SIAM 2022

Searchable Abstracts Document

Conference on Uncertainty Quantification

April 12–15, 2022

This document was current as of April 4, 2022. Abstracts appear as submitted.

SIAM
3600 Market Street, 6th Floor
Philadelphia, PA 19104-2688 U.S.
Telephone: 800-447-7426 (U.S. & Canada) +1-215-382-9800 (Worldwide)
meetings@siam.org

IP1

Opening Remarks and Presentation: Transport Methods for Stochastic Modeling and Inference

Transportation of measure underlies many powerful tools for Bayesian inference, density estimation, and stochastic modeling. A central idea is to deterministically couple a probability measure of interest with a tractable reference measure (e.g., a standard Gaussian). Such couplings are induced by transport maps, and enable direct simulation from the desired measure simply by evaluating the transport map at samples from the reference. In recent years, an enormous variety of representations and constructions for such transport maps have been proposedranging from monotone polynomials, invertible neural networks, and normalizing flows to the flows of ODEs. Within this framework, one can describe many useful notions of low-dimensional structure: for instance, sparse or decomposable transports underpin modeling and computation with non-Gaussian Markov random fields, and low-rank transports arise frequently in inverse problems. I will present an overview of this framework, and then focus on new developments for nonlinear ensemble filtering and likelihood-free inference (LFI). Some of associated algorithms can be understood as the natural nonlinear generalization of the ensemble Kalman filter. Motivated by broader applications in LFI and generative modeling, I will also discuss methods for estimating monotone triangular maps from few samples, and joint dimension reduction of parameters and data in inference applications.

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

IP2

Machine Learning in Data Assimilation and Inverse Problems

Machine learning is emerging as an essential tool in many science and engineering domains, fueled by extraordinarily powerful computers as well as advanced instruments capable of collecting high-resolution and high-dimensional experimental data. However, using off-the-shelf machine learning methods for analyzing scientific and engineering data fails to leverage our vast, collective (albeit partial) understanding of the underlying physical phenomenon or models of sensor systems. Reconstructing physical phenomena from indirect scientific observations is at the heart of scientific measurement and discovery, and so a pervasive challenge is to develop new methodologies capable of combining such physical models with training data to yield more rapid, accurate inferences. We will explore these ideas in the context of inverse problems and data assimilation challenges; examples include climate forecasting, uncovering material structure and properties, and medical image reconstruction. Classical approaches to such inverse problems and data assimilation approaches have relied upon insights from optimization, signal processing, and the careful exploitation of forward models. In this talk, we will see how these insights and tools can be integrated into machine learning systems to yield novel methods with significant accuracy and computational advantages over nave applications of machine learning.

Rebecca Willett
University of Chicago
willett@uchicago.edu

IP3

Learning Physics-based Models from Data: Perspectives from Model Reduction

Operator Inference is a method for learning predictive reduced-order models from data. The method targets the derivation of a reduced-order model of an expensive high-fidelity simulator that solves known governing equations. Rather than learn a generic approximation with weak enforcement of the physics, we learn low-dimensional operators whose structure is defined by the physical problem being modeled. These reduced operators are determined by solving a linear least squares problem, making Operator Inference scalable to high-dimensional problems. The method is entirely non-intrusive, meaning that it requires simulation snapshot data but does not require access to or modification of the high-fidelity source code. For problems where the complexity of the physics does not admit a global low-rank structure, we construct a nonlinear approximation space. This is achieved via clustering to obtain localized Operator Inference models, or by approximation in a quadratic manifold. The methodology is demonstrated on challenging large-scale problems arising in rocket combustion and materials phase-field applications.

Karen E. Willcox
UT Austin
kwillcox@oden.utexas.edu

IP4

Remarks and Presentation: A New Age of Climate Modeling: Evolving Risk Assessment in the Post-Paris World

The infrastructure and underlying statistical assumptions which informed the assessment of future climate impacts, and how they relate to policy, remained fundamentally unchanged for several decades. However, the period after the Paris Agreement in 2016 has seen multiple changes to the landscape which have altered how climate model output is interpreted, and shifted what is urgently needed from the climate modeling community. Firstly, emerging climate signals have enabled an increasing capacity for constraining response from the multi-model ensemble of state of art models. Second, rapidly evolving climate policy has shifted the plausible scenario landscape, highlighting aspects of model uncertain response which received little attention in the past. Thirdly, a new generation of emulation tools have provided a mechanism for rapid scenario testing and uncertainty quantification, in some cases replacing the ensemble of opportunity of complex models as the primary source of information in assessment. These changes carry both new opportunities and capacity for rapid policy impact assessment, but also risks a systematic underestimation of tail risks. In this talk, we discuss the rapidly evolving role of statistics in climate assessment, and urgent challenges to be addressed by the uncertainty quantification community in a critical decade for climate action.

Benjamin Sanderson
Centre for International Climate and Environmental Research
(CICERO), Oslo, Norway

sanderson@cerfacs.fr

IP5

Uncertainty Quantification for Remote Sensing Data

Remote sensing data are the premier source of information available today for studying global environmental and climate phenomena. Disruptions brought about by changes in the climate system on both short- and long-term scales demand that we exploit this rich data source, but its complexity, heterogeneity, and massive size present challenges. Uncertainty quantification (UQ) is a key to overcoming these issues because probabilistically defined uncertainty is a common currency for understanding relationships between data and unknown, underlying truth. In remote sensing today there are a raft of approaches to assigning uncertainties, ranging from merely reporting empirically derived biases and variances over specific spatio-temporal domains, to full probabilistic characterization through inverse UQ methods such as MCMC. There are no generally agreed-upon methods or protocols, and the user community remains unsure of how to make use of uncertainty information. More broadly, there are deep philosophical questions about how to frame UQ in the remote sensing context. In this talk I will discuss this framing, and use it as a vehicle to identify leading research problems in the development of uncertainty quantification methods for remote sensing data and analysis.

Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology
Amy.Braverman@jpl.nasa.gov

IP6

Closing Remarks and Presentation: Structural Uncertainty Matters: How Strong Modelling Assumptions Misled on the Role of Lockdowns in the UK Covid-19 Epidemic

Prominent studies from early (e.g. Flaxman et al. 2020, Nature 584, 257-261) and later (e.g. Knock et al. 2021, Science Translational Medicine) in the English Covid-19 epidemic conclude that full lockdowns were essential for epidemic control. The studies used modern Bayesian computing technology for parametric inference about epidemic models, from a range of clinical and surveillance data. Such models can be impressively rich in epidemic process detail, but critical dynamic components are constrained to follow simple parametric sub-models. These sub models are often over-constrained, overstating structural knowledge and understating uncertainty. In particular, simple ad hoc parametric representations of the time course of the pathogen reproductive number, R, or of contact rates, almost completely drive the conclusions about English lockdown efficacy in the above studies. Replacing these restrictive sub models with data driven spline based alternatives, leads to the conclusion that infections were in decline and R¡1 before each English lockdown (Wood, 2021, Biometrics; Wood and Wit, 2021, PLOS1). This confirms a May 2020 Bayesian deconvolution analysis estimating daily new infections (incidence) from daily death data and direct incidence estimates from the major statistical surveys of UK Covid prevalence.

Simon Wood
Mathematical Sciences
University of Bath
simon.wood@ed.ac.uk

IP7

Title Tbd - Lakshminarayanan

Balaji Lakshminarayanan
Google Brain
balaji@gatsby.ucl.ac.uk; balajiln@google.com

SP1

Remarks and Presentation: SIAG/Uncertainty Quantification Early Career Prize Lecture - Scalable Gaussian Process for Computer Model Emulation and Uncertainty Quantification

Computer experiments are ubiquitously used in science and engineering, yet the computational complexity can prohibit its applications for many large-scale systems. In this talk, we will introduce the Gaussian process emulator as a surrogate model for approximating computer model simulations and uncertainty quantification. The new results concern emulating computer models with massive coordinates, high-dimensional inputs and functionals. We will review the stochastic differential equation representation of Gaussian processes with Matern covariance, and applications of Kalman filter, as exact, computationally efficient alternatives. Our extensions include the generalized probabilistic principal component analysis and Gaussian orthogonal latent factor processes, for modeling incomplete matrix observations of correlated data. We have developed software packages in R and MATLAB that implement the Gaussian process emulator and some of the computationally scalable alternatives. Applications include emulating the TITAN2D model of pyroclastic flows, ground deformation simulation by COMSOL Multiphysics, ab initio molecular dynamics simulations, and cellular migration simulations.

Mengyang Gu
Department of Statistics and Applied Probability
University of California, Santa Barbara
mengyang@pstat.ucsb.edu

CP1

Robust Prediction Interval Estimation for Gaussian Processes by Cross-Validation Method

Gaussian Processes are considered as one of the most important Bayesian Machine Learning methods. They typically use the Maximum Likelihood Estimation or Cross-Validation to fit parameters. Unfortunately, these methods may give advantage to the solutions that fit observations in average, but they do not pay attention to the coverage and the width of Prediction Intervals. This may be inadmissible, especially for systems that require risk management. Indeed, an interval is crucial and offers valuable information that helps for better management than just predicting a single value. In this work [N. Acharki et al., Robust Prediction Interval estimation for Gaussian Processes by Cross-Validation method, arXiv : 2106.05396], we address the question of adjusting and calibrating Prediction Intervals for Gaussian Processes Regression. First, we determine the model's parameters by a standard Cross-Validation or Maximum Likelihood Estimation method then we adjust the parameters to assess the optimal type II Coverage Probability to a nominal level. We apply a relaxation method to choose parameters that minimize the Wasserstein distance between the Gaussian distribution of

the initial parameters (Cross-Validation or Maximum Likelihood Estimation) and the proposed Gaussian distribution among the set of parameters that achieved the desired Coverage Probability.

Naoufal Acharki
Ecole Polytechnique
TotalEnergies SE
naoufal.acharki@polytechnique.edu

Josselin Garnier
Ecole Polytechnique
josselin.garnier@polytechnique.edu

Antoine Bertoncello
TotalEnergies SE
antoine.bertoncello@totalenergies.com

CP1

Multi-Fidelity Gaussian Process Regression for High-Dimensional Code Outputs Using Wavelet Transform Covariance

Gaussian process regression is widely used to emulate the output of an expensive computer code. Our work focuses on 1) multifidelity computer codes, which can be run with different levels of accuracy and cost and 2) high-dimensional code outputs, typically time series. We propose to extend the auto-regressive co-kriging model originally proposed by Kennedy-O'Hagan to high-dimensional outputs by choosing an appropriate covariance structure in the space of wavelet coefficients. We show that the resulting surrogate model performs better in terms of prediction errors and uncertainty quantification than standard dimension reduction techniques. This talk extends the recent preprint [arXiv:2109.11374].

Baptiste Kerleguer
CEA, DAM, DIF
CMAP, Ecole polytechnique
baptiste.kerleguer@polytechnique.edu

Claire Cannamela
CEA, DAM, DIF
F-91297 Arpajon, France
claire.cannamela@cea.fr

Josselin Garnier
Ecole Polytechnique
josselin.garnier@polytechnique.edu

CP1

Emulating Computationally Expensive Dynamical Simulators Using Gaussian Processes

A Gaussian process (GP)-based methodology is proposed to emulate computationally expensive dynamical computer models or simulators. The method relies on emulating the short-time numerical flow map of the model. The flow map returns the solution of a dynamic system at an arbitrary time for a given initial condition. The prediction of the flow map is performed via a GP whose kernel is estimated using random Fourier features. This gives a distribution over the flow map such that each realisation serves as an approximation to the flow map. A realisation is then employed in an iterative manner to perform one-step ahead predictions and forecast the whole time series. Repeating this procedure with multiple draws from the emulated flow map provides a probability distribution over the time series. The mean and variance of that distribution are used as the model output prediction and a measure of the associated uncertainty, respectively. The proposed method is used to emulate several dynamic non-linear simulators including the well-known Lorenz attractor and van der Pol oscillator. The results show that our approach has a high prediction performance in emulating such systems with an accurate representation of the prediction uncertainty.

Hossein Mohammadi
University of Exeter
h.mohammadi@exeter.ac.uk

Peter Challenor
College of Engineering, Mathematics and Physical Sciences
University of Exeter
p.g.challenor@exeter.ac.uk

CP1

Automatic Covariance Structure Discovery for Gaussian Processes Though MCMC Sampling of Kernel Algebra

A recurring problem in Gaussian Process (GP) models is the specification of an appropriate covariance kernel. Current designs are dominated by pre-existing generic forms and as such, even with expert knowledge, may produce sub-optimal fits. As new kernels can be synthesised from existing ones, for any given modelling problem there is an infinite kernel space and associated optimality distribution which can be learned. This allows for uncertainty quantification (UQ) over the kernel space, with significant implications for GP inference. A novel MCMC scheme is proposed that operates over kernel algebra to sample its posterior distribution, automatically discovering novel covariance structures. This approach constructs complex kernels without requiring Deep GP modelling techniques, which may significantly prevent loss of interpretability. Furthermore, the sampler has the potential to discover approximations or equivalencies to deep GP models, fostering new discussions on hierarchical modelling practices. A complete workflow inclusive of proposal construction, control priors and convergence metrics is presented. Sampling is verified exhaustively over subsets of kernel space for case studies up to 10 dimensions. Modal kernels from the posterior are compared against competing kernel designs from statistics and machine learning fields, including deep GP models and the Automatic Statistician greedy search methodology an important precursor to our MCMC development.

Christopher Parton-Fenton, Daniel Williamson
University of Exeter
cf502@exeter.ac.uk, d.williamson@exeter.ac.uk

Derek Bingham
Simon Fraser University
derek_bingham@sfu.ca

CP1

Mixed-Variable Latent Variable Gaussian Process Modeling for Adaptive Learning and Bayesian Optimization

Real engineering applications of surrogate modeling often involve both discrete (categorical) and continuous inputs. Existing machine learning (ML) models that can handle mixed variables as inputs require a large amount of data

but do not provide uncertainty quantification that is crucial for sequential (adaptive) design of experiments. We have developed a novel Latent Variable Gaussian Process (LVGP) based ML approach that involves a latent variable (LV) representation of qualitative inputs, and automatically discovers a categorical-to-numerical nonlinear map that transforms the underlying high dimensional physical attributes into the LV space. The nonlinear mapping also provides an inherent ordering and structure for the levels of the qualitative factor(s), which leads to substantial insight and explainable ML. Our LVGP approach inherits all advantages of GP methods. Specifically, LVGP provides uncertainty quantification of prediction which is critical for adaptive sampling to sequentially choose samples based on current observations and the method also offers easy integration with Bayesian optimization (BO) or other reinforcement learning strategies for the purpose of design optimization. The benefits of our method for surrogate modeling, adaptive learning, and Bayesian Optimization will be demonstrated through examples in design of engineered materials systems.

Wei Chen, Anton Van Beek
Northwestern University
weichen@northwestern.edu, antonvanbeek2022@u.northwstern.edu

CP1

Representation of Stochastic Processes via Coupling of State-Dependent Random Fields with Time-Dependent Stochastic Processes of State

In this paper, a methodology is presented for the multi-fidelity modeling of stochastic processes through the coupling of dependencies of random responses on state and the state on time. For multi-fidelity modeling, phenomenologically similar states have to be evaluated together and the feasibility of efficient numerical approximation of highly dynamic transient behavior has to be ensured. For this purpose, the approximation of non-Gaussian random fields of Quantities of Interest is conducted using Karhunen-Love Expansions (KLE) on a discrete state space. In addition, the stochastic process of state with respect to time is also approximated using KLE. The random variables within the KLE are represented by Polynomial Chaos Expansions (PCE), thus creating the connection between the KLE and the physical uncertain parameters. By substitution of the stochastic state, the state-dependent random fields describing Quantities of Interest along with the time-dependent stochastic process of state, yield the time-dependent stochastic process of Quantities of Interest. The developed methodology is applied for the assessment of uncertainties within a buoyancy-driven mixing process inside the Differentially Heated Cavity (DHC) of aspect ratio 4. It is found that stochastic processes of Quantities of Interest like the Nusselt-number are well represented through this methodology.

Philipp J. Wenig
Bundeswehr University Munich
Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany
philipp.wenig@unibw.de

Stephan Kelm
Forschungszentrum Jülich GmbH
52425 Jülich, Germany
s.kelm@fz-juelich.de

Markus Klein
Bundeswehr University Munich
Werner-Heisenberg-Weg 39, 85577 Neubiberg, Germany
markus.klein@unibw.de

CP2

Estimating Time-Varying Parameters via Computational Filtering

Estimating and quantifying uncertainty in unknown system parameters remains a big challenge in many scientific applications. In addition to static parameters, systems may rely on parameters that vary with time but have unknown (or uncertain) dynamics and/or cannot be directly measured. This talk will address new approaches using computational filtering to address time-varying parameter estimation inverse problems in deterministic dynamical systems, with emphasis on how uncertainty in the parameter estimates affects the corresponding model predictions. Results will be demonstrated with numerical examples from nonlinear dynamical systems.

Andrea Arnold
Worcester Polytechnic Institute
anarnold@wpi.edu

CP2

Machine Learning-Based Ensemble Kalman Filters for Joint Learning of States and Dynamics

Learning both latent dynamics and states from time-evolving high-dimensional observation data is a challenging problem in the fields of data assimilation, time series modeling and uncertainty quantification. The ensemble Kalman filters (EnKFs) have been developed as a powerful tool to infer latent states when one has full knowledge about the latent dynamics. However, in real world applications one may only have partial or zero knowledge about the latent dynamics. To tackle this issue, we propose *auto-differentiable ensemble Kalman filters* (AD-EnKFs), where the unknown parameters of the latent dynamics can be identified through gradient-based maximum likelihood estimation, and the data log-likelihood can be estimated as a by-product of EnKF. In doing so, AD-EnKFs leverage the ability of ensemble Kalman filters to scale to high-dimensional latent states and the power of automatic differentiation to train high-dimensional surrogate models for the latent dynamics. Numerical results on chaotic systems like Lorenz-63 and Lorenz-96 show that AD-EnKFs outperform existing methods that use expectation-maximization or particle filters to learn both latent dynamics and states. Moreover, AD-EnKFs can be applied to learn data-driven reduced order models, where the dimension of latent states can be far less than the dimension of observation data.

Yuming Chen, Daniel Sanz-Alonso, Rebecca Willett
University of Chicago
ymchen@uchicago.edu, sanzalonso@uchicago.edu, willett@uchicago.edu

CP2

Adaptive Importance Sampling for Seismic Fragility Curves Estimation

As part of Probabilistic Risk Assessment studies, it is necessary to study the fragility of mechanical and civil engineered structures when subjected to seismic loads. This risk can be measured with fragility curves, which express the probability of failure of the structure conditionally to a seismic intensity measure. The estimation of fragility

curves relies on time-consuming numerical simulations, so that careful experimental design is required in order to gain the maximum information on the structures fragility with a limited number of code evaluations. We propose and implement an active learning methodology based on adaptive importance sampling in order to reduce the variance of the training loss. Numerical simulations of an elasto-plastic oscillator and an industrial water piping system have been performed in order to validate our active learning procedure. Theoretical guarantees of convergence have also been assessed. A preprint is available on Arxiv (https://arxiv.org/abs/2109.04323).

Clement Gauchy
CEA, Service d'Etudes Mecaniques et Thermiques
clement.gauchy@polytechnique.edu

CP2

UQ Explorations: Atmospheric Temperature and Humidity Profile Retrievals from High-Resolution Infrared Radiation Sounder

A long-term atmospheric temperature and humidity profile dataset based on high-resolution infrared radiation sounder (HIRS) observations from 1979-present has been developed for climate applications. The remote sensing retrieval algorithm is based on a neural network scheme where the networks are trained with emulated satellite sounder measurements from radiative transfer model simulations and the input atmospheric profiles. Previous work has demonstrated the stability and accuracy of the multi-satellite time series for the more than a dozen satellites constituting the 40+ year time series. Independent observation platforms, including other satellite retrievals as well as radiosondes, have validated the approach. To further improve the utility of the dataset for the user community, we are developing associated uncertainty quantification (UQ) measures. UQ methods that may be applied to deep learning algorithms are abundant in the literature. For this application, we explore several candidate methodologies: bootstrap model, Monte Carlo dropout, and deep ensembles. Evaluation of the candidate UQ approaches are made based on the ability of the method to correctly estimate prediction interval coverages.

Jessica L. Matthews
North Carolina State University
jessica.matthews@noaa.gov

Lei Shi
NOAA
lei.shi@noaa.gov

Yuhan Rao
North Carolina Institute for Climate Studies
North Carolina State University
yrao5@ncsu.edu

CP2

Addressing Model Inadequacy in Inverse Uncertainty Quantification for Fuel Performance Modeling by using a Hierarchical Statistical Formulation

The nuclear industry uses fuel performance modeling to demonstrate integrity preservation of fuel rods. These codes include a complex system of models with many empirical constants that one must calibrate so that the output of the code agrees with measurement data. This process is called code calibration and has traditionally included fitting best-estimate and conservative models. Lately, the focus has increased on improved inverse-uncertainty quantification (IUQ) methods to replace over-conservative models developed in the past. However, IUQ is challenging for fuel modeling for several reasons, e.g., many models approximate the physics, many nuisance parameters are present, etc. This study investigates a method that addresses unknown uncertainties in code calibration by adapting the covariance matrix of the model parameters so that the propagated uncertainty conforms with the spread of the residuals. We do this by assuming a hierarchical model and a Gaussian variability in the calibration parameters. We then use a marginalized formulation proportional to the posterior density of the hyper-parameters (mean and covariance) and sample this using Markov Chain Monte Carlo. The is applied to various synthetic test-beds relevant to fuel performance models (e.g., cladding oxidation, fission gas release, etc.) to demonstrate its applicability.

Gustav Robertson, Henrik Sjöstrand, Peter Andersson
Uppsala University
gustav.robertson@gmail.com,
henrik.sjostrand@physics.uu.se,
peter.andersson@physics.uu.se

Paul Blair
Westinghouse Electric Sweden, Västerås, Sweden
blairpr@westinghouse.com

CP3

Prior Covariance Constraints for Bayesian Stochastic Inversion of Computer Models

Stochastic inversion problems are typically encountered when it is wanted to quantify the uncertainty affecting the inputs of computer models. They consist in estimating input distributions from noisy, observable outputs, and such problems are increasingly examined in Bayesian contexts where the targeted inputs are affected by a mixture of aleatory and epistemic uncertainties. While they are characterized by identifiability conditions, constraints of "signal to noise" have to be took into account within the definition of the model, prior to inference. In addition to numeric conditioning notions and regularization techniques used in inverse problems, this article proposes and investigates an interpretation of a meaningful solution, in the context of parametric uncertainty quantification and global sensitivity analysis, based on the degradation of uncertainty-related indicators (variance, entropy, Fisher information), which can be translated by constraints on the input covariance. Such prior constraints can be made explicit considering linear or linearizable operators. Simulated experiments indicate that, when injected into the modeling process, these constraints can limit the influence of measurement or process noise on the estimation of the input distribution, and let hope for future extensions in a full non-linear framework, for example through the use of linear Gaussian mixtures.

Nicolas Bousquet
EDF & Sorbonne Université
nicolas.bousquet@edf.fr

CP3

A Sur Version of the Bichon Criterion for Inversion

Nowadays, many inversion issues are present in industry. In these problems, the goal is to find all sets of parameters such that a quantity of interest respects a given constraint,

for example remains below a threshold. In the field of floating wind for instance, a pre-calibration step consists in estimating model parameters that fit the measured data (e.g., accelerations), with a given accuracy. An effective way to solve this problem is to use Gaussian process regression with a Design of Experiment (DoE), sequentially enriched by an inversion-adapted acquisition criterion such as Bichon. A class of enrichment criteria is the SUR (Stepwise Uncertainty Reduction) class, which quantify the uncertainty reduction that can be achieved by the addition of a new point. The goal of this work is to propose a SUR version of the Bichon criterion. The proposed SUR Bichon strategy is defined by integrating the Bichon criterion on the design space. Performances of the SUR Bichon criterion are compared to some state-of-the-art criteria, on common test functions. The SUR Bichon criterion shows encouraging results compared to other conventional DoE enrichment criteria, like the SUR Vorob'ev criterion. The prospects for this work are adapting this criterion to more complex data like functional uncertain input variables.

Clément Duhamel
Université Grenoble Alpes
clement.duhamel@inria.fr

Clémentine Prieur
Grenoble Alpes University
Jean Kunzmann Lab, INRIA project/team AIRSEA
clementine.prieur@univ-grenoble-alpes.fr

Céline Helbert
Institut Camille Jordan Ecole Centrale Lyon
celine.helbert@ec-lyon.fr

Miguel Munoz Zuniga
IFP Energies Nouvelles
miguel.munoz-zuniga@ifpen.fr

Delphine Sinoquet
IFP Energies nouvelles
delphine.sinoquet@ifpen.fr

CP3

Estimating Hyperparameters in Hierarchical Bayesian Linear Inverse Problems

We consider a hierarchical formulation of the Bayesian inverse problems in which the forward operator is assumed to be linear, and we treat unknown parameters associated with the noise and the prior as hyperparameters. We marginalize over the unknown parameters and compute the maximum a posteriori estimator over the marginalized distribution. We use Krylov subspace methods to accelerate the computations of the nonlinear optimization technique. We demonstrate the performance of our approach on several synthetic test problems.

Khalil A. Hall-Hooper, Arvind Saibaba
North Carolina State University
kahallho@ncsu.edu, asaibab@ncsu.edu

Julianne Chung
Department of Mathematics
Virginia Tech
jmchung@vt.edu

Scot Miller
The Johns Hopkins University
scot.m.miller@gmail.com

CP3

Universal Upper Estimate for Prediction Errors under Moderate Model Uncertainty

We present a method of sensitivity analysis for general dynamical systems subjected to deterministic or stochastic modeling uncertainty. Using the properties of the unperturbed, idealized dynamics, we derive a universal bound for the leading-order prediction error. Specifically, our estimates give upper bounds on the leading order trajectory-uncertainty arising along model trajectories, solely as functions of the invariants of the known Cauchy-Green strain tensor of the idealized model. Our bounds turn out to be optimal, which means that they cannot be improved for general systems. This bound motivates the definition of the model sensitivity (MS), a scalar quantity closely related to the finite time Lyapunov exponents (FTLE) of the idealized model. We use nonlinear numerical models of various complexities to demonstrate that the model sensitivity provides both a global view over the phase space of the dynamical system and, in some situations, a localized, time-dependent predictor of uncertainties along trajectories. We show that the mean-squared trajectory uncertainty qualitatively follows the leading-order bound for surprisingly long time intervals. In addition, we use simple energy balance type models of the climate system to show that the method is expected to scale to models of higher dimensions.

Balint Kaszas, Philipp Bucholtz
ETH Zurich
bkaszas@ethz.ch, bucphili@student.ethz.ch

George Haller
ETH, Zurich
Switzerland
georgehaller@ethz.ch

CP3

A Variational Inference Approach to Inverse Problems with Gamma Hyperpriors

Hierarchical models with gamma hyperpriors provide a flexible, sparse-promoting framework to bridge L^1 and L^2 regularizations in Bayesian formulations to inverse problems. Despite the Bayesian motivation for these models, existing methodologies are limited to maximum a posteriori estimation. The potential to perform uncertainty quantification has not yet been realized. In this talk, we introduce a variational iterative alternating scheme for hierarchical inverse problems with gamma hyperpriors. The proposed variational inference approach yields accurate reconstruction, provides meaningful uncertainty quantification, and is easy to implement. In addition, it lends itself naturally to conduct model selection for the choice of hyperparameters. If time permits, we will illustrate the performance of our methodology in several computed examples, including a deconvolution problem and sparse identification of dynamical systems from time series data.

Shiv Agrawal, Hwanwoo Kim, Daniel Sanz-Alonso, Alexander Strang
University of Chicago
shivagr@uchicago.edu, hwkim@uchicago.edu, sanza-

lonso@uchicago.edu, alexstrang@uchicago.edu

CP3

Bayesian Parameter Inference on Arbitrary Multi-Resolution Polynomial Chaos Expansion Based Surrogate Models

Bayesian parameter inference based techniques became an integral part of the machine learning and data-driven uncertainty quantification in the last years. Despite the increasing power of the computational resources in the past decades the straightforward approach, which requires high number of model evaluations is still unfeasible for many problems. Therefore, data-driven surrogate models are very common for a variety of real-world applications. In particular, the traditionally data-poor geophysical applications poses the challenge for the construction of an appropriate stochastic discretization. Arbitrary multi-resolution polynomial chaos (aMR-PC) expansion based surrogate models, which combine the data-driven principles of the arbitrary polynomial chaos and multi-resolution based localization, which allows reducing the Gibbs phenomena of the polynomial representation of non-linear convection dominated problems. The goal of our work is on the one hand to provide the aMR-PC surrogate model based Bayesian parameter inference and on the other hand to reduce the number of the model evaluations, which are necessarily to build a reliable surrogate model. Numerical experiments demonstrate the performance and validate the accuracy of the Bayesian parameter inference techniques on non-intrusive implementation of aMR-PC in a non-linear hyperbolic scenario driven by several uncertain parameters, which is relevant also in the field of environmental sciences.

Ilja Kröker
University of Stuttgart
Institute for Applied Analysis and Numerical Simulation
ilja.kroeker@iws.uni-stuttgart.de

CP4

Bacterial Strain Design Using Active Subspaces

Salmonella can utilize a biodiesel waste product, glycerol, to produce 1,3-propanediol (1,3-PDO), a common commercial solvent. Experimental collaborators modify Salmonella to sequester the two enzymes needed for this pathway in microcompartments (MCPs), small protein-bound shells that spatially segregate reactions. I formulated a differential equation model of this system to compare five candidate strains with different concentrations of 1,3-PDO producing enzymes in the MCP. To rank the strains, I considered quantities of interest (QoIs): product yield, rate of production and toxicity level. Evaluating QoIs on uniformly sampled parameters, restricted by physical constraints and prior measurements, is computationally intractable. To efficiently generate QoI distributions, I used Active Subspaces to identify parameter directions that most affect each QoI. I then used maximin sampling to produce a space-filling spread of parameter samples in the active subspace while randomly perturbing the samples in the inactive subspace. This sampling reduced the computational load by at most 5 orders of magnitude when compared to a coarse grid sampling. The QoI distributions converged with increasing numbers of samples. I used hypothesis testing on the QoI distributions to predict optimal producing strains.

Andre Archer, Taylor Nichols
Northwestern University
andrearcher2017@u.northwestern.edu,
tay.nichols@u.northwestern.edu

Niall M. Mangan
Dept. Eng. Sci. and Applied Mathematics
Northwestern University
niall.mangan@northwestern.edu

Danielle Tullman-Ercek
Northwestern University
ercek@northwestern.edu

CP4

Modeling Transient Dynamics of Coarse-Grained Molecular Systems

In recent years, the development of model reduction and coarse-grained methodologies for studying large-scale molecular systems that cannot be practically studied with atom detailed molecular dynamics simulations is an active research field. Defining the new effective coarse-grained system, means finding the model which best represents the reference system both in structure and dynamic properties. In the present work, we approximate the dynamics of coarse-grained systems at the transient regime. We approximate the short time non-Markovian dynamics by Markovian dynamics with a time-dependent force field. We present the application of the path-space force matching method to retrieve the coarse space parametrized drift. We follow a data-driven approach to estimate the friction kernel, calculating correlation functions directly from the underlying all-atom Molecular Dynamics simulations. The proposed model's effectiveness is examined by comparing its structural and dynamical properties with the corresponding reference system. We also quantify the uncertainties in the effective model due to the involved approximations and the limited size of the available microscopic configurations by employing statistical methods. The methodology is illustrated for the molecular water system.

Georgia Baxevani
Department of Mathematics and Applied Mathematics, UoC
Institute of Applied and Computational Mathematics FORTH
g.baxevani@iacm.forth.gr

Evangelia Kalligiannaki
Institute of Applied and Computational Mathematics FORTH GR-
evangelia.kalligiannaki@iacm.forth.gr

Vagelis Harmandaris
University of Crete
harman@uoc.gr

CP4

Bayesian Learning Machines for Glioblastoma Multiforme Brain Tumor Evolution

With less than 5% of patients surviving 5 years following diagnosis, Glioblastoma multiforme (GBM) is the most common and aggressive form of primary brain tumor. We attempt to model the evolution of Glioblastoma Multiforme (GBM), the most common (and most aggressive) type of human brain cancer. However, a significant amount of uncertainty exists in the functional form of model equa-

tions and parameterizations. This is due to the complexity and lack of understanding of the processes involved, along with patient-specific differences in brain cell density and geometry. These challenges motivate the objective of the present work, in which we implement the mathematical model for GBM in a rigorous PDE-based machine-learning framework, utilizing the Dynamically Orthogonal (DO) evolution equations corresponding to the GBM model for effective dimensional reduction of the stochastic model. We then try to integrate noisy and sparse (in time) data extracted from magnetic resonance (MR) images of the patient into the model using a dynamics-based Bayesian learning framework. Such a predictive, mathematical model of the tumor growth could be very useful in a clinical setting, in coming up with a patient-specific treatment plan such as a chemotherapy schedule, or for the timing to schedule a surgery.

Pierre F. Lermusiaux
MIT
pierrel@mit.edu

Abhinav Gupta
Massachusetts Institute of Technology
guptaa@mit.edu

Ali Daher
MIT
adaher@mit.edu

Wael H. Ali
Massachusetts Institute of Technology
whajjali@mit.edu

CP4

Revisiting Distribution Discrepancies in Nonparametric Bayesian Classifiers for Microbial Abundance Data

Amplicon analysis of targeted gene sequences remains a popular technique for the identification of micro- and macrobiota due to its utility in environmental and climate research, public safety, and forensic analysis. Performing meaningful and unbiased classifications on such data, however, is difficult due both to the structure of the data and a series of complex and uncertainty-prone procedures required to generate the data. To address this challenge, Bayesian methods, especially nonparametric models, have recently been favored in the literature for robust analysis and classification of microbiome datasets. Unfortunately, the utility of this approach can be limited by the distributional assumptions made for raw and relative abundances, which may heavily underestimate higher-order moments and sparsity properties of this data. In this work, using both simulated datasets and data from environmental microbiomes, we used fitted beta mixture distributions to better attain realistic priors and uncertainty quantification. Samples were varied by location as well as by level of noise reduction and model outcomