knowledge $\boldsymbol{\Theta}_{\boldsymbol{X}} \sim \pi_{\boldsymbol{\Theta}_{\boldsymbol{X}}}(\boldsymbol{\theta}_{\boldsymbol{X}})$ is available. Experimental conditions $\boldsymbol{d}_i$ possibly differ throughout the experiments yet they are (deterministic) perfectly known values.

A "complex" inverse problem is posed when model responses $\tilde{\boldsymbol{y}}_i = \mathcal{M}(\boldsymbol{m}, \boldsymbol{x}_i, \boldsymbol{\zeta}_i, \boldsymbol{d}_i)$ are measured in n experiments, forward model inputs comply with the aforementioned uncertainty model and inference focuses on the unknowns $(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$. While classical Bayesian multilevel modeling deals with the analysis of "imperfect" data $\boldsymbol{y}_i = \tilde{\boldsymbol{y}}_i + \boldsymbol{\varepsilon}_i$, i.e. model-measurement discrepancy is accounted for by residual terms that are modeled as outcomes $\boldsymbol{\varepsilon}_i$ of a random variables $\boldsymbol{E}_i \sim f_{\boldsymbol{E}_i}(\boldsymbol{\varepsilon}_i)$ with distributions $f_{\boldsymbol{E}_i}(\boldsymbol{\varepsilon}_i)$, the problem formulation at hand deals with "perfect" data $\tilde{\boldsymbol{y}}_i$. Interestingly, in the context present the analysis of "perfect" data is more involved than the analysis of "imperfect" data in mathematical and numerical terms. Thus firstly we will devise a Bayesian multilevel model involving "perfect" data. Subsequently we will show how Bayesian calibration of the formulated multilevel model can be accomplished by analyzing the entirety of collected data $\langle \tilde{\boldsymbol{y}}_i \rangle = (\tilde{\boldsymbol{y}}_1, \dots, \tilde{\boldsymbol{y}}_n)$. The inferential prior distribution $\pi(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$, that represents the knowledge about the quantities of interest $(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}})$ prior to analyzing the data, will be updated in order to obtain the posterior distribution $\pi(\boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}|\langle \tilde{\boldsymbol{y}}_i \rangle)$. To that end a likelihood function $\mathcal{L}(\langle \tilde{\boldsymbol{y}}_i \rangle | \boldsymbol{m}, \boldsymbol{\theta}_{\boldsymbol{X}}; \boldsymbol{\theta}_{\boldsymbol{Z}})$ has to be formulated as well as a means for its efficient evaluation.

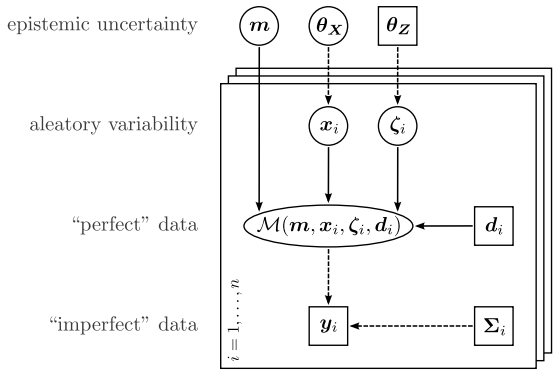


Figure 1: **DAG of the generic multilevel model.** Vertices symbolize unknown (○) or known (□) quantities and directed edges represent their deterministic (⟶) or probabilistic (⇢) relations. Quantities are shown in a way that reflects their uncertainty.

Since for the specific problem at hand such a likelihood function is not available in closed-form, we will propose a statistical simulator of the likelihood which is based on Monte Carlo (MC) sampling and kernel density estimation (KDE). Moreover, in order to explore the posterior of the quantities of interest, we will devise dedicated Markov chain Monte Carlo (MCMC) algorithms. The very principle of MCMC is to construct a Markov chain whose long-run distribution approaches the desired posterior. By virtue of Bayes' law closed-form approximations of the likelihood directly induce approximations on the level of the posterior. However, if calls to the likelihood function $\mathcal{L}$, over the course of the Markov chain, are replaced by calls to a statistical estimator $\hat{\mathcal{L}}$, an approximation is introduced on the level of the Markov chain transition kernel. This raises the distinctly important question as to which degree the induced equilibrium distribution is in conformity with the true posterior, i.e. the issue of *posterior fidelity*. In turn the practical question becomes how to "optimally" tune free algorithmic parameters, e.g. the number of MC samples, the bandwidth of the KDE and parameters of the MCMC simulation. We will present a heuristic way of approaching those delicate issues. Beyond that, we will demonstrate how data augmentation [7] can be utilized in the outlined multilevel context. Data augmentation is a powerful technique from the vast MCMC toolkit that traditionally aims at enhancing MCMC efficiency by introducing hidden data as auxiliary variables. Instead we will herein introduce latent data as auxiliary variables in order to enhance the adequacy of likelihood estimations and the fidelity of the posterior densities that are eventually obtained.

References:

[1] E. de Rocquigny and S. Cambier. Inverse probabilistic modelling of the sources of uncertainty: A non-parametric simulated-likelihood method with application to an industrial turbine vibration assessment. *Inverse Prob. Sci. Eng.*, 17(7):937–959, 2009.

[2] P. Barbillon, G. Celeux, A. Grimaud, Y. Lefèbvre, and E. de Rocquigny. Nonlinear methods for inverse statistical problems. *Comput. Stat. Data Anal.*, 55(1):132–142, 2011.

[3] H. T. Banks, Z. R. Kenz, and W. C. Thompson. A review of selected techniques in inverse problem nonparametric probability distribution estimation. *J. Inverse Ill-posed Prob.*, 20(4):429–460, 2012.

[4] D. J. Lunn. Bayesian analysis of population pharmacokinetic/pharmacodynamic models. In D. Husmeier, R. Dybowski, and S. Roberts, editors, *Probabilistic Modeling in Bioinformatics and Medical Informatics*, Advanced Information and Knowledge Processing, pages 351–370. Springer, London, 2005.

[5] L. G. Crespo, S. P. Kenny, and D. P. Giesy. The NASA Langley multidisciplinary uncertainty quantification challenge, 2013. NASA Langley Research Center.

[6] J. B. Nagel and B. Sudret. A Bayesian multilevel framework for uncertainty characterization and the NASA Langley multidisciplinary UQ challenge. In *Proc. 16th AIAA Non-Deterministic Approaches Conference (SciTech 2014), National Harbor, Maryland*, January 13-17, 2014.

[7] D. A. Dyk and X.-L. Meng. The art of data augmentation. *J. Comp. Graph. Stat.*, 10(1):1–50, 2001.

Short biography – Joseph Benjamin Nagel studied theoretical physics and wrote his diploma thesis about quantum field theories, MCMC simulations and GPU computations. After working on geophysical variational data assimilation he joined the Chair of Risk, Safety and Uncertainty Quantification, where his research focuses on "Bayesian techniques for model calibration and stochastic inverse problems".

Multivariate Quantile Surfaces and Application to an Aircraft Problem

Adil Ahidar
Université Toulouse III - Paul Sabatier

Affiliation: - *Institut de Mathématiques de Toulouse (IMT)*, Université Paul Sabatier 118, route de Narbonne 31062 Toulouse, France
- *Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique (CERFACS)*, 42 Avenue Gaspard Coriolis, 31057 Toulouse, France

Email: adil.ahidar@math.univ-toulouse.fr – **URL**: http://www.math.univ-toulouse.fr

Master: Université Toulouse III - Paul Sabatier

Ph.D. (2011-2014): Université Toulouse III - Paul Sabatier

Supervisor(s): Prof. Philippe Berthet (Université Toulouse III) and Dr. Thierry Klein (Université Toulouse III) and Prof. Serge Gratton (CERFACS) and Dr. Jayant Sengupta (EADS IW)

Abstract:

Introduction During an aircraft design process, an engineer would be interested to distribute margins on different suppliers in an optimal way, i.e. giving more margins to those who do not control their own design process (because of a new technology for instance) and less margins to those who can easily or cheaply commit on a good confidence. The design, influenced by the uncertainty of the suppliers margins, has to comply to some performance constraints taking into account a risk measure. In the case of multiple criteria, this leads to the definition of a multivariate quantile. The main contribution of this PhD thesis is on a new definition and on the associated convergence theroem of a multivariate quantile.

Multivariate Quantile Let $X_1, ..., X_n$ be i.i.d. $\mathbb{R}^d$ valued random variables with common law $P = \mathbb{P}^X$. Many approaches in multivariate data analysis have been proposed to describe the structure of such a data cloud. In connection with our work it is worth mentioning spatial quantiles, data depth, level sets, mode localization, shorth sets, classification, k-means, trimming, among others.

In this presentation we introduce an original notion of multivariate quantiles that is not based on a global M-estimation but rather on a directional M-estimation which is also a natural generalization of the univariate quantiles.

For $d = 1$, let's first recall the definition of the univariate quantile $Q(\alpha) = \inf\{y \in \mathbb{R} : \mathbb{P}(X \leq y) \geq \alpha\}$

Facing the usual fact that $\mathbb{R}^d$ is not ordered our idea is simply to admit subjectivity and thus to define a local viewpoint rather than a global one, anchored at some point of reference O and arbitrary shape φ with the motivation of crossing information gathered by changing viewpoint O, shape φ and α-th order of quantile. Since these viewpoints are correlated, the next key steps will be to properly compare them to automatically learn about P, but this is beyond our current scope. In this presentation we mainly focus on the special case of half-spaces instead of the general shape φ.

For $d \geq 1$, we define the half-space standing at distance $y \in \mathbb{R}$ from $O \in \mathbb{R}^d$ in the direction $u \in \mathbb{S}_{d-1}$

$$H(O,\, u,\, y) = \left\{x \in \mathbb{R}^d : \langle x - O,\, u\rangle \leq y\right\}$$

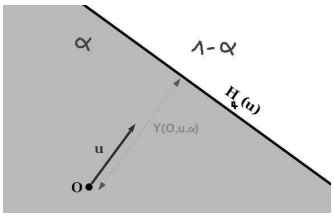
now, given $\alpha \in (1/2, 1]$, we define the α-th quantile range from O in the direction u

$$Y(O, u, \alpha) = \inf\{y \in \mathbb{R} : \mathbb{P}(H(O,\, u,\, y)) \geq \alpha\}$$

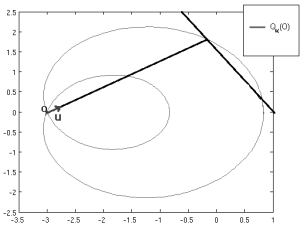
and the α-th quantile point seen from O in the direction u

$$Q(O, u, \alpha) = O + Y(O, u, \alpha) \cdot u$$

Imagine an observer located in $O \in \mathbb{R}^d$ looking at the sample in all directions $u \in \mathbb{S}_{d-1}$ where $\mathbb{S}_{d-1}$ is the unit sphere of $\mathbb{R}^d$. How can he picture out the empirical mass localization in $\mathbb{R}^d$? We propose to draw a



(a) in pink: space region of mass α
in black: half-space H(O,y,α)



(b) in red: The quantile surface
$Q(O,\alpha) = \{Q(O,u,\alpha) : \mathbb{S}_{d-1}\}$

collection of u-directional α-th quantile points that we call a subjective α-th quantile surface, for a fixed $\alpha \in (1/2, 1)$.

Each point of this surface (b) is the univariate α-th quantile of the projected distribution on the line (O, u). How can the observer catch more information on the law P? Keeping O fixed and letting α vary determines P. Thus analyzing a collection of such surfaces can be viewed as some purely non-parametric and non-linear data analysis approach in which we keep track of more than a few orthogonal directions. A second motivation for these α-th quantile surfaces is that the algorithms computing them at a reasonable precision are simple and fast, even for large samples, faster than level sets estimators or plug-in estimators based on density estimators – however density may even not exist.

Convergence Results The graphical representations of these spatial quantiles are random closed surfaces generated by the sample for which we establish the following theorems:

1- Almost sure consistency ; 2- Uniform central limit theorem with some rate ; 3-Uniform Strong approximation with rate ; 4- Bahadur-Kiefer type representation

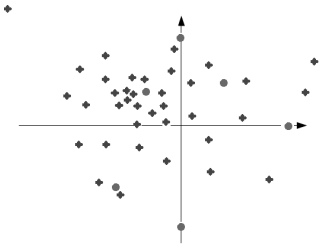
Back to the industrial problem The industrial problem can be reformulated (or generalized) in the folowing way:

$$\begin{cases} \max_{\sigma \in \mathbb{R}^n} C(\sigma) \\ s.t.\, Z_\sigma^\alpha \subset \mathcal{M} \end{cases}$$

$C : \sigma \mapsto C(\sigma)$ is a controlability function and Z_σ^α some transformation of our quantile of level α and the set $\mathcal{M}$ can be seen as a confident region.

In other words, on the space of the outputs some regions are critical (in a weaker version we only consider some critical points) and one wants to calibrate σ such that with high probability the outputs are far from the critical region.

(a) in red: boundary of critical region
in blue: Sample points



(b) in red: critical points
in blue: Sample points

We show that the quantiles introduced in the previous part are well-suited to handle with this type of applied problems.

References:

[1] D. Nolan, Asymptotics for multivariate trimming, Stochastic Processes and their Applications. *Stochastic Processes and their Applications, Volume 42, Issue 1, Pages 157-169*, August 1992.

[2] J. Tukey, Mathematics and the picturing of data. *Proc. 1975 Inter. Cong. Math., Vancouver 523531. Montreal: Canad. Math. Congress.*, 1975.

Short biography – The PhD is a follow-up to an internship completed on a landing gear problem in EADS IW in 2010/11. the PhD is founded at "CERFACS" by "EADS Innovation Works".

Modelling dependence under constraint

R. Fischer
Université Paris-Est

Affiliation: *Department of Industrial Risk Management*, EDF R&D, 1 Avenue du Général de Gaulle, 92140 Clamart, France

Email: richard.fischer@edf.fr – **URL**: http://www.cermics.enpc.fr/~fischerr/home.html

Master: Université Paris-Dauphine

Ph.D. (2013-2016): Université Paris-Est

Supervisor(s): Prof. Jean-François Delmas (ENPC) and Prof. Cristina Butucea (UPE-MLV) and Anne Dutfoy (EDF R&D)

Abstract:

In various cases of modelling the propagation of uncertainity in complex numerical simulation schemes, the engineer needs to create a probabilistic model that takes into consideration the available statistical information (or expert judgement). In most of the cases this means that the marginal distribution of each uncertain quantity is determined, but it can also be a physical constraint between these quantities, for example an order relationship. The construction of such a model is quintessential if we want to evaluate a probabilistic reliability criterion by Monte Carlo methods. In order to estimate the expectation of a random variable derived from the uncertain quantities, or the probability that it exceeds a certain threshold, we need to generate i.i.d. samples, which requires the complete knowledge of the joint distribution of these quantities.

My thesis focuses on the modelling of multivariate random vectors where the marginal distributions of the components are given and the components verify certain deterministic constraints. As a first problem we consider ordering constraits, that is the components of the random vector are ordered almost surely. Among the possible joint distributions (if there exists any), our aim is to find the most "random" possible. To measure the uncertanity of our joint model, we use the Shannon entropy defined for a d-dimensional random variable $X = (X_1, \dots, X_d)$ as:

$$H(X) = -\int f_X(x) \log\left(f_X(x)\right)\, dx,$$

where f_X denotes the density of X. We are looking for the joint distribution of X that maximizes this measure. From an information theory point of view, this is the distribution that adds the least amount of information in addition to the initial constraints, leading to a more realistic modelling of the random vector.

Since the marginal distributions are fixed, the theory of copula functions seems appropriate to model the dependence structure between the uncertain quantities. Copulas are multivariate distribution functions whose marginals are uniformly distributed on $[0, 1]$. Sklar's theorem states that the distribution function F of a d-dimensional random variable X with continuous marginal distribution functions F_i, $i = 1 \dots d$ can be written as:

$$F(x_1, \dots, x_d) = C\left(F_1(x_1), \dots, F_d(x_d)\right)$$

where C is a uniquely determined copula. This allows us to decompose the information on the marginals and the information on the dependence, which is incorporated in the copula. Furthermore, we have the following decomposition of the entropy of X:

$$H(X) = \sum_{i=1}^{d} H(X_i) + H(U),$$

where $H(X_i)$ is the entropy of the univariate random variable X_i, and $H(U)$ is the entropy of a d-dimensional random variable U whose distribution function is C, the copula associated to X. Since the

marginals X_i are given, we can observe that $H(X)$ is maximal if and only if $H(U)$ is maximal. Therefore we turned our attention to finding the copula that maximizes the entropy under the constraints imposed by the marginals and the ordering. We have found that in the case of $d = 2$, this copula can be expressed with the help of the copula which maximizes the entropy under the constraint that its diagonal section, defined by $\delta(t) = C(t, t)$, $t \in [0, 1]$, is given. The function δ only depends on the marginals distribution functions of X. We present the explicit density of the unique optimal copula with given diagonal section δ. We give an explicit criterion on the diagonal section for the existence of the optimal copula as well as the closed formula for its entropy. Using this copula we determine explicitly the distribution function of the random variable X which maximizes the entropy under the ordering and marginal constraints. For $d > 2$, we have shown that the maximum entropy copula of ordered random vectors with given marginals is related to the maximum entropy symmetric copula whose order statistics have fixed distributions. That is, if U is a random variable with distribution function C, then the distribution function of the i-th largest element of U is fixed and given by $\delta_{(i)}$ for all $1 \leq i \leq d$. As in the case of $d = 2$, we give the explicit density of the optimal copula with a closed formula for its entropy for a special class of marginal distributions. We also present the density of the random variable X which maximizes the entropy under the ordering and marginal constraints for general $d \geq 2$ for this special class of marginals.

The results will be illustrated on various choices for the marginal distributions, and an example arising from an industrial problem will also be presented.

Short biography – I'm a second-year Ph.D. student in Applied Probability from Hungary at Université Paris-Est, affiliated to the laboratories CERMICS of ENPC and LAMA of UPE-MLV. My work is a collaboration between the Department of Industrial Risk Management of EDF R&D and the aforementioned laboratories under a CIFRE contract. My studies are jointly funded by the organization ANRT and EDF R&D.

Multi-fidelity regression using a non-parametric relationship

F. Zertuche
Université Joseph Fourier, Grenoble

Affiliation: *Laboratoire Jean Kuntzmann*, UJF Grenoble, 51 rue des Mathmatiques, BP 53 38041 Saint Martin d'Hres, France

Email: zertuche.federico@imag.fr – **URL**: https://team.inria.fr/moise/en/

Master: Université Jean Monnet, Saint-Etienne

Ph.D. (2012-2014): Université Joseph Fourier, Grenoble

Supervisor(s): Prof. A. Antoniadis (UJF) and Dr. C. Helbert (ICJ)

Abstract:

We study the synthesis of data from different experiments. These experiments are very complex computer simulations that take several hours to produce a response for a given input. Understanding the phenomenon modeled by the simulation requires a large number of responses and in practice having all of them is unfeasible due to time constraints. This is why the computer simulation is often replaced by a simpler probabilistic model, also known as metamodel, that is faster to run.

The studied metamodel is based on the hypothesis that the computer simulation is in fact the realization of a gaussian process indexed by the inputs and defined by a parametric mean function and a parametric covariance function. A small number of responses produced by the computer code are used to determine the values of the parameters of the mean and covariance functions. Given a new input, the predicted value is the expectation of the stochastic process at that input conditioned by the responses available. Since the stochastic process is gaussian, there is a formula for this expectation and the error of prediction.

When the precision of the output produced by the computer code can be tuned it is possible to incorporate responses with different levels of fidelity to enhance the prediction of the most accurate simulation at a new input while respecting the time constraints. This is usually done by adding several imprecise responses instead of a few precise ones. The main example for this type of computer experiments are the numerical solutions of differential equations. The precision can depend on the size of the mesh of the domain of resolution used to produce the response; on the space where the solution is projected or even whether a part of the physical model involved is left aside. The problem is how to take into account all the information available. This problem has been studied by many authors, most notably by LeGratiet in [1] and by Kennedy and O'Hagan in [3].

In the present work, we propose a new approch that is different from the existing ones.

For ease of notation only two precision or fidelity levels are considered: 1 for the least accurate and 2 for the most precise. First we will assume that the most precise level is a function of the least accurate. The difference between the two will be modeled by the gaussian process $Z_{(2,x)}$. If we suppose that $Y_{(1,x)}$ is the gaussian process related to 1, then $Y_{(2,x)}$ defined by equation (1) is also a gaussian process. It will model the outcomes of 2.

$$Y_{(2,x)} = \varphi(Y_{(1,x)}) + Z_{(2,x)} \tag{1}$$

Generalizing the results in [1], we propose a non-parametric approach where we compute a locally linear approximation of the function φ. We estimate the relationship and build a predictor by using all the responses to compute the conditional expected values for $Y_{(1,x)}$ and $Z_{(2,x)}$. The prediction error is built using the predicting errors of $Y_{(1,x)}$ and $Z_{(2,x)}$.

Then, we study an analog model based in [2] where the difference between the two levels is no longer a gaussian process. This time the difference between the two computer simulations will be modeled by

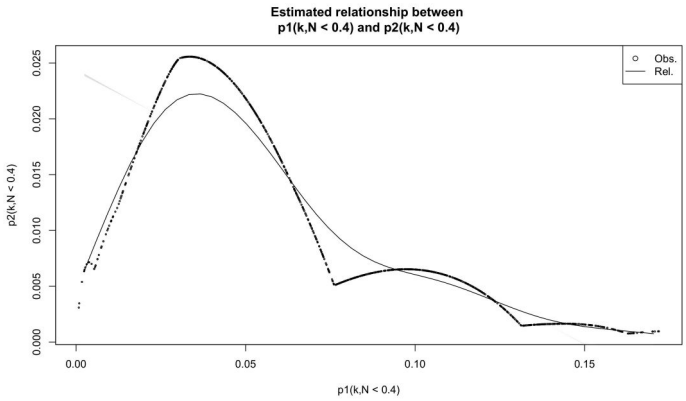


Figure 1: Estimated relationship between two successive levels of a computer code that simulates the pressure transient in a porous media.

the correlated errors ϵ_y. The correlation structure of the errors will depend on the distance between the outputs of 1. The new probabilistic model for the second simulator is given by equation (2) where $Y_{(1,x)}$ is still the gaussian process related to 1.

$$Y_{(2,x)} = \varphi(Y_{(1,x)}) + \epsilon_y \tag{2}$$

Once again we will estimate φ by using locally linear polynomials. Since we considered a particular correlation structure for the errors, we use the algorithm described by Fernandez in [2] to correct the bias in the estimation of the smoothing parameter of the non-parametric regression.

Finally, the two models are tested to illustrate their advantages and shortcomings. First by simulating the computer codes as gaussian processes we find that assuming that φ is linear when it is not can affect the results of the predictions. By using physical models we notice that the relationship between two fidelity levels of a computer code can be non-linear - as shown in Figure 1 - and in some cases not even function-like. Then, we develop briefly a case study related to a diphasic air-water flow in a rectangular domain.

References:

[1] Le Gratiet, L. Recursive co-kriging model for Design of Computer experiments with multiple levels of fidelity with an application to hydrodynamic, arXiv:1210.0686; 2012.

[2] Fernandez, F and Opsomer, J. Smoothing Parameter Selection Methods for Nonparametric Regression with Spatially Correlated Errors. Canad. J. Statist., 33(2): p.279-295; 2004.

[3] Kennedy, M and O'Hagan, A. Predicting the output from a complex computer code when fast approximations are available. Biometrika. 87(1): p.1-13; 2000.

Short biography – Federico Zertuche is a third year PHD student at the Laboratoire Jean Kuntzmann (Université Joseph Fourier, Grenoble) under the supervision of Céline Helbert and Anestis Antoniadis. He is part of the INRIA team MOISE whose main research theme is the development of mathematical methods for modeling environmental phenomena.

Uncertainty quantification and visualization for functional random variables

S. Nanty
Université Joseph Fourier, Grenoble, France

Affiliation: *CEA, DEN, DER*, F-13108 Saint Paul-lez-Durance, France

Email: simon.nanty@cea.fr – **URL**: http://www-ljk.imag.fr/

Master: Université Jean Monnet, Saint-Etienne, France

Ph.D. (2012-2015): Université Joseph Fourier, Grenoble, France

Supervisor(s): C. Prieur (Laboratoire Jean Kuntzman) and C. Helbert (Ecole Centrale Lyon) and A. Marrel (CEA) and N. Pérot (CEA)

Abstract:

In the framework of industrial risk assessment studies, the reliability of a component is evaluated during accidental conditions. A numerical model provides the thermal-hydraulic (T-H) loading of this component subjected to highly hypothetical conditions. Then, a numerical model for the mechanical analysis of components and structures (called hereafter the T-M code), taking as input the T-H loading, calculates the breaking strength of the component and the thermo-mechanical actual applied load. From these two elements, a safety criterion is deduced.

Under the hypothesis of accidental conditions, the component behavior depends on many uncertain parameters, related to the initial plant conditions or to the safety system characteristics. These uncertain parameters are input variables of the T-H code. They can be of various types: scalar, functional, categorical... It can be important to assess how these uncertainties can affect the code forecasts. To deal with all these uncertainties, computer experiment methodologies like uncertainty propagation and sensitivity analysis are useful. As the T-H code is much more time expensive than the T-M code, the uncertainties on the results of the T-H code which are of functional type, are directly characterized, so that uncertainty propagation and sensitivity analysis can be applied on the T-M code.

The uncertainty characterization of functional input variables has already been investigated by a few authors. A common way to study functional variables is to decompose them on a functional basis. Collin et al. [1] decompose the functional variable under study thanks to Functional Principal Component Analysis (FPCA), developed by Ramsay and Silverman [2]. As they consider that the functional random variable is a Gaussian process, the coefficients of the FPCA basis functions are independent and follow a centered and standardized Gaussian distribution. However, the Gaussian process hypothesis is quite restrictive in practice. Hyndman and Shang [3] propose a method close to the previous one. The functional variable is first decomposed on a FPCA basis. The joint probability density function of the coefficients on this basis is then estimated thanks to a kernel density estimator (Rosenblatt [4]). However, the kernel density estimation is inefficient in high dimension, so that the number of functions in the FPCA basis must be small to apply this method in practice.

In the present work, the problem under consideration is different from the previously studied ones in the sense that the functional random variables to be characterized are dependent upon one another and are linked to a scalar (or vectorial) variable, called hereafter a covariable. The considered covariable is here the output of the second code. This covariable can be, for instance, the output of a computer code which takes as inputs the functional random variables. The main objective of this work is thus to provide a new methodology to characterize the uncertainties associated to dependent functional variables linked to a covariable. These functional variables are discretized in practice. As in the two presented methodologies, the proposed characterization process is composed of two parts. First, the dimension of the problem is reduced by decomposing the functional random variables on a functional basis. In order to take into account the dependence between the functional random variables, the decomposition is done simultaneously on all

the studied functional random variables. This means that the decomposition is done on a vector of functional random variables instead of a functional random variable. We propose a simultaneous version of classical Partial Least Squares (Wold [5]), denoted SPLS. This SPLS method enables to take into account the link between the functional random variables and the covariable. The functional random variables are approximated by their coefficients on the SPLS basis. The problem becomes then multivariate instead of functional. The second step of the characterization procedure consists in estimating the joint probability density function of the basis coefficients. The distribution of these coefficients is modeled by a Gaussian mixture, whose parameters are estimated by the Expectation-Maximization algorithm (Dempster et al. [6]). Thanks to this modeling, the probability density function can be estimated in higher dimension than with the kernel density estimation, which is usually inefficient above dimension 6, so that a higher number of SPLS basis functions can be selected. This methodology gives a statistical characterization of the uncertainties on the studied functional random variables. New realizations of these random variables can also be simulated: coefficients are sampled from the estimated Gaussian mixture model, and the corresponding functions are then built by multiplying the new coefficients with the SPLS basis functions. The presented methodology has been tested and validated on a numerical example with two dependent functional variables and on a nuclear application with three dependent functional variables.

References:

[1] F. Collin, J. Goffart, T. Mara and L. Denis-Vidal. Uasa of complex models: Coping with dynamic and static inputs. sensitivity analysis of model outputs. In *7th International Conference on Sensitivity Analysis of Model Output*, 2013.

[2] J. O. Ramsay and B. W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2nd edition, 2005.

[3] R. J. Hyndman and H. L. Shang. Rainbow plots, bagplots, and boxplots for functional data. *Journal of Computational and Graphical Statistics*, 19(1):29–45, 2010.

[4] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *The Annals of Mathematical Statistics*, pages 832–837, 1956.

[5] H. Wold. Path models with latent variables: The nipals approach. *Quantitative sociology: international perspectives on mathematical and statistical modelling*, 1975.

[6] A. P. Dempster, N. M. Laird and D. B. Rubin. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, Series B (Methodological):1–38, 1977.

Short biography – After a master's degree in applied mathematics in Saint-Etienne, Simon Nanty began his PhD in October 2012 with Grenoble university and the comissariat à l'énergie atomique. This PhD thesis is funded by the CEA, and its objective is to quantify the uncertainties associated to a computer code with functional inputs and scalar output.

PhD Posters Session

List of Abstracts

Robust Design and Optimization of a Jet Engine Low Pressure Turbine Rotor

Giulia Antinori
Technische Universität München, Germany

Affiliation: *Chair of Computational Mechanics*, Technische Universität München, Arcistrasse 21, 80333 Münich, Germany

Email: giulia.antinori@tum.de – **URL**: http://www.cm.bgu.tum.de/index.php?id=59&L=1

Master: University of Padua, Italy

Ph.D. (2012-2015): Technische Universität München, Germany

Supervisor(s): Prof. Fabian Duddeck (TUM) and Dr. Andreas Fischersworring-Bunk (MTU Aero-Engines)

Abstract:

Aim of the PhD project is the evaluation of several stochastic methods with respect to their applicability for the uncertainty quantification of a coupled flow-thermo-mechanical model of a jet engine low pressure turbine rotor. Stochastic methods to quantify uncertainty are well established in literature, however their application for the multidisciplinary design of a low pressure turbine rotor is not common. In fact, the design of a new part is usually based on a deterministic model. The stochastic analysis is comprised of a sensitivity analysis followed by an uncertainty analysis. The sensitivity analysis is performed to gain a better understanding of the coupled flow-thermo-mechanical system robustness, to identify the important variables and to reduce the number of design parameters which will be used in the optimization. The uncertainty analysis using probability distributions derived from the manufacturing process, allows to predict the effect of the input uncertainties on the life duration of the rotor. The flow system models the Secondary Air System of the aircraft engine, which performs essential tasks most of them related to cooling or sealing applications. It is represented as a succession of chambers, with finite volume, linked by complex flow passages (holes, ducts, pipes, seals, discs) which possess their own pressure loss characteristics. The thermal system models the heat conduction, convection phenomena between surface and fluids, and the radiation phenomena which take place inside the low pressure turbine rotor. The temperature field is used to compute the stresses due to thermal and dynamic loads, which are then used to produce an estimate of life expectancy. The flow-thermo-mechanical system inputs, which primarily consist of geometric characteristics as well as of the performance parameters are varied in a certain range ($\pm x\%$ from nominal value) on the basis of the manufacturing tolerances. Using some knowledge on the system, physically meaningful dependencies among the inputs (among performance parameters as well as among some of the geometric features) has been introduced before the sensitivity analysis starts. In order to simulate the system the remaining independent input variables are sampled uniformly within their range of variation. A review of the most common sampling methods is performed. The study shows that some of the sampling methods cannot be recommended since they produce spurious correlations between independent input variables.

With regards to the sensitivity analysis, many literature sources state that the Pearson correlation method is only valid for linear models when assessing the importance of input variables. As the coupled flow-thermo-mechanical system might behave non-linearly and interactions among the different parameters should be also determined, non parametric variance based methods are introduced to make up for the limitations of the correlation method. Following the result of the study, it is recommended to combine the correlation computation with a non parametric variance based method.

Once the main players have been identified, uncertainty quantification analysis is performed. There are mainly two kinds of uncertainty: epistemic uncertainty, which is reducible to a lack of knowledge, and aleatory uncertainty, which is due to random variability inherent in nature. In this work we are interested in the quantification of aleatory uncertainty. This aim can be fulfilled by using sampling methods, i.e

by generating sets of samples according to the probability distribution of the uncertain variables and by mapping them into corresponding sets of response functions. The probability distribution of the input variables can be determined by using measurement data or by making some assumptions. Due to the lack of measurements, a normal probability distribution has been assigned to the input variables. The mean is given by the input parameters' nominal value and the standard deviation by the parameter variation, which for the geometric parameters corresponds to the manufacturing tolerances and it amounts to $\pm 3\sigma$. For what concerns the performance parameters, at the beginning of the analysis an estimate of the variation based on experience was given. However, it resulted to be too inaccurate and it led to erroneous results. Therefore the estimation has been refined and in the new definition it corresponds to a $\pm 2\sigma$ variation. Additionally, correlations (which were neglected in the first approximation) among the different performance parameters has been introduced by using the Iman and Conover modification of the Latin Hypercube sampling technique. Probably a more rigorous definition would be given by the copula theory, which will be objective of future investigations. The introduction of correlations among some of the input parameters requires a new investigation of the sensitivity analysis methods. This is due to the fact that both the EE and the Sobol methods assume the independence of the input parameters.

The objective of the uncertainty analysis is to check the robustness of the system which can be measured by assessing that the current solution differs from the nominal one, i.e. from the design intent solution, less than $\pm 3\sigma$. As first, the probability distribution of the output variables is identified through an Anderson-Darling test. Almost all the responses results to be normally distributed, hence a t-test is performed to check if the sampled mean is equal to the nominal solution. The confidence level chosen for the test is 95%. The result of the test, is that the nominal value lies between the first and third quartile, i.e. inside a $\pm 1\sigma$ variation, which proves the robustness of the system.

The next step is the optimization. Life expectancy is one of the objective functions to be optimized. In addition, the amount of cooling air employed by the SAS should be minimized as well as the temperature gradient between bore and rim locations in every turbine stage. Additionally, the parameters which represent the system topology are affected by uncertainties due to manufacturing tolerances, engine-to-engine variations and ambient conditions. Thus, in order to include the uncertainties, a probability distribution is assigned to each parameter and the optimization is performed in the frame of robustness, i.e. it should be assured that the optimized solution remains relatively unchanged when exposed to uncertain conditions. Therefore, the optimization problem is not only multidisciplinary, but also multi-objective, since both the mean and variance of the objective function have to be minimized. Since the simulation of the physical system is computationally quite demanding, Monte Carlo methods for the computations of robustness measures are not affordable. A faster method, based on sparse polynomial chaos expansions, is implemented instead to guarantee higher efficiency.

Short biography – I studied Mathematics at the University of Padua gaining my Bachelor degree in 2009 and my Master degree in 2012. I have been always interested in the application of mathematics to real life problems. I am currently PhD student at the Technische Universität in Munich and my research project is the robust design and optimization of a low pressure turbine rotor. The project is part of the LuFo aeronautics research programme funded by the German Federal Ministry of Economics and Technology and it is supported by the MTU-Aero Engines and the Technische Universität München.

Kriging for non-parametric ML estimation from region-censored data

Y. Bennani
Université de Nice Sophia Antipolis

Affiliation: *Laboratoire I3S*

Email: "bennani@i3s.unice.frBennani"<bennani@i3s.unice.fr> – **URL**: http://www.i3s.unice.fr/I3S/labos/labo4.html

Master: M2 Modélisation Aléatoire, Université Paris 7

Ph.D. (2012-2015): Université de Nice Sophia Antipolis

Supervisor(s): Dr. Luc Pronzato (UNS-CNRS) and Dr. Maria-João Rendas (UNS-CNRS)

Abstract:

We consider a non-parametric density estimation problem with region-censored observations. The study is motivated by prevention of decompression sickness accidents through prediction of the amount of nitrogen bubbles produced during deep-sea diving. It relies on measurements of bubble grades – which reflect the peak gas volume in the diver's body – on a set of dives made by individuals in the population under analysis.

It is assumed that the instantaneous volume of gas $B(\theta, P(\cdot), t)$ flowing through the right-ventricule of a diver characterised by a set of biophysical parameters θ when executing dive profile $P(\cdot)$ (a function of time) is well described by a known mathematical model [1]. Bubble grades, $G \in \{0, \ldots, 4\}$, are a strongly quantified version of peak gas volume (see Fig. 1):

$$G(\theta, P(.)) = i \Leftrightarrow \tau_i \leq \max_t (B(\theta, P(\cdot), t)\} < \tau_{i+1},$$

where $\tau_0 = 0 < \tau_1 < \cdots < \tau_4 < \tau_5 = \infty$ is a set of thresholds assumed known.

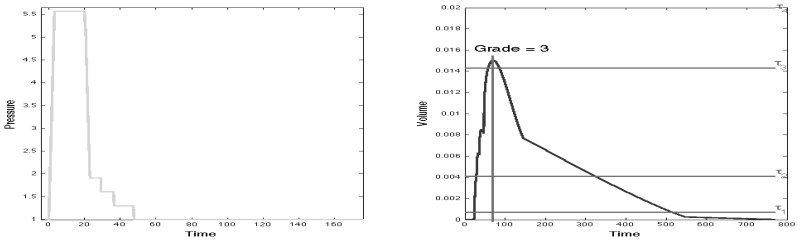


Figure 1: Left: dive profile $P(\cdot)$. Right: model response (blue) and grade computation (threshold are shown in red). Observed grade is equal to 3 in this case.

We address determination of $\hat{\pi}_\theta, \theta \in \Theta$, the non-parametric Maximum Likelihood estimate (NPMLE) of the distribution of θ in the population under study. Observation of grade $G = i$ when executing $P(\cdot)$ only indicates that $\theta \in \mathcal{R}_i(P(\cdot)) = \{\theta \in \Theta \ : \ \tau_i \leq \max_t (B(\theta, P(\cdot), t)\} < \tau_{i+1}\}$ and thus we face a problem of density estimation from (region-)censored observations. For interval-censored observations it is known [3, 4] that $\hat{\pi}_\theta$ is affected by several forms of indeterminacy, the major being that only the probability mass over the (finitely many) elements of a partition $\mathcal{P}$ of the parameter space, determined by the set $\mathcal{R}$ of observed regions $\mathcal{R}_i(P(\cdot))$, can be estimated. Moreover, $\hat{\pi}_\theta$ is concentrated on a subset of the elements of $\mathcal{P}$, that can be found from the intersection graph of $\mathcal{R}$, see Fig. 2. These features carry over to censoring by regions of arbitrary geometry, as in our case, requiring only a slightly more complex determination of the support of $\hat{\pi}_\theta$.

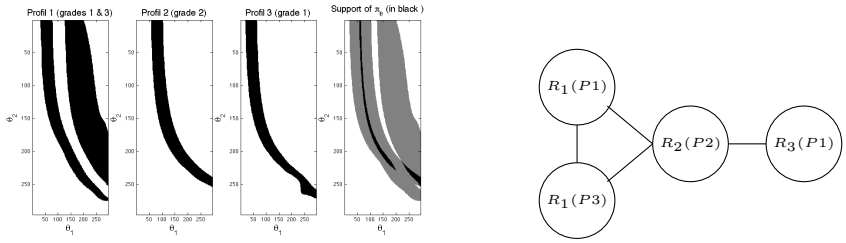


Figure 2: Definition of $\mathcal{P}$ for the observation of grades 1 and 3 for profile P_1, grade 2 for P_2 and grade 1 for P_3 ($\theta \in \Theta \subset R^2$). Left: observed regions and partition $\mathcal{P}$ (the support of the $\hat{\pi}_\theta$ is indicated in black). Right: intersection graph of $\mathcal{R}$ (support of $\hat{\pi}_\theta$ is determined by its cliques).

Central to the determination of the NPMLE is the identification of the regions $\mathcal{R}_i(P(\cdot))$, that must resort to numerical methods, requiring computation of the model response to $P(\cdot)$ over a dense grid covering Θ. In our case, we have 444 measures of grades along 48 different decompression profiles, rendering impractical direct use of the biophysical model. To overcome this problem, we rely on a set of kriged observation models, that predict the value of $\max_t B(\theta, P(\cdot), t)$, from the model response over a sparse (11×11) grid. The response surface was estimated by simple kriging for an isotropic Mattern kernel using the package STK [2].

We present the NPML estimate of the probability mass of π_θ over the elements of $\mathcal{P}$, which is based on a fast multiplicative algorithm [5]. We illustrate the pathological behaviour of this estimator, in particular its sensitivity to the detailed geometry of $\mathcal{P}$, and propose alternative (regularised) solutions that account for the entropy of the estimated distribution.

Our ultimate goal is to predict the distribution of grades for an arbitrary profile $P(\cdot)$: $\hat{p}_P(i) = \hat{\pi}_\theta(\mathcal{R}_i(P(\cdot))) = \sum_{A \in \mathcal{P}} \hat{\pi}_\theta(A \cap \mathcal{R}_i(P(\cdot)))$. These estimates are affected by two distinct uncertainties: *(a)* we do not know $\hat{\pi}_\theta(A \cap \mathcal{R}_i(P(\cdot)))$ since $\mathcal{R}_i(P(\cdot)) \notin \mathcal{P}$ for new profiles; *(b)* the identification of $\mathcal{R}_i(P(\cdot))$ relies on kriging and is thus uncertain. We present upper and lower bounds on each $\hat{p}_P(i)$ that take into account *(a)*. Assessment of uncertainty source *(b)* concerns the determination of level sets based on kriging, as approached e.g. in [6] using the notions of Vorob'ev expectation and deviation and will be considered in the near future.

References:

[1] Julien Hugon. *Vers une modélisation biophysique de la décompression*. PhD Dissertation, University of Aix-Marseille 2, 2010.

[2] Small Tool for Kriging (STK), `http://sourceforge.net/projects/kriging/`

[3] Bruce W. Turnbull. *The empirical distribution function with arbitrarily grouped, censored and truncated data*. Journal of the Royal Statistical Society. Series B (Methodological), 1976, 290–295.

[4] Robert Gentleman and Alain C. Vandal. *Nonparametric estimation of the bivariate CDF for arbitrarily censored data*. Canadian Journal of Statistics, Vol.30, No. 4, 2002, 557–571.

[5] Radoslav Harman and Luc Pronzato. *Improvements on removing nonoptimal support points in D-optimum design algorithms*. Statistics & probability letters, Vol. 77, No. 1, 2007, 90–94.

[6] Clément Chevalier, David Ginsbourger, Julien Bect, Ilya Molchanov *Estimating and quantifying uncertainties on level sets using the Vorob'ev expectation and deviance with Gaussian process models*. In proceeding of: mODa 10, Advances in Model-Oriented Design and Analysis, Physica-Verlag HD, 35-43.

Short biography – After obtaining an engineer diploma (2008) and the master degree (2009), Y. Bennani worked in banking before starting a PhD thesis at Laboratory I3S (2012) on the estimation of the risk of decompression accidents among deep-sea divers. His thesis is conducted in the framework of contract DGA-DGCIS SAFE DIVE, a joint partnership between the company BF-Systèmes, Institut Langevin (ESPCI Paristech), and the laboratory I3S (UNS-CNRS).

Sequential sensitivity analysis for computer experiments with functional inputs

J. Fruth
TU Dortmund University

Affiliation: *Faculty of Statistics*, TU Dortmund University, Vogelpothsweg 87, 44227 Dortmund, Germany

Email: fruth@statistik.tu-dortmund.de – **URL**: http://www.statistik.tu-dortmund.de/1553.html

Master: TU Dortmund University

Ph.D. (2012-2015): TU Dortmund University

Supervisor(s): Prof. S. Kuhnt (Dortmund University of Applied Sciences and Arts) and Prof. O. Roustant (Ecole des Mines, Saint Etienne)

Abstract:

Sensitivity analysis [1] studies the influence of input factors of a computer experiment on a given output. A computer experiment usually simulates a real-life process which runs for a certain time. Input values are usually fixed at the beginning of the simulation as single input numbers and are not altered any more during the process. However, in various contexts, especially in mechanical processes, input values could indeed be varied during the process. Examples might be speed (e.g. of a car or a robot arm), the dose of a physical agent or temperature. Varying such factors during the process and analysing the functional sensitivity can strongly improve the understanding of the process and can lead to much more precise settings.

Our research is motivated by a sheet metal forming process [2], where two parameters, which were previously kept constant, can now be varied in time: the blankholder force, which keeps the sheet metal in place during the forming, and the friction between the tool and the sheet metal. See Fig. 1 (left) for a representation of a forming press. As output value the forming accuracy is measured in terms of springback, the amount of deformation of the flange after the forming. Two example runs can be seen in the table in Fig. 1 on the right. The first column shows the functional path of the parameter friction, the second the resulting springback value. The blankholder force is kept constant. For both runs, the friction mean over time is equal to 12.5, but the resulting springback values differ considerably. This emphasizes the possibilities that lie in a functional exploration.

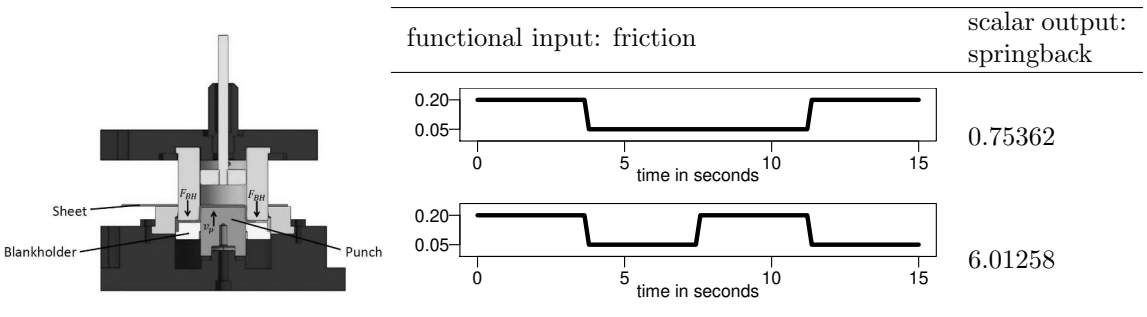


Figure 1: Representation of a sheet metal forming press (left). Two example runs with varying friction and constant blankholder force (right).

In the poster a very economical sequential approach for the sensitivity analysis of time-varying factors in computer experiments is introduced and developed. It does not only return a single influence value

per input, but instead gives the influence of each input over the whole time scale. This allows for a clear graphical representation of the functional influence. The idea is to reduce the dimension of the functional analysis problem [3] by exploring whole intervals of time and sequentially decreasing the interval size for interesting time spaces. A sequential design approach making use of ideas from group factor screening [4] is developed, resulting in very few runs necessary to explore the interesting time regions of the functional inputs. In addition the analysis and visualization of interactions between time regions is adressed.

In the sheet metal forming application five sequential steps have been performed taking only 40 runs in total. The size of the bars in Fig. 2 represents the linear influence of the time interval. A stronger influence of friction compared to blankholder force is noticeable as well as, for both inputs a clear difference between the first and second half of the punch time: in the first half the influence is mainly positive whereas in the second half it is strongly negative with the last time points showing the most influence. The results improved the engineers understanding of the time dependent influence. They have been validated afterwards in real forming experiments and also led to a better springback reduction.

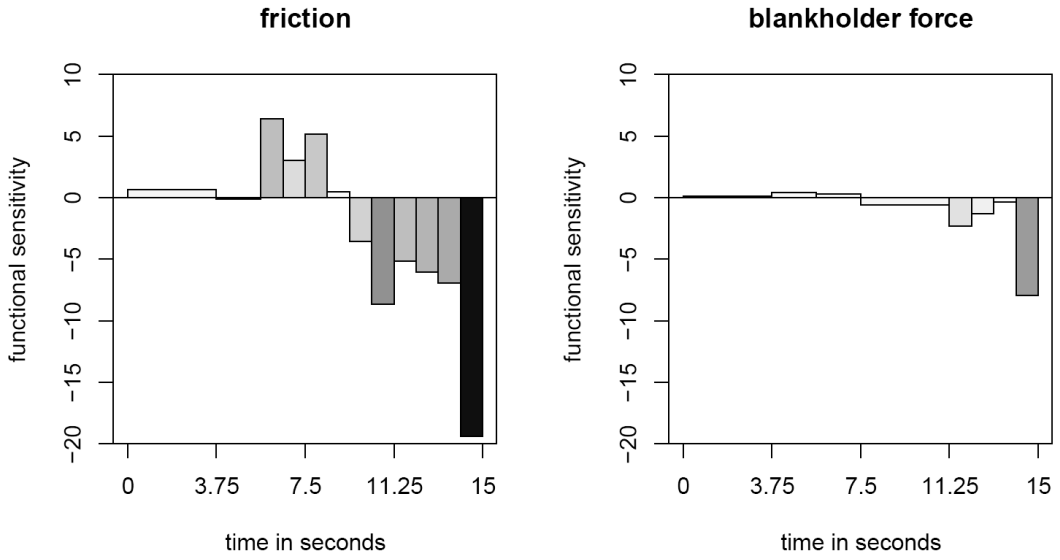


Figure 2: Graphical representation of the functional sensitivities of the input parameters friction and blankholder force in a sheet metal forming process.

References:

[1] Saltelli, A., Chan, K., Scott, E. M. (2000): Sensitivity analysis. Chichester: Wiley.

[2] Hassan, H., Fruth, J., Guener, A., Tekkaya, A. E. (2013): Finite Element Simulations for Sheet Metal Forming Process with Functional Input for the Minimization of Springback. In: IDDRG conference 2013. Zurich, June 2-5, 393-398.

[3] Ramsay, J. O., Silverman, B. W. (2005): Functional data analysis. 2nd ed. New York: Springer.

[4] Morris, M. D. (2006): An overview of group factor screening. In: A. Dean und S. Lewis (eds.): Screening methods for experimentation in industry, drug discovery, and genetics. New York: Springer, 191-206.

Short biography – Jana Fruth did her Master of Statistics at TU Dortmund University in Dortmund, Germany. In her PhD thesis she works on sensitivity analysis methods for various situations in sheet metal forming, including interaction analysis, analysis of functional inputs, and derivative-based indices. The project is a cooperation with the TU Dortmund Engineering Department as part of the collaborative research centre SFB 708, funded by the Deutsche Forschungsgemeinschaft (DFG).

Sensitivity analysis of tire model micro-coefficients

Sabra Hamza
Université de Haute-Alsace, Mulhouse, France

Affiliation: *Modélisation Intelligence Processus Systèmes (MIPS) laboratory*, Université de Haute-Alsace, 12 Rue des Frères Lumière 68093, Mulhouse, France

Email: sabra.hamza@uha.fr – **URL**: –

Master: -

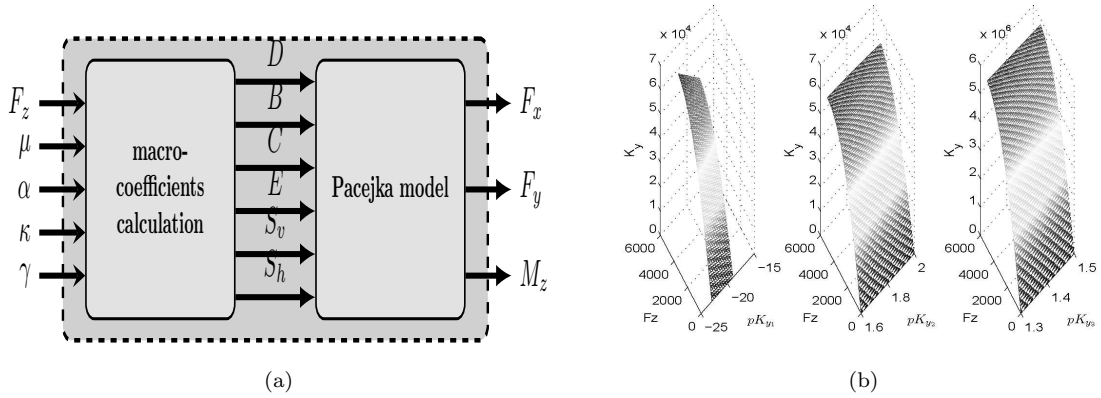
Ph.D. (2011-2014): Université de Haute-Alsace, Mulhouse, France

Supervisor(s): Abderazik BIROUCHE, Michel BASSET (Université de Haute-Alsace, Mulhouse, France) and Floriane ANSTETT-COLLIN (Université de Lorraine, France)

Abstract:

Problem statement

Tires involve the vehicles' most important safety features. Indeed, tires are required to produce the forces necessary to control the vehicle. Various models have been proposed to describe the behavior of the tire on the ground. These models depend of numerous parameters. One semi-empirical model commonly used in vehicle dynamics simulations, was developed by Pacejka [1]. It is widely used to calculate steady-state tire force and moment characteristics. This model depends on various parameters. An overview of Pacejka tire model is given in Fig. 1(a) :



(a) (b)

Figure 1: (a) An overview of the structure of Pacejka tire model (b) Graphical representation of lateral stiffness Ky as a function of vertical load F_z and micro-coefficients pKy_1, pKy_2 and pKy_3

Input variables are the vertical load F_z, the friction factor μ, the slip angle α and the camber angle γ. The coefficients B, C, D, E, S_h and S_v are the macro-coefficients which depend of the set of parameters called micro-coefficients. The outputs of the model are the longitudinal tire force F_x, the lateral tire force F_y and the tire self-aligning moment M_z. In [2], it has been shown that the lateral stiffness K_y and the slip angle α are the parameters affecting the lateral force variation. However, the lateral stiffness K_y depends on numerous parameters. In Fig. 1(b) the lateral stiffness Ky is illustrated as a function of the vertical load F_z and the micro-parameters pKy_1, pKy_2 and pKy_3 in their entire range of variation. Through Fig. 1(b), one can observe that the lateral stiffness increases when the vertical load increases. The impact of pKy_1, pKy_2 and pKy_3 on the lateral stiffness cannot be clearly distinguished. Thus, this work is an extension study of parameters influence of Pacejka tire model [1,2]. The aim is to quantify the influence of micro-parameters pKy_1, pKy_2 and pKy_3 on the lateral stiffness Ky and, therefore, on F_y.

Method and result

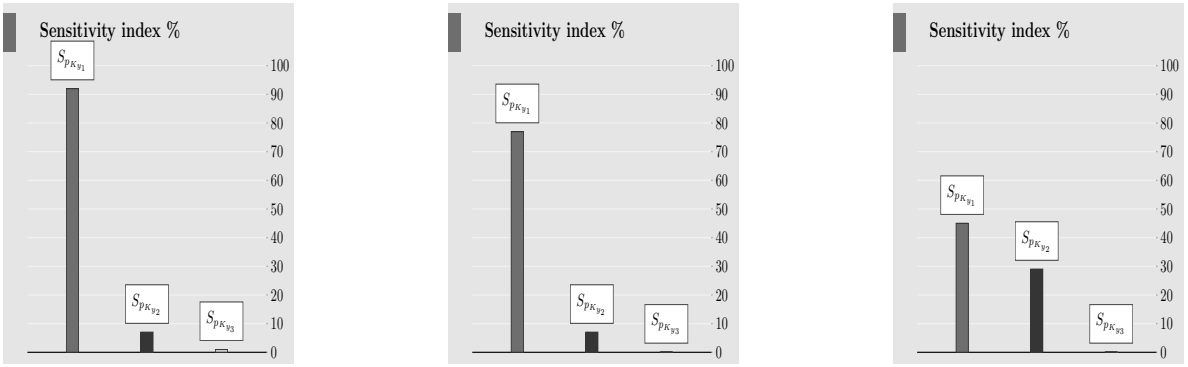
Polynomial chaos approach (PC) as a global sensitivity analysis method is applied. This method consists of approximate Ky into a sum of PC as follows [4]:

$$Ky \approx \sum_{j=0}^{\infty} c_j \psi_j(pKy_1, pKy_2, pKy_3) \tag{1}$$

with c_j the unknown deterministic coefficients and ψ_j the multi-variate orthonormale polynomial basis for (pKy_1, pKy_2, pKy_3) including all cross-terms between different parameters.
Since the lateral stiffness Ky depends on the vertical load F_z, this study has been made during different situations and for a small value of slip angle α. Depending on the value of F_z, three cases are considered:

- $F_z = F_{z_0}$: corresponding to situation without acceleration or braking.
- $F_z >> F_{z_0}$: corresponding to braking situation for tires of the front axle or acceleration for tires of the rear axle.
- $F_z << F_{z_0}$: corresponding to acceleration situation for tires of the front axle or braking for tires of the rear axle.

Sensitivity index of micro-parameters for different values of vertical load F_z are given in Fig. 2.



(a) $F_z >> F_{z_0}$ (b) $F_z = F_{z_0}$ (c) $F_z << F_{z_0}$

Figure 2: Sensitivity index

This result highlights the contribution of parameter pKy_1 on the lateral stiffness Ky variation for hight values of vertical load $F_z >> F_{z_0}$ and a negligible influence of other parameters. However, the parameter pKy_2 become influent when $F_z << F_{z_0}$.

References:

[1] H. B. Pacejka, Tyre and Vehicle Dynamics. Elsevier, 2006.

[2] S. Hamza, A. Birouche, F. Anstett-Collin and M. Basset, Sensitivity analysis for the study of a tire model with correlated parameters and an arbitrary distribution. *International Conference on Sensitivity Analysis of Model Output*, Nice, France, 1-4 July, 2013.

[3] R. Kiebre, F. Anstett-Collin, and M. Basset, Sensitivity analysis for tire/road interface model. *International Conference on Sensitivity Analysis of Model Output*, Milan, Italy, 19-22 July, 2010.

[4] G. Blatman, B. Sudret, Sparse polynomial chaos expansion and adaptive stochastic finite element using a regression approach, Comptes rendus de Mécanique 336 (6) (2008) 518-523.

Short biography – The objective of this thesis is to develop approaches for global sensitivity analysis for dynamic models. These approaches will be applied in the automobile domain.

Use of Chaos Polynomials in a Universal Kriging Model: Application to the Numerical Dosimetry

P. Kersaudi
Orange Labs, Whist Lab and Université Paris-Est

Affiliation: *Orange Labs*, 38 rue du Général Leclerc, 92130 Issy-les-Moulineaux, France

Email: pierric.kersaudy@orange.com – **URL**: http://www.univ-paris-est.fr/fr/

Master: Ecole Centrale de Nantes

Ph.D. (2012-2015): Orange Labs, Whist Lab and Université Paris-Est

Supervisor(s): Dr J. WIART (Orange Labs, Whist Lab) and Pr. O. PICON (Université Paris-Est)

Abstract:

If recent advances in terms of IT resources led to strongly reduce computational time for numerical dosimetry, the calculation of the Specific Absorption Rate that assess the human exposure to electromagnetic fields (EMF) remains very costly (a few hours per calculation). Consequently usual methods as the Monte Carlo cannot be used to study the influence of random input parameters variability on the SAR (the output variable). Then, optimal metamodeling strategies have to be employed to model the output response depending on the input parameters.

Universal Kriging and Polynomial Chaos are strategies that both aim at modeling the relation between an output process Y an some random input parameters $X = (x_1, ..., x_M) \in \mathbb{R}^M$.

On one side, Universal Kriging theory [1] models the relation between the output and the input parameters by the following expression:

$$Y(X) = \sum_{k=0}^{P-1} \beta_k \psi_k(X) + Z(X) \tag{1}$$

where $\psi = \{\psi_k, k = 0...P-1\}$ is a collection of regression functions, $\beta = \{\beta_k, k = 0...P-1\}$ are the regression coefficients and Z is a Gaussian process depending on X. In practice, the regression functions are chosen with respect to an a priori knowledge about the evolution of the output process. Most of the time, there is no such obvious a priori knowledge and the regression function ensemble is reduced to the unit function. The resulting meta model in generally called Ordinary Kriging model.

On the other side, Polynomial Chaos uses the Wiener polynomial expansion [2] to model the relation between the output variable and the input parameters:

$$Y = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^M} \beta_{\boldsymbol{\alpha}}(X) \psi_{\boldsymbol{\alpha}}(X) \tag{2}$$

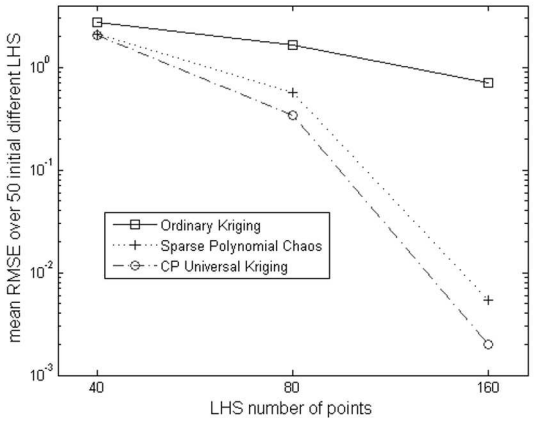
Where $\boldsymbol{\alpha} = [\alpha_1 ... \alpha_M]$ is the multi-index, the $\beta_{\boldsymbol{\alpha}}$ are deterministic coefficients to compute and the $\psi_{\boldsymbol{\alpha}}$ are multivariate orthonormal polynomials with respect to the probability measure associated with the random input parameters X.

Considering a sparse representation of the Polynomial Chaos expansion using Least Angle Regression technique [3] to select the most influencing polynomials, we propose to use these polynomials as regression functions in the Universal Kriging model. This Universal Kriging enriched with chaos polynomials will be called PC Universal Kriging.

The leave-one-out cross validation is used to assess respective accuracies of sparse Polynomial Chaos, Ordinary Kriging and PC Universal Kriging techniques.

This approach is illustrated with several benchmark functions used in the literature and with a dosimetry example aiming at assessing the fetus exposure to electromagnetic fields. The optimal nature of the PC Universal Kriging approach is illustrated.

As example, for the well-known Ishigami function [4], Figure 1 presents the mean of each metamodeling technique Root Mean Square Errors (RMSE) over 50 different initial LHS designs of experiments that are iteratively increased by the NLHS technique [5]. RMSE is computed using 100000 Monte Carlo points of the generated metamodels. Figure 1 shows that PC Universal Kriging brings about better accuracy in average than the two other techniques for the different numbers of points in the LHS design: slightly better than sparse Polynomial Chaos and much better than Ordinary Kriging for 80 and 160 points.



,

Figure 1: Accuracy comparison between Spase PC, Ordinary Kriging and PC Universal Kriging depending on the number of points in the LHS for the Ishigami function

References:

[1] G. Matheron, “Le krigeage universel," *Cahiers du centre de morphologie mathématique* 1, 1969.

[2] R. Ghanem and P. Spanos, *Stochastic Finite Elements: A Spectral Approach*, Courier Dover Publications, 2003.

[3] G. Blatman and B. Sudret, “Adaptative sparse Polynomial Chaos expansion based on least angle regression," *Journal of Computational Physics* 230, no. 6, pp. 2345–2367, 2011.

[4] A. Saltelli, K. Chan and E. M. Scott, *Sensitivity analysis*, Wiley, New-York, 2000.

[5] G. Blatman and B. Sudret, “An adaptative algorithm to build up sparse polynomial chaos expansions for stochastic finite element analysis," *Probabilistic Engineering Mechanics* 25, no. 2, pp. 183–197, 2010.

Short biography – From a master in Signal Processing, I am currently involved in the PhD thesis: "Statistical Analysis of People Exposure via the Numerical Dosimetry and the Design of Experiments in Orange Labs and Université Paris-Est". The main purpose of this thesis is to statistically characterize the exposure induced by wireless communication systems.

Gaussian processes for computer experiments with monotonicity information

H. Maatouk
Ecole Nationale Supérieure des Mines, Saint-Etienne, France

Affiliation: *Ecole Nationale Supérieure des Mines*, 158 cours Fauriel, 42 023 St-Etienne cedex 2, France

Email: hassan.maatouk@mines-stetienne.fr – **URL**: http://www.mines-stetienne.fr/en

Master: Université Lille 1, Sciences et Technologies, France

Ph.D. (2012-2015): Ecole Nationale Supérieure des Mines, Saint-Etienne, France

Supervisor(s): Xavier Bay (Ecole Nationale Supérieure des Mines, Saint-Etienne, France), Laurence Grammont (Université Jean Monnet, Saint-Etienne, Institut Camille Jordan, Lyon 1), Yann Richet (Institut de Radioprotection et de Sûreté Nuclaire (IRSN), Paris) and Olivier Roustant (Ecole Nationale Supérieure des Mines, Saint-Etienne, France)

Abstract:

Recently, statistical researchers have shown increased interest in Gaussian process modeling with monotonicity constraints (see [1], [2] and [3]). In computer experiments, the true function (scalar output) may be known to be monotone with respect to some or all input variables. We propose a new methodology based on the Bayesian Gaussian process metamodeling to sample from posterior distribution including monotonicity information in the monovariate case.

Let $y = f(x)$ be a monotonic increasing function where the input x is assumed to be scalar and in the domain $[0, 1]$. We consider a set of computer experiments $\{(x_i, y_i) \mid i = 1, \cdots, n\}$ of size n and assume that

$$y_i = f(x_i),\ 1 \leq i \leq n. \tag{1}$$

Also suppose that $\mathcal{M}$ is the space of increasing functions and $(Y_x)_{x\in[0,1]}$ is a zero-mean Gaussian process (GP) with kernel $k(x, x')$ given by a priori knowledge about the relationship between the input x and the output y. The following experimental results (see figures below) are obtained with the classical Gaussian kernel.

We are interested in the simulation of the conditional (or posterior) distribution of the GP Y given data and monotonicity information

$$\begin{array}{l} Y_{x_i} = y_i,\ 1 \leq i \leq n, \\ Y \in \mathcal{M}. \end{array} \tag{2}$$

The important step is to approximate the GP by a finite-dimensional GP Y^N

$$Y_x^N = \sum_{j=1}^{N} \xi_j \phi_j(x), \tag{3}$$

in which $\xi = \begin{pmatrix} \xi_1 \\ \vdots \\ \xi_N \end{pmatrix}$ is a zero-mean Gaussian random vector with covariance matrix Γ_N.

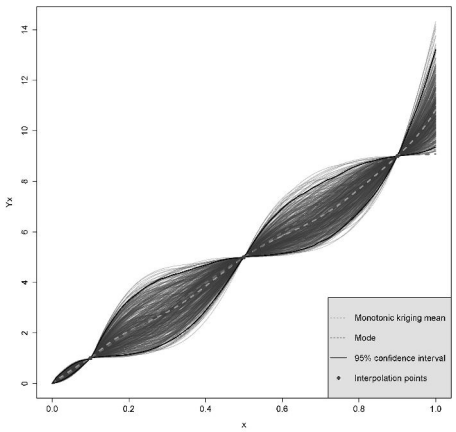
Such a decomposition can be seen as a Karhunen-Loève approximation of the process Y where the deterministic basis functions ϕ_j $(1 \leq j \leq N)$ are chosen in the space $\mathcal{M}$ of increasing functions and where random coefficients ξ_j partially reflect the randomness of the Gaussian process Y.

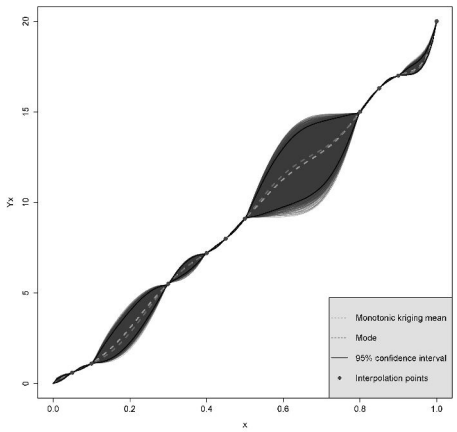
Due to the special choice of the basis functions ϕ_j, the crucial property here is that Y^N should be a monotonic increasing function **if and only if** the N coefficients ξ_j are all nonnegative. Now, we are

mainly interested in the new formulation of the problem : simulate the conditional distribution of the random Gaussian vector ξ given

$$\begin{array}{ll} \sum_{j=1}^{N} \xi_j \phi_j(x_i) = y_i,\ 1 \leq i \leq n & \text{(n interpolation linear equations)} \\ \xi_j \geq 0,\ 1 \leq j \leq N & \text{(N inequality conditions)} \end{array} \tag{4}$$

The advantage of such a methodology is that any posterior sample of the vector ξ leads to a monotone interpolating function. By a Monte Carlo technique, the conditional mean value of the function could be computed and be thought as the conditional monotone kriging mean. The conditional monotone kriging variance and confidence bounds can be calculated as well.





References:

[1] Marrel Amandine Da Veiga Sbastien. Gaussian process modeling with inequality constraints. *Annales de la facult des sciences de Toulouse Mathmatiques*, 21(3):529–555, 4 2012.

[2] Shirin Golchi, Derek R. Bingham, Hugh Chipman, and David A. Campbell. Monotone function estimation for computer experiments, September 2013.

[3] Jaakko Riihimaki and Aki Vehtari. Gaussian processes with monotonicity information. In Yee Whye Teh and D. Mike Titterington, editors, *AISTATS*, volume 9 of *JMLR Proceedings*, pages 645–652. JMLR.org, 2010.

[4] Luc Devroye. *Non-Uniform Random Variate Generation*. Springer-Verlag, 1986.

[5] J.A. Breslaw. Random sampling from a truncated multivariate normal distribution. *Applied Mathematics Letters*, 7(1):1 – 6, 1994.

[6] ChristianP. Robert. Simulation of truncated normal variables. *Statistics and Computing*, 5(2), 1995.

[7] John von Neumann. Various Techniques Used in Connection with Random Digits. *J. Res. Nat. Bur. Stand.*, 12:36–38, 1951.

Estimation of rare events probability and quantiles under monotonicity constraints

V. Moutoussamy
Université Toulouse Paul Sabatier

Affiliation: *EDF R&D - MRI*, 6 quai Watier, 78400 Chatou, France

Email: vincent.moutoussamy@edf.fr – **URL**: http://www.univ-tlse3.fr/

Master: Université Paris Descartes

Ph.D. (2012-2015): Université Toulouse Paul Sabatier

Supervisor(s): Prof. F. Gamboa (Université Toulouse Paul Sabatier) and Prof. T. Klein (Université Toulouse Paul Sabatier) and Dr. N. Bousquet (EDF)

Abstract:

A common task in structural reliability studies is to use a numerical computer code which simulates the physical behavior of a component. That computer code depends of inputs which represent some physical situation, an event is said undesirable if a function of an output is higher than a fixed threshold. The input parameters of the computer code can be uncertain, thus they are represented by random variables, so then the output become random. We are interested in the estimation of the probability that the computer code exceed the threshold. Here, the code is considered as black-box, deterministic (a set of input gives always the same output), time consuming and can be discontinous, so classical method to estimate probabilities are not adapted.

The knowledge of the component gives structural informations, here assumes that numerical code is monotonic. That allows us obtain deterministic bounds of the probability and delimits a set of input where it is not necessary to make new calls of the numerical code. To accelerate the convergence of these deterministic bounds, theoretical results shows that standard Monte Carlo method is not efficient. One propose here a general strategy to construct optimal design of experiment based on a sequential sampling, and construct a statistical estimator associated to the bounds.

As in the estimation of probability, the monotonic hypothesis provide information for the estimation of quantile. One propose a geometrical method to get deterministic bounds for quantile. Knowing a probability, that method is based on the construction of two sets which the volume is respectively lower and greater than the known probability. These bounds can be obtain from a given design of experiments or can be construct by a sequantial sampling.

We will compare classical methods dedicated to estimation of probability without monotonic hypothesis and their adaptation with monotonic assumption.

Short biography – During the previous year of the beginning of my thesis, I have worked on the development of that new class of method based on monotonic hypothesis for my final year internship. That thesis is a partnership between EDF and the University Toulouse Paul Sabatier and takes place on structural reliability of some passive components of an energy production unit.

Dynamical low rank approximation of time dependent PDEs with random data

E. Musharbash, F. Nobile, T. Zhou
École Polytechnique Fédérale de Lausanne

Affiliation: *MATHICSE*, École Polytechnique Fédérale de Lausanne, Station 8, CH-1015 Lausanne, Switzerland

Email: eleonora.musharbash@epfl.ch – **URL**: http://csqi.epfl.ch

Master: Politecnico di Milano, Milano, Italy

Ph.D. (2013-2016): École Polytechnique Fédérale de Lausanne

Supervisor(s): Prof. F. Nobile, CSQI, MATHICSE, École Polytechnique Fédérale de Lausanne

Abstract:

Mathematical models, often based on partial differential equations, are widely used to describe and predict the behaviour of physical and engineering systems. However, in many situations, some parameters entering the model can not be exactly quantified. This uncertainty on the data may reflect, on the one side, our ignorance or inability to properly characterize all input parameters of the mathematical model; on the other side, it may describe an intrinsic variability of the physical system. A convenient framework to include uncertainty in the mathematical model is offered by probability theory, where all uncertain input parameters are treated as random variables or random fields. The aim in this case, is to quantify the effects of the uncertainty on the predicted quantities of interest relevant for the applications at hand. We focus on possibly non-linear time dependent partial differential equations of diffusion reaction type with stochastic parameters where the randomness can appear in the initial data, in the coefficients of the differential operator, in the forcing term, etc.

In the last decades, differential equations with random data received a lot of attention in the field of scientific computing and numerical analysis. Since the number of the stochastic variables that are involved in a single phenomenon is often of the order of tens/hundreds or even more, the numerical approximation of these differential problems remains a challenging task. A great effort has been devoted to develop methods that are more efficient than classical Monte Carlo approaches. Alternatively to sampling methods, one could consider generalized Polynomial Chaos (gPC) approximations of the solution with respect to a finite number of random variables parameterizing the probability space. At this step practical approximations can be obtained by different numerical techniques such as Galerkin projection or stochastic collocation method. However this approach could still be affected by the so called 'curse of dimensionality', especially when the solution features a complex dependence on the input parameters. This can be an issue in evolution problems, as the dependence of the solution on the random parameters can significantly vary in time . Then the approximation of the solution by means of fixed polynomial basis functions requires during the evolution an increasing number of terms to maintain a proper level of accuracy, which possibly implies a too high computational effort.

We propose here a low rank method based on the Dynamically Orthogonal Field (DOF) approach, according to which the solution is approximated as a linear combination of a small number of deterministic orthogonal basis functions multiplied by random coefficients, both of them evolving in time in order to keep the dimensionality of the approximate problem as low as possible. To fix the idea we briefly introduce the mathematical settings. We consider the following time dependent stochastic PDE:

$$\begin{cases} \frac{\partial u(x,t,\omega)}{\partial t} = \mathcal{L}[u(x,t,\omega),\omega], & x \in D,\ t \in [0,T],\ \omega \in \Omega \\ \mathcal{B}\left[u(\xi,t,\omega)\right] = h(\xi,t), & \xi \in \partial D,\ t \in [0,T], \\ u(x,t=0,\omega) = u_0(x,\omega), & x \in D,\ \omega \in \Omega, \end{cases} \tag{1}$$

where D is the physical domain, ω is a random elementary event in a complete probability space $(\Omega, \mathcal{A}, \mathcal{P})$ and $\mathcal{L}$ is a general (linear or non-linear) differential operator. Now we look for an approximate solution

of rank S of the form:

$$u_S(x,t,\omega) = \bar{u}(x,t) + \sum_{i=1}^{S} u_i(x,t) y_i(t,\omega). \tag{2}$$

where $\bar{u} \simeq \mathbf{E}[u]$, $u_1, ...u_S$ are $L^2(D)-$orthogonal deterministic basis functions and $y_1, ..., y_S$ are zero mean stochastic variables. From a variational point of view, if $\mathcal{M}_S$ denotes the manifold of all the functions of rank S, the DO approximate solution u_S is obtained by projecting the residual of the governing equation onto the tangent space to $\mathcal{M}_S$ at $u_S(t)$ at each time, i.e.

$$\mathbb{E}\left[\left\langle \frac{\partial u_S(\cdot,t,\omega)}{\partial t} - \mathcal{L}(u_S(\cdot,t,\omega);\omega),\, v(\cdot,\omega)\right\rangle\right] = 0, \quad \forall v \in \mathcal{T}_{u_S(t)}\mathcal{M}_{\mathcal{S}} \tag{3}$$

The approximate solution is therefore forced to belong to a S dimensional manifold but at the same time the stochastic coefficients and the deterministic basis functions adapt to the structure of the solution at each time, in order to be as close as possible to the best rank S approximation, which is given by the truncated $S-$terms Karhunen-Lòeve expansion.

Our goal is to find a good balance between computational saving and effectiveness of the approximation. Indeed even if the DO expansion does not necessarily coincide with the Karhunen-Lòeve decomposition, the DO approach allows us to derive directly from the governing equation a coupled system of $S+1$ deterministic and S stochastic equations that completely characterize our approximate solution u_S at each time. An analogous approach in the deterministic framework, based on the Dirac-Frenkel variational principle, is present in the literature [1]-[3] and it is adopted in the field of the quantum dynamics.

By investigating the nature of the manifold of the solution we analyse the effectiveness of the DO approximate solution. In particular we provide an error analysis of the DO approximation of stochastic parabolic equations. We exploit the curvature bounds of the manifold given in [3] to show that, under suitable assumptions, the DO error can be bounded in terms of the best rank S approximation error. Moreover, we show that in the case of deterministic linear operator, with an $S-$ dimensional stochastic initial datum, the DO approximate solution coincides with the exact solution.

Finally we propose a numerical method, able to handle the problem of the over approximation. In that case the number of modes S is bigger then the effective dimension of the exact solution and the covariance matrix of the stochastic coefficients is singular. This point has a great importance since it might happen that the exact solution has a rank smaller than the one used in the DO approximation at some time instant. This is typically the case when starting the algorithm from a deterministic initial condition or when looking at a system of equations converging asymptotically to a deterministic equilibrium, in which case the rank tends asymptotically to zero. The strategy we adopt consists in diagonalizing the covariance matrix, or rather re-orthogonalizing the random coefficients at each time step. Then only the modes associated to stochastic variables with variance bigger than zero will evolve while the other remain constant. From the numerical point of view, the DO system is decoupled in deterministic and stochastic equations, for the former we use Finite Element Methods and for the latter the Stochastic Collocation. The DO method and its convergence properties are assessed with several numerical examples.

References:

[1] Othmar Koch and Christian Lubich, Dynamically low-rank approximation, SIAM J. Matrix Anal.Appl., Vol.29, No.2(2007), pp.434-454.

[2] Themistoklis P. Sapsis and Pierre F.J. Lermusiaux, Dynamically orthogonal field equations for continuous stochastic dynamical systems, Physica D, 238 (2009) 2347-2360,

[3] Dajana Conte and Christian Lubich, An error analysis of the multi-configuration time-dependent Hartree method of quantum dynamics, ESAIM: M2AN. Vol.44, 2010, pp.759-780.

Short biography – Undergraduate Studies: (*2005 - 2008*) Bachelor Degree in Applied Mathematics, Università degli Studi di Perugia, Italy; (*2009 - 2010*) curriculum Integrations of Engineering Mathematics (single courses), Politecnico di Milano; (*2010 - 2012*) Master Degree in Engineering Mathematics, Politecnico di Milano, Italy. Actual Position: PhD Student at the first year. The PhD project focuses on dynamical low rank approximations of PDEs with random parameters and it is financially supported by FNS.

Multilevel Monte Carlo for Bayesian Inverse Problems

R. Gantner
ETH Zürich

Affiliation: *Seminar for Applied Mathematics*, ETH Zürich, Rämistrasse 101, 8092 Zürich, Switzerland

Email: robert.gantner@sam.math.ethz.ch – **URL**: http://www.sam.math.ethz.ch

Master: MSc. CSE ETH Zürich

Ph.D. (2013-2016): ETH Zürich

Supervisor(s): Prof. Christoph Schwab (ETH Zürich) and Dr. Claudia Schillings (ETH Zürich)

Abstract:

Introduction
In recent years, various methods have been developed for solving parametric operator equations, focusing on the estimation of parameters given measurements of the parametric solution, subject to a stochastic observation error model. The *Bayesian approach* [1] to such inverse problems for PDEs will be considered here and solved using adaptive, deterministic sparse tensor Smolyak quadrature schemes from [2, 3]. Multiple solutions of the Bayesian inverse problem based on different measurements are often averaged using a standard Monte Carlo approach. We develop a multilevel Monte Carlo method achieving an error of the same order while requiring less work [4, 5, 6].

Bayesian Inversion of Parametric Operator Equations
We assume an operator equation depending on a distributed, uncertain parameter u with values in a separable Banach space X of the form

$$\text{Given } u \in \tilde{X} \subseteq X, \text{ find } q \in \mathcal{X}: \quad A(u;q) = F(u) \quad \text{in} \quad \mathcal{Y}' \tag{1}$$

where we denote by $\mathcal{X}$ and $\mathcal{Y}$ two reflexive Banach spaces over $\mathbb{R}$ with (topological) duals $\mathcal{X}'$ and $\mathcal{Y}'$, respectively and $A(u;\cdot) \in \mathcal{L}(\mathcal{X}, \mathcal{Y}')$. Assuming that the forcing function $F : \tilde{X} \mapsto \mathcal{Y}'$ is known, and the uncertain operator $A(u;\cdot) : \mathcal{X} \mapsto \mathcal{Y}'$ is locally boundedly invertible for uncertain input u in a sufficiently small neighborhood $\tilde{X}$, we define the *uncertainty-to-observation map* $\mathcal{G} : \tilde{X} \mapsto \mathbb{R}^K$ with the structure

$$X \supseteq \tilde{X} \ni u \mapsto \mathcal{G}(u) := \mathcal{O}(G(u;F)) \in Y\ . \tag{2}$$

Here, $\tilde{X} \ni u \mapsto q(u) = G(u;F) \in \mathcal{X}$ denotes the response of the forward problem for a given instance of $u \in \tilde{X}$ and $\mathcal{O}$ an *observation operator* $\mathcal{O} \in \mathcal{L}(\mathcal{X}, \mathbb{R}^K)$, $K < \infty$. The goal of computation is the low-order statistics of a *quantity of interest* (QoI) ϕ given noisy observational data δ of the form $\delta = \mathcal{G}(u) + \eta$, where δ represents the observation $\mathcal{G}(u)$ perturbed by the normally distributed noise η. We assume u to be parametrized by $u = u(\mathbf{y}) := \langle u \rangle + \sum_{j\in\mathbb{J}} y_j \psi_j \in X$ for some "nominal" value $\langle u \rangle$ and coefficient sequence $\mathbf{y} = (y_j)_{j\in\mathbb{J}}$, $\mathbb{J} = \{1, \ldots, J\}$ where the y_j are uniformly distributed on $[-1, 1]$.

Bayes' theorem characterizes moments of the QoI as mathematical expectations with respect to the prior measure μ_0 on U, which here is given as the countable product of uniform measures. In particular, we are interested in $\phi = G$, the response of the system. To this end, we use Bayes' Theorem to obtain an expression for $\mathbf{y}|\delta$, as in [7, 1]. This yields our desired expectation as an integral over the prior measure μ_0. Defining $Z_\delta := \int_U \exp\left(-\Phi(\mathbf{y};\delta)\right) \mu_0(\mathrm{d}\mathbf{y}) > 0$, we obtain

$$\mathbb{E}^{\mu^\delta}[\phi] = \int_U \phi(\mathbf{y})\, \mu^\delta(\mathrm{d}\mathbf{y}) = \frac{1}{Z_\delta} \int_U \phi(\mathbf{y}) \exp\left(-\frac{1}{2}\|\delta - \mathcal{G}(\mathbf{y})\|_\Gamma^2\right) \mu_0(\mathrm{d}\mathbf{y}) =: \frac{Z'_\delta}{Z_\delta}. \tag{3}$$

This formulation of the expectation $\mathbb{E}^{\mu^\delta}[\cdot]$ is based on just one measurement δ. For a given model for the measurement errors η, we would like to additionally compute the expectation over the assumed error distribution, in this case $\gamma_\Gamma^K(\eta)$, the K-variate Gaussian measure with s.p.d. covariance matrix Γ.

Here, we assume the observation noise η to be statistically independent from the uncertain parameter u in (1). Thus, the total expectation of the QoI ϕ in terms of Z'_δ and Z_δ is $\mathbb{E}^{\gamma_\Gamma^K}\left[\mathbb{E}^{\mu^\delta}[\phi]\right] = \int_{\mathbb{R}^K} \frac{Z'_\delta}{Z_\delta}\Big|_{\delta=\mathcal{G}(\mathbf{y}_0)+\eta} \gamma_\Gamma^K(\mathrm{d}\eta)$, where $\mathcal{G}(\mathbf{y}_0)$ denotes the observation at the unknown, exact parameter $\mathbf{y}_0$.

In practice, we are given a set of measurements $\Delta := \{\delta_i,\, i = 1, \ldots, M\}$ with which this outer expectation should be approximated. The measurements can be taken at different positions, i.e. with respect to different observation maps $\mathcal{O}_i$ in (2). We consider the notationally more convenient case where the measurements are all obtained using the same observation map. We do, however, impose the restriction that the measurements are homoscedastic, i.e. δ_i is Gaussian with the same covariance Γ for all $\delta_i \in \Delta$.

Approximation of Posterior Expectation

The inner expectation over the posterior distribution μ^δ is replaced by an approximation $E^{\mu^\delta}_{\tau_L}[\phi]$ with tolerance parameter $\tau_L > 0$. We assume the work required to compute this approximation to be bounded by $C(\Gamma)\tau_L^{-s}$, with $C(\Gamma) > 0$ independent of τ_L and $s > 0$. Our method of choice for approximating $\mathbb{E}^{\mu^\delta}[\phi]$ is the adaptive Smolyak quadrature algorithm developed in [2, 3], which adaptively constructs a sparse tensor quadrature rule that approximates Z_δ and Z'_δ. For forward problems belonging to a certain sparsity class, analytic regularity of the Bayesian posterior suggests dimension-independent convergence rates for the adaptive, deterministic Smolyak quadrature fulfilling the work bound $C(\Gamma)\tau_L^{-s}$, where s depends on the sparsity class.

Binned Multilevel Monte Carlo

The approach proposed here is based on the multilevel Monte Carlo method originally applied by [6] and formulated in the current form for PDEs by [4]. Our approximation to $\mathbb{E}^{\gamma_\Gamma^K}\left[\mathbb{E}^{\mu^\delta}[\phi]\right]$ is given by

$$E^{\gamma_\Gamma^K}_{\mathrm{ML},L}[E^{\mu^\delta}_{\tau_L}[\phi]] := \sum_{\ell=0}^{L} E^{\gamma_\Gamma^K}_{M_\ell}\left[E^{\mu^\delta}_{\tau_\ell}[\phi] - E^{\mu^\delta}_{\tau_{\ell-1}}[\phi]\right], \tag{4}$$

where $E^{\gamma_\Gamma^K}_{M_\ell}[\cdot]$ denotes the sample mean over M_ℓ samples and $E^{\mu^\delta}_{\tau_\ell}[\cdot]$ denotes the posterior expectation approximation introduced above. We show that, assuming a certain distribution of samples per level, one can find a tolerance for each level such that the rate of convergence of the error e_{tot} vs. the work W^L_{tot} fulfills the optimal relationship $e_{\mathrm{tot}} = \mathcal{O}\left(\left(W^L_{\mathrm{tot}}\right)^{-\frac{1}{2}}\right)$, which is superior to Monte Carlo depending on the sparsity class of the underlying problem.

Applications

The proposed approach is applicable for instance for definite and indefinite elliptic and parabolic evolution problems with scalar and tensoral unknowns. Furthermore, uncertainty in domains and high-dimensional initial value problems can be treated. Numerical experiments yielding the optimal rate of convergence when using the binned multilevel Monte Carlo algorithm will be presented and compared to standard Monte Carlo simulations.

References:

[1] A. M. Stuart. Inverse problems: a Bayesian perspective. *Acta Numer.*, 19:451–559, 2010.

[2] Cl. Schillings and Ch. Schwab. Sparse, adaptive Smolyak quadratures for Bayesian inverse problems. *Inverse Problems*, 29(6):065011, 2013.

[3] Cl. Schillings and Ch. Schwab. Sparsity in Bayesian inversion of parametric operator equations. Technical Report 2013-17, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2013.

[4] A. Barth, Ch. Schwab, and N. Zollinger. Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients. *Numer. Math.*, 119(1):123–161, September 2011.

[5] R. Gantner, Cl. Schillings, and Ch. Schwab. Multilevel Monte Carlo for Bayesian inverse problems. (submitted), 2014.

[6] M. B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, pages 1–25, 2008.

[7] Ch. Schwab and A. M. Stuart. Sparse deterministic approximation of Bayesian inverse problems. *Inverse Problems*, 28(4):045003, 2012.

Short biography – After completing his master's degree in Computational Science and Engineering at ETH Zurich, Robert Gantner is now pursuing a PhD in the group of Prof. Schwab at the Seminar for Applied Mathematics. The main focus is the development and high-performance implementation of novel algorithms for Bayesian inverse problems, including Multilevel Monte Carlo and Smolyak quadratures.

An intuitive variance based variable screening method

R. Sala
Universita degli Studi di Firenze

Affiliation: *Dipartimento di Ingegneria Industriale*, Universita degli Studi di Firenze, Via di Santa Marta 3, 50139 Firenze, Italy

Email: ramses.sala@unifi.it – **URL**: http://www.dief.unifi.it

Master: Technische Universitaet Muenchen

Ph.D. (2012-2015): Universita degli Studi di Firenze

Supervisor(s): Prof. M. Pierini (Universita degli Studi di Firenze) and Prof. N. Baldanzini (Universita degli Studi di Firenze)

Abstract:

Variance based sensitivity analysis methods are a valuable tool for modelers in engineering and other disciplines to quantitatively determine the strength of the relation between input factors and model outputs. There is a great variety of variance based sensitivity methods [4]. For some complex simulation based models (such as for example full vehicle crashworthiness simulation) it can be infeasible to apply established quantitative variance based sensitivity because the models are too computationally expensive to perform enough function evaluations that would meet the sample requirements. For such expensive models qualitative screening methods can be applied instead. In [2] a unified approach was presented for variable screening using a radial design sampling strategy to estimate the elementary effects, and for an increasing number of samples the total sensitivity indices based could be estimated.

In this communication a simple and intuitive variance based variable screening method is presented and applied to test functions and multidisciplinary automotive simulation models. Although provided enough samples, the approach could approximate the first order sensitivity indices with arbitrary close accuracy; the main reason to bring this approach under attention is for variable screening purposes, when the number of samples is too small to apply established quantitative first order index estimation methods such as EFAST [5], RBD [6] and EASI [7]. Similarly to the previously referenced unified method [2], also this method enables a seamless extension of the qualitative variable screening to a quantitative sensitivity analysis, all by itself or more efficient when combined with for example the EASI method.

The general idea of the approach is to estimate for each variable the first order contribution to the output variance, by approximating conditioned means and their variances, over discrete intervals of finite size in scatterplot projections. For a high number of intervals the method converges to the definition of the first order effect sensitivity index or main effect index given in [3]. For a low number of discrete intervals the resulting indices can be too distantly related to the sensitivity indices to label them sensitivity index estimates. The theoretical values towards which the resulting indices, for fixed intervals converge when the number of sample point increases are however extensive properties of the model with respect to the fixed set of intervals chosen. The approach allows selecting the degree of spatial discretization to which the first order sensitivities are approximated. For the remaining of this document the converged values of those (degenerate) indices will be called Discrete Interval based Sensitivity indices (DIS) indices.

For a realistic case with a limited fixed number of samples available, that is insufficient to estimate first order indices without a large statistical error, a compromise has to be made. Smaller intervals, lead to theoretical DIS indices that are closer to the sensitivity indices, but since the number of samples per interval is low the accuracy of the estimation will be low. Whereas few but large intervals contain more samples per interval and the resulting DIS index estimations are more accurate estimates for those DIS indices that correspond to the larger intervals. The tradeoff between statistical accuracy and spatial discretization resolution can be selected.

The approach is extendable for higher order interaction effects, by creating subsequent projections and divisions in discrete intervals. However, inevitably smaller discrete intervals and subintervals require a

larger sample density to obtain relevant results, and therefore only interactions of second are considered. The method is tested to work for data obtained with sampling strategies that result in evenly distributed sample spacing, with a low correlation such as obtained by pseudo-random and quasi-random sampling methods, and combinations thereof.

An intuitive overview of the method will be provided, followed by case studies using a high dimensional instance of the Sobol-G function and vehicle simulation models. Convergence comparisons with quantitative estimation methods for first order sensitivity indices will be provided, and the screening effectiveness will be compared with the modified elementary effect method and regression methods. Besides presenting the practical value of coarse discrete interval based variance contribution screening, the intuitive principle of the method could also lower the threshold to start using variance based sensitivity analysis methods, for practitioners that are less experienced in the field of sensitivity analysis.

References:

[1]Campolongo, Francesca, Jessica Cariboni, and Andrea Saltelli. "An effective screening design for sensitivity analysis of large models." Environmental modelling and software 22.10 (2007): 1509-1518.

[2] Campolongo, Francesca, Andrea Saltelli, and Jessica Cariboni. "From screening to quantitative sensitivity analysis. A unified approach." Computer Physics Communications 182.4 (2011): 978-988.

[3] Sobol, Ilya M. "Sensitivity estimates for nonlinear mathematical models." Mathematical Modelling and Computational Experiments 1.4 (1993): 407-414.

[4] Saltelli, Andrea, and Paola Annoni. "How to avoid a perfunctory sensitivity analysis." Environmental Modelling & Software 25.12 (2010): 1508-1517.

[5] Saltelli, Andrea, Stefano Tarantola, and KP-S. Chan. "A quantitative model-independent method for global sensitivity analysis of model output." Technometrics 41.1 (1999): 39-56.

[6] Tarantola S., Gatelli, D. and T. Mara (2006) Random Balance Designs for the Estimation of First Order Global Sensitivity Indices, Reliability Engineering & System Safety, 91:6, 717-727 Morris, Max D. "Factorial sampling plans for preliminary computational experiments." Technometrics 33.2 (1991): 161-174.

[7] Plischke, Elmar. "An effective algorithm for computing global sensitivity indices (EASI)." Reliability Engineering & System Safety 95.4 (2010): 354-360.

Short biography – The work is performed in the scope of the task "managing the conflict between lightweight design, NVH and Crashworthiness requirements in vehicle design" within the GRESIMO project that is funded by the EC under grand agreement 290050. The presented variable screening approach resulted as a side product during the activities for this task which the focus is more on multi-disciplinary optimization of automotive vehicle structures.

PC-Kriging: Combining Polynomial Chaos Expansions and Universal Kriging

R. Schöbi
ETH Zürich

Affiliation: *Chair of Risk, Safety and Uncertainty Quantification*, ETH Zürich, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland

Email: schoebi@ibk.baug.ethz.ch – **URL**: http://www.ibk.ethz.ch/su/people/schoebir/index_EN

Master: ETH Zürich

Ph.D. (2013-2016): ETH Zürich

Supervisor(s): Prof. B. Sudret (ETH Zürich)

Abstract:

There is an ongoing trend for replacing real life experiments by computer simulations. The physical behaviour is substituted by a computational model which approximates the system response of the physical system. Advances in research lead to more complex and more accurate computational models which are at the same time more costly to evaluate, *i.e.* time-consuming. There is a conflicting situation between accuracy and speed. Applications such as reliability analysis or optimization algorithms require a large number of model evaluations, *e.g.* the computation of a system's failure probability or the optimal value of a parameter. These operations are reasonable when the computational model is easy-to-evaluate, *i.e.* when a model evaluation is inexpensive and the system response of a large number of input samples is processed fast.

Metamodelling provides a framework for replacing an expensive-to-evaluate computational model $Y = \mathcal{M}(\boldsymbol{X})$ by a simple although approximative surrogate model. The surrogate model (also called metamodel) allows one to predict the system response of a large number of input samples at low cost. The metamodel is built from a *small* number N of support points called the experimental design $\boldsymbol{\mathcal{X}} = \{\boldsymbol{\chi}^{(i)}, i = 1, \ldots, N\}$ for which the original model is evaluated. The input and output values/vectors are used to determine an appropriate metamodel with a certain metamodelling technique. Two of the more popular non-intrusive metamodelling techniques are *Polynomial Chaos Expansions* (PCE) and *Kriging* (also called Gaussian process modelling).

PCE surrogates the computational model $\mathcal{M}$ by a finite set of orthonormal polynomials in the input variables [1]. In the context of uncertainty quantification, the latter are defined in coherency with the probability distribution functions of those input variables. The coefficients of a PC expansion may be computed using *e.g.* least-square minimization algorithms. PCE assumes that the computational model is a black-box model, *i.e.* only information about the input values and model response are available (the inner structure and features of the model (nonlinearity, interaction between parameters, etc.) are assumed unknown).

Kriging is called a stochastic metamodelling technique which assumes that the computational model is a realization of a Gaussian random field whose properties are inferred from the experimental design and the associated model output [2]. The experimental design points provide the information to compute the optimal correlation parameters by *e.g.* maximum likelihood method. The prediction of the surrogate at a new point results in a Gaussian variable represented by its mean value and variance value called Kriging mean prediction and prediction variance.

Although these two techniques have become popular for solving uncertainties propagation, optimization or sensitivity problems, their combination has not been considered yet. In this paper, the new metamodelling technique *Polynomial-Chaos-Kriging* (PC-Kriging) is proposed. This metamodel is based on the classical universal Kriging approach where the trend (regression part) is a sum of functions in the general case. In

PC-Kriging a sparse set of orthonormal polynomials serves as the trend of the universal Kriging model. The general formulation of the metamodel is then:

$$\mathcal{M}(\boldsymbol{X}) \approx \mathcal{M}^{(\mathsf{PCK})}(\boldsymbol{x}) = \sum_{k=1}^{P} \beta_k \, f_k(\boldsymbol{x}) + \sigma^2 \, Z(\boldsymbol{x}, \omega) \tag{1}$$

where $\sum_{k=1}^{P} \beta_k \, f_k(\boldsymbol{x})$ is the mean value of the Gaussian process (so-called trend) and $Z(\boldsymbol{x}, \omega)$ is a zero mean, unit variance Gaussian process described by a set of hyper-parameters $\omega = \{\boldsymbol{\theta}, R\}$. The auto-correlation function $R(\boldsymbol{x}, \boldsymbol{x}'; \boldsymbol{\theta})$ describes the correlation between two samples given its parameters $\boldsymbol{\theta}$. $\boldsymbol{f}(\boldsymbol{x}) = \{f_k(\boldsymbol{x}), k = 1, \ldots, P\}$ are the multidimensional orthonormal polynomials, its coefficients are β_k, $\boldsymbol{x}$ are realizations of the input variables $\boldsymbol{X}$, σ^2 is the Kriging variance. The sparse set of P polynomials is determined by using hyperbolic index sets and least angle regression based of the experimental design $\{\boldsymbol{\chi}^{(i)}\}$ [3]. In an iterative manner and one-by-one, a polynomial out of the determined sparse set is added to the Kriging model. The iteration starts with the polynomial which is the most correlated to the system response. Then the classical equations to fit the Kriging model are used to determine the correlation parameters, the Kriging variance and the trend coefficients [4]. The P PC-Kriging models are then compared by means of the leave-one-out error and the optimal PC-Kriging metamodel (with minimal leave-one-out error) is chosen.

The performance of PC-Kriging is compared to ordinary Kriging (constant trend β_0) and pure PCE on six easy-to-evaluate benchmark problems in the field of optimization and metamodelling. The results show that PC-Kriging performs better or at least as good as PCE and/or Kriging. Especially, for small experimental designs PC-Kriging is preferable to the two distinct approaches. From the numerical experiments is appears that some problems are better suited for PCE whereas some problems are better handled by Kriging. PC-Kriging converges to the best of the two simple approaches and leads to a smaller squared residual error. For large sample sizes, PC-Kriging performs similar to PCE, so that the added value of PC-Kriging is questionable in that case.

Heuristically, the behaviour of PC-Kriging can be explained as follows: the set of polynomials approximates the global behaviour whereas the correlation part models the local variabilities between the support points. The combination of these two effects leads to a higher accuracy and thus to a better metamodel. The validation of PC-Kriging is shown on numerous analytical benchmark functions which are easy-to-evaluate.

References:

[1] R. Ghanem and P. Spanos. *Stochastic Finite Elements : A Spectral Approach*. Courier Dover Publications, 2003.

[2] T.J. Santner, B.J. Williams, and W.I. Notz. *The Design and Analysis of Computer Experiments*. Springer, New York, 2003.

[3] G. Blatman and B. Sudret. Adaptive sparse polynomial chaos expansion based on Least Angle Regression. *J. Comput. Phys*, 230:2345–2367, 2011.

[4] F. Bachoc. Cross validation and maximum likelihood estimations of hyper-parameters of Gaussian processes with model misspecifications. *Comp. Stat. Data An.*, 66:55–69, 2013.

Short biography – R. Schöbi graduated in 2012 from ETH Zürich with a Master's degree in civil engineering where his Master's Thesis was entitled "Subset Simulation in Engineering Problems." Since May 2013, he is a Ph.D. student at the Chair of Risk, Safety and Uncertainty Quantification of ETH Zürich under the supervision of Prof. B. Sudret. His research topics include the quantification and propagation of epistemic uncertainty using mixed probabilistic/nonprobabilistic approaches.

Uncertainty quantification and visualization for functional random variables

Shorash Miro
Ruhr-Universität Bochum

Affiliation: *Chair of Computing in Engineering*, Ruhr-Universität Bochum, Universitätstrasse 150, D-44801 Bochum, Germany

Email: Shorash.Miro@ruhr-uni-bochum.de – **URL**: http://www.inf.bi.ruhr-uni-bochum.de/index.php?option=com_staff&id=116&Itemid=145&lang=en

Master: Ruhr-Universität Bochum

Ph.D. (11.2010-11.2014): Ruhr-Universität Bochum

Supervisor(s): Prof. D. Hartmann (Ruhr-Universität Bochum) and Prof. T. Schanz (Ruhr-Universität Bochum)

Abstract:

The ground model is central to computational tunneling, where, a realistic ground model is crucial for predicting the distributions and magnitudes of the strains and, consequently, reducing the surface settlements caused by the Tunnel Boring Machine (TBM) propagation. However, due to the complex interactions between the ground, the driving machine, the lining tube and the built environment, the accurate assignment of in-situ system parameters for numerical simulation in mechanized tunneling is always bounded with tremendous difficulties. Furthermore, the more accurate these parameters are, the more applicable the responses gained from computations will be. In particular, if the entire length of the tunnel lining is considered, then the appropriate selection of various kinds of ground parameters is accountable for the success of a tunnel project and, more importantly, will prevent serious casualties. As a consequence, only the realization of the system identification approach can result in improved and more sophisticated numerical predictions of the spatio-temporal ground behavior induced by driving the tunnel. In this context, methods of system identification for the adaptation of numerical simulation ground models are presented. These methods consider both deterministic and probabilistic approaches for typical scenarios for variations or changes in the ground model.

In the deterministic approach, measurements of system responses obtained during the tunneling process (e.g. surface subsidence) are compared to their numerical counterparts computed from the numerical simulation (forward calculation). If there is no match, the serious deviations (defects or residuals) are iteratively minimized by applying a different derivative-free optimization algorithm and taking the parameters of the ground model, used in the numerical simulation, as minimizing parameters for the minimization problem (inverse problem). Hereby, the underlying optimization problem constitutes a highly nonlinear as well as non-smooth optimization problem associated with multiple optima (nonstandard optimization problem), due to the ground model that represents a complex, mechanically-hydraulically coupled and a three dimensional boundary value problem.

Customarily, the geotechnical applications are associated with uncertainties, where, the significantly scattered material properties (e.g. stiffness, strength, permeability) of the soil, due to the inherent randomness of its nature, yields aleatoric (objective) uncertainty. In addition to that, lack of data, of information about events and processes and of understanding the physical laws results in Knowledge (epistemic/subjective) uncertainty. In order to obtain realistic results and to quantify the underlying uncertainties in the parameter estimation process, a Bayesian probabilistic approach for the inverse problem is introduced. This approach is able to include the prior information about the parameters, which may be captured from bore holes sunken in the target area of the tunnel alignment. To this end, a numerical approximation of the probabilistic solution using Markov chain Monte Carlo is carried out.

In order to make the identification process as efficient and robust as possible, it is favorable to reduce the number of the parameters to be identified (consequently, the number of numerical simulation runs) by

performing an effective sensitivity analysis. With this analysis, the importance of each unknown model parameter with respect to the system response is evaluated such that an effective selection of the dominating model parameters is achieved. For this purpose, two global sensitivity analysis approaches, namely elementary effect and variance-based analyses, have been adopted and validated. In the elementary effect approach a strategy for considering the dependencies, which result from a set of constraints between different parameters, is proposed. As a result, the propagating sensitivities of subsoil parameters during the excavation process of the mechanized tunnel are achieved. On the basis of this, an efficient choice of the parameters that have to be identified is enabled.

The computationally expensive finite element tunnel simulation has been replaced by an accurate meta-model. In this regard, and in order to construct robust and reliable meta-models, three meta-modeling approaches have been implemented and tested. These approaches are quadratic polynomial regression QPR, moving least squares MLS, and proper orthogonal decomposition with radial basis functions POD-RBF. Furthermore, an extended version of the latest approach is proposed and implemented. This version constitutes a combination of proper orthogonal decomposition with extended radial basis functions and is abbreviated as POD-ERBF. Its performance has been systematically compared with the three aforementioned methods through a comparative study utilizing pure mathematical test functions. With this study the best performing meta-model, which is the extended version of POD-RBF, is selected to replace the tunneling simulation model in the system identification and global sensitivity analysis.

For the evaluation of influences of subsoil parameter uncertainties on the mechanized tunnel safety and stability, a probabilistic analysis has been performed. In this analysis, the input parameter uncertainties are mathematically represented by adequately chosen probability density functions, consequently, the propagation of these uncertainties are evaluated by performing a Monte Carlo-based simulation of the computationally cheap surrogate model that is developed to replace the computationally expensive finite element tunnel simulation. In addition to that, a global sensitivity analysis is conducted for quantifying the impact of each uncertain parameter on different system responses that are considered in this study. The variations of system responses, which result from input parameters propagating uncertainties, are compared with predetermined threshold values, and based on that, reliability-based failure criteria of the tunneling system are defined and probabilistically quantified. As a result, the influence of subsoil parameter uncertainty representation on failure probabilities is addressed and evaluated.

Short biography – After obtaining a M.Sc. degree in Computational Engineering at Ruhr-Universität Bochum in Germany in October 2010, I started my PhD study in November 2010. My research work is part of the Collaborative Research Center SFB 837 "Interaction Modeling in Mechanized Tunneling", funded by the German Research Foundation DFG. I work within sub-project C2 "System identification methods for the adaptation of numerical simulation models". My research focuses are: System identification considering uncertainties, Bayesian inverse analysis, global sensitivity analysis, meta-modeling, and failure analysis of the mechanized tunnel due to subsoil parameter uncertainties. After obtaining a M.Sc. degree in Computational Engineering at Ruhr-Universität Bochum in Germany in October 2010, I started my PhD study in November 2010. My research work is part of the Collaborative Research Center SFB 837 "Interaction Modeling in Mechanized Tunneling", funded by the German Research Foundation DFG. I work within sub-project C2 "System identification methods for the adaptation of numerical simulation models". My research focuses are: System identification considering uncertainties, Bayesian inverse analysis, global sensitivity analysis, meta-modeling, and failure analysis of the mechanized tunnel due to subsoil parameter uncertainties.

Multi Level Monte Carlo methods with Control Variate for elliptic Stochastic Partial Differential Equations

F. Tesei
École Polytechnique Fédérale de Lausanne

Affiliation: *Chair of Scientific computing and uncertainty quantification*, EPF Lausanne, Bâtiment MA Station 8, CH-1015 Lausanne, Switzerland

Email: francesco.tesei@epfl.ch – **URL**: http://www.csqi.epfl.ch

Master: Politecnico di Milano

Ph.D. (2012-2015): École Polytechnique Fédérale de Lausanne

Supervisor(s): Prof. F. Nobile

Abstract:

We consider the numerical approximation of a partial differential equation (PDE) with random coefficients. These type of problems can be found in many applications in which the lack of available measurements makes an accurate reconstruction of the coefficients appearing in the mathematical model unfeasible. In particular we focus on the model problem of an elliptic partial differential equation with random diffusion coefficient, modeled as a random field with limited spatial regularity. This approach is inspired by the groundwater flow problem which has a great importance in hydrology: in this context the diffusion coefficient describes the permeability of the subsoil and is often modeled as a lognormal random field.

More precisely, we consider the problem defined on a physical domain $D \in \mathbb{R}^d$ and on the set of all possible events Ω:

$$\begin{cases} -\text{div}(a(x,\omega)\nabla u(x,\omega)) = f(x), & x \in D, \omega \in \Omega, \\ u(x,\omega) = g(x) & x \in \Gamma_D \subset \partial D, \omega \in \Omega, \\ a(x,\omega)\nabla u(x,\omega)\cdot \mathbf{n} = 0 & x \in \Gamma_N \subset \partial D, \omega \in \Omega. \end{cases}$$

where $a(x,\omega) = e^{\gamma(x,\omega)}$ and $\gamma(x,\omega)$ is a Gaussian random field. Several models have been proposed in the literature for the covariance function of the log-permeability γ leading to realizations having varying spatial smoothness. In particular, a widely used covariance function is the exponential one, $\text{cov}_\gamma(x_1,x_2) = \sigma^2 e^{-\frac{\|x_1-x_2\|}{l_c}}$, that has realizations with Hölder continuity $\mathcal{C}^{0,\alpha}$ with $\alpha < \frac{1}{2}$. In this work we focus on covariance functions belonging to the so called Matérn family. These covariance functions depend on a parameter ν that defines the spatial smoothness of the field, ranging from very low spatial regularity as in the exponential covariance case ($\nu = 0.5$) to very high spatial regularity as in the gaussian covariance case ($\nu \to \infty$).

Models with limited spatial smoothness pose great numerical challenges. The first step of their numerical approximation consists in building a series expansion of the input coefficient; here we use a Fourier expansion. Whenever the random field has low regularity, such expansions converge very slowly and this makes the use of deterministic methods such as Stochastic Collocation on sparse grids highly problematic since it is not possible to parametrize the problem with a relatively small number of random variables without a significant loss of accuracy. A natural choice is to try to solve such problems with a Monte Carlo (MC) type method. In formulas, let $Q(u)$ be the quantity of interest (QoI) related to the solution u of the elliptic stochastic PDE. The MC estimator of the QoI and its corresponding mean square error (MSE) are given by:

$$\hat{Q}^{MC}_{h,M} = \frac{1}{M}\sum_{i=1}^{M}\left(Q^i_h\right), \qquad e(\hat{Q}^{MC}_{h,M})^2 = \frac{\mathbb{V}\text{ar}(Q_h)}{M} + \mathbb{E}[Q_h - Q]^2,$$

where Q_h denotes the approximate evaluation of Q computed from the finite element solution u_h of the PDE with mesh size h and Q_h^i are i.i.d. replica of Q_h corresponding to i.i.d. realizations of the log-permeability field γ. On the other hand it is well known that the convergence rate of the standard Monte Carlo method is quite slow, making it impractical to obtain an accurate solution. Indeed, the computational cost of a Monte Carlo simulation is given by the number of samples of the random field multiplied by the cost needed to solve a single deterministic PDE, which requires a very fine mesh due to the roughness of the coefficient. Multilevel Monte Carlo methods (MLMC) have already been proposed in the literature in order to reduce the variance of the Monte Carlo estimator, and consequently reduce the number of solves on the fine grid. These methods introduce a sequence of increasingly fine grids $\mathcal{T}_{h_0}, ..., \mathcal{T}_{h_L}$ and, thanks to the linearity of the expectation operator, they split the total work on different levels in order to get a cheaper estimator $\hat{Q}^{MLMC}_{h,\{M_l\}}$ in terms of computational cost.

In this work we propose to use a MLMC approach combined with an additional control variate variance reduction technique on each level in order to solve the elliptic SPDE for different choices of the covariance function of the input field, within the Matérn family. The control variate is obtained as the solution of the PDE with a regularized version of the lognormal random field $\gamma^\epsilon(x,\omega) = \gamma(x,\omega) * \phi_\epsilon(x)$ with $\phi_\epsilon(x)$ a smooth convolution kernel. Since γ^ϵ is smooth, the mean of the correspondent QoI Q^ϵ_h can be successfully computed with a Stochastic Collocation method on each level. The solution of this regularized problem turns out to be highly positively correlated with the solution of the original problem on each level, which makes the control variate technique very effective.
Within this Monte Carlo framework the choice of a suitable regularized version of the input random field is the key element of the method; we propose to regularize the random field by convolving the log-permeability with a centered Gaussian kernel having ϵ^2 variance. We analyze the mean square error of the estimator and the overall complexity of the algorithm. The MLCV estimator of the QoI and its corresponding MSE bound are given by:

$$\hat{Q}^{MLCV}_{h,\{M_l\}} = \sum_{l=0}^{L} \frac{1}{M_l} \sum_{i=1}^{M_l} \left(Q^i_{h_l} - Q^i_{h_{l-1}} - (Q^{\epsilon,i}_{h_l} - Q^{\epsilon,i}_{h_{l-1}}) \right) + \mathbb{E}[Q^{\epsilon,SC}_{h_L}],$$

$$e(\hat{Q}^{MLCV}_{h_L,\{M_l\}})^2 \leqslant \sum_{l=0}^{L} \frac{\mathbb{V}\mathrm{ar}(Y^{CV}_{h_l})}{M_l} + 2\mathbb{E}[Q^\epsilon_{h_L} - Q^{\epsilon,SC}_{h_L}]^2 + 2\mathbb{E}[Q_{h_L} - Q]^2.$$

We also propose possible choices of the regularization parameter and of the number of samples per grid so as to equilibrate the space discretization error, the statistical error and the error in the computation of the expected value of the control variate by Stochastic Collocation. Numerical examples demonstrate the effectiveness of the method. A comparison with the standard Multi Level Monte Carlo method is also presented for different choices of the covariance function of the input field.

References:

[1] A.L. Teckentrup, R. Scheichl, M.B. Giles, E. Ullmann. *Further analysis of multilevel Monte Carlo methods for elliptic PDEs with random coefficients*. Numerische Mathematik 125: 569-600, 2013

[2] A. Barth, C. Schwab, N. Zollinger, *Multi-level Monte Carlo finite element method for elliptic PDEs with stochastic coefficients*. Numerische Mathematik, 119(1):123-161, 2011

[3] I. Babuška, F. Nobile, R. Tempone, *A stochastic collocation method for elliptic partial differential equations with random input data*, SIAM REVIEW Vol. 52, No. 2, pp. 317-355

[4] J. Charrier, R. Scheichl, A. Teckentrup, *Finite element error analysis of elliptic PDEs with random coefficients and its application to multilevel Monte Carlo methods*, SIAM Journal on Numerical Analysis (SINUM), 51 (1), pp. 322-352, 2013

Short biography – Currently I am a PhD student at EPFL in the Chair of scientific computing and uncertainty quantification. My research is devoted to the development of efficient numerical methods for flow and transport phenomena in heterogeneous random porous media. This project is funded by the Fonds National Suisse (FNS). I got my bachelor and master diplomas at Politecnico di Milano in Mathematical Engineering.

Getting into the failure domain: application in estimation of extreme quantiles and probabilities

C. Walter
Universit Paris Diderot - Paris 7

Affiliation: *Commissariat à l'énergie atomique et aux énergies alternatives*, Bruyère-le-Châtel, 91297 Arpajon Cedex, France

Email: clement.walter@cea.fr – **URL**: http://www.lpma.univ-paris-diderot.fr

Master: Mines ParisTech, Paris, France

Ph.D. (2013-2016): Universit Paris Diderot - Paris 7

Supervisor(s): Prof. Josselin Garnier (Universit Paris Diderot - Paris 7) and Gilles Defaux (CEA)

Abstract:

In the context of reliability analysis, complex models representing physical situations are used to determine failure domain via a given threshold not to be overpassed. These models are often very time consuming and no analytical expression is available. Thus, the evaluation of a failure probability given a threshold or the estimation of a quantile given a targeted probability cannot be done by usual computation tools. Some developments have been done in two different directions : first the use of sequential Monte-Carlo methods allows to decrease the number of necessary calls to the model ([1], [2]) ; in this mood Guyader *et al.* [3] have shown that the optimal algorithm works with limit subsets, *ie.* fixing the current threshold at the minimum of the working population. Another approach is to use the given computational budget to fit a surrogate model and then to use it instead of the real model to evaluate quantities of interest (see [4], [5] or [6] for examples). Both approaches suffer from several limitations.

Sequential Monte-Carlo would still require quite a lot of calls to the limit-state function, while fitted surrogate models can be very far from the original ones without any possibility to control precision. However, in the context of probability or quantile estimation, only the boundary delimiting failure and safety domains is to be well approximated ; especially, no good approximation of the model in both domains is necessary if data points are well classified.

The novelty of our work comes from our understanding of these two approaches in the rare event simulation context as a *run to the failure domain*. This means that sequential Monte-Carlo appears indeed like a move of particles from an initial random state to the failure domain and that meta-modelling needs only pairs of points on each side of the boundary. The basic result is that with a sequential sampling strategy, the theoretical number of samples needed to get into a failure domain of probability measure p follows a Poisson law with parameter $\log(1/p)$. This number is to be compared with a classical Monte-Carlo sampling with a theoretical $1/p$ number of samples needed.

A consequence for sequential Monte-Carlo algorithms is the theoretically full parallelisation of Guyader *et al.* ([3]) optimal algorithm for rare event probability estimation with the limit case of algorithms with only 1 particle. On the other hand quantile estimation requires a small adaptation and we thus present a modified algorithm which enables full parallelisation. This modification is of very little effect as in the test cases it increases the total number of calls by approximately 10% while parallelisation can reduce computational budget by a factor up to population size (*eg.* 100, 1000...). However because of robustness issue in the Metropolis-Hastings generation step, it appears that population size for one algorithm shouldn't be smaller than 10 to 20.

On the other hand we use this *run to the failure domain* approach to derive a new strategy for first Design of Experiment (DoE) in the case of meta-modelling. In fact, it is well noticed that surrogate models require failing data points in their learning database to perform well ([5]). Unlike Space-filling strategies which aim at giving an overall knowledge of the model or Gaussian sampling techniques which "overlearn" the model in the safety domain, we propose to use a sequential strategy to get into the failure

domain quickly. More precisely, by replacing the expensive-to-evaluate model by the cheap surrogate in the Metropolis-Hastings step, we get similar results in terms of number of calls before the failure domain while limiting drastically the total number of calls usually driven by a *burn-in* parameter. Finally, we get a new relation for the size of the first DoE, it is : $N_{DoE} \approx d + 1 + N \log(1/p)$ with N the number of willing failing points in the DoE and d the dimension of the input space. Practically, it appears that given N and d, total number of calls is indeed lower than this theoretical value.

References:

[1] F Cérou, P Del Moral, T Furon, and A Guyader. Sequential monte carlo for rare event estimation. *Statistics and Computing*, 22(3):795–808, 2012.

[2] S-K Au and J L Beck. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*, 16(4):263–277, 2001.

[3] A Guyader, N Hengartner, and E Matzner-Løber. Simulation and estimation of extreme quantiles and extreme probabilities. *Applied Mathematics & Optimization*, 64(2):171–196, 2011.

[4] J-M Bourinet, F Deheeger, and M Lemaire. Assessing small failure probabilities by combined subset simulation and support vector machines. *Structural Safety*, 33(6):343–353, 2011.

[5] B Echard, N Gayton, and M Lemaire. Ak-mcs: An active learning reliability method combining kriging and monte carlo simulation. *Structural Safety*, 33(2):145–154, 2011.

[6] V Dubourg, F Deheeger, and B Sudret. Metamodel-based importance sampling for the simulation of rare events. *arXiv preprint arXiv:1104.3476*, 2011.

Short biography – I am Clément Walter, 24 years old. I just graduated from Mines ParisTech which I integrated after 2 years of prep classes, section Math and Physics. At Mines Paristech I specialised in geostatistics and used it for coal resources evaluation and drill hole campaign planing. For my 4 months final internship I started working at CEA on rare event simulation and got the opportunity to continue here in a PhD, which I officially started in November.

Goal-oriented low-rank approximations for high dimensional stochastic problems

O. Zahm
École Centrale de Nantes

Affiliation: *GeM (UMR CNRS 6183)*, École Centrale de Nantes, 1 rue de la Noë, 44321 Nantes, France.

Email: olivier.zahm@ec-nantes.fr – **URL**: http://gem.ec-nantes.fr/

Master: École Normale Supérieure de Cachan

Ph.D. (2011-2015): École Centrale de Nantes

Supervisor(s): A. Nouy and M. Billaud-Friess (École Centrale de Nantes)

Abstract:

Due to the need of more realistic numerical simulations, models presenting uncertainties are receiving a growing interest. Various numerical methods attempt to quantify the probabilistic response of some physical phenomena, often modeled by partial differential equations with random coefficients. Here we adopt a functional point of view of the uncertainty, meaning that we look for an approximation of the solution that is a function of the random coefficients (seen as new variables). Low-rank tensor methods appear as an efficient way to solve the resulting high dimensional problems, and can also be interpreted as model reduction (see recent surveys [2, 6, 5]).

Different strategies have been proposed for the solution of equations in tensor format. One can obtain an approximation of the solution in low-rank tensor subsets through the direct minimization of some residual norm. Using this approximation, estimation of quantities of interest (expectation, variance or sensitivity indices) can be computed. However, there is no guaranty that classical minimal residual formulations provide accurate reduced order models for the estimation of quantities of interest.

The basic idea of the present work (inspired from [3, 4]) is to introduce an ideal minimal residual formulation such that the optimality of the approximation is achieved with respect to a specified norm. We propose and analyze in [1] an algorithm that provide a quasi-optimal low-rank approximation of the solution defined by the ideal formulation.

This new approach offers the possibility to choose a norm so that the optimality of the approximation is achieved with respect to some quantity of interest. We investigate different constructions of such "goal-oriented norm" in the case of linear and quadratic quantities of interest. The resulting method can be seen as an optimal goal-oriented model reduction method.

References:

[1] M. Billaud-Friess, A. Nouy and O. Zahm. *A tensor approximation method based on ideal minimal residual formulations for the solution of high dimensional problems*, in preparation, 2013.

[2] F. Chinesta, P. Ladeveze, and E. Cueto. *A short review on model order reduction based on proper generalized decomposition*, Archives of Computational Methods in Engineering, 18(4):395–404, 2011.

[3] A. Cohen, W. Dahmen and G. Welper *Adaptivity and Variational Stabilization for Convection-Diffusion Equations*, ESAIM: Mathematical Modelling and Numerical Analysis, 2012.

[4] Dahmen W., Huang C. and Schwab C. *Adaptative petrov-galerkin methods for first order transport equations*, IGPM Report 321, RWTH Aachen, 425–467, 2011.

[5] L. Grasedyck, D. Kressner, and C. Tobler. *A literature survey of low-rank tensor approximation techniques.* GAMM-Mitteilungen, 2013.

[6] W. Hackbusch. *Tensor Spaces and Numerical Tensor Calculus*, Series in Computational Mathematics volume 42. Springer, 2012.

Short biography – After a master in advanced numerical methods for computational mechanics at the École Normale Supérieure de Cachan, I started my PhD in 2011 in the GeM laboratory at the École Centrale de Nantes (http://gem.ec-nantes.fr/). I am also a teaching assistant for a course of applied mathematics at the third-year Bachelor level (numerical analysis, optimization, probability and statistics). My PhD is funded by the French ministry of research.

Wednesday, April 23rd, 2014

08:30 **Registration**

09:15 **Opening talk - B. Sudret & D. Ginsbourger**

09:30 **Opening lecture - C. Schillings** (ETH Zürich). *Exploiting sparsity in Bayesian inverse problems of parametric operator equations*

10:30 **Coffee break**

10:50 **S. Sukys** (ETH Zürich). *Multi-level Monte Carlo finite volume methods for stochastic systems of hyperbolic conservation laws*

11:20 **D. Jacquemart** (University Rennes I). *Adaptive particle methods for rare event simulation in a Markovian framework*

11:50 **J. Nagel** (ETH Zürich). *Bayesian multilevel model calibration for inversion of "perfect" data in the presence of uncertainty*

12:20 **Lunch Break**

13:30 ***PhD Students - Posters session***

15:00 **A. Ahidar** (Université Paul Sabatier). *Multivariate quantile surfaces and application to an aircraft problem*

15:30 **R. Fischer** (Université Paris-Est). *Modeling dependence under constraint*

16:00 **Coffee Break**

16:20 **F. Zertuche** (Université Joseph Fourier). *Multi-fidelity regression using a non-parametric relationship*

16:50 **S. Nanty** (Université Joseph Fourier). *Uncertainty quantification and visualization for functional random variables*

Thursday, April 24th, 2014

08:30 **Registration**

09:00 **Conference opening -** W. Kröger (ETH Risk Center)

09:15 **C. Prieur** (Université Joseph Fourier). *MascotNum news*

09:30 **F. Nobile** (EPF Lausanne). *Discrete least-square polynomial approximations for high dimensional uncertainty propagation*

10:30 **Coffee break**

10:50 **J.P. Vert** (Institut Curie). *Machine Learning for personalised genomics*

11:50 **A. Krause** (ETH Zürich). *Focusing exploration with confidence*

12:50 **Lunch Break**

14:15 **F.S. Koutsourelakis** (TUM). *Simulation-based, highdimensional stochastic optimization: application in robust topology optimization under large material uncertainties*

15:15 **S. Marelli** (ETH Zürich). *UQLab: A framework for uncertainty Quantification in Matlab*

16:00 **Coffee Break**

16:20 **D. Ginsbourger** (University of Bern). *Incorporating structural priors in Gaussian random field models*

Friday, April 25th, 2014

09:00 **P. Koumoutsakos** (ETH Zürich). *Data driven, molecular dynamics for nanoscale fluid mechanic*

10:00 **C. Chevalier** (University of Bern)**, Y. Richet** (IRSN)**, O.Roustant** (EMSE)**.** *ReDICE Consortium*

10:30 **Coffee break**

10:50 **P. Renard** (Université de Neuchâtel). *Multiple-point statistics as a tool to assess complex spatial uncertainty*

11:50 **J. Irving** (University of Lausanne). *Stochastic inverse methods for near-surface geophysical problems*

12:50 **Lunch Break**

14:15 **J. Wiart** (Orange Labs). *Contribution of Statistics to the numerical assessment of the electromagnetic fields human exposure.*

15:15 **Closure**

MascotNum 2014 Website:

http://www.ibk.ethz.ch/su/mascotnum2014/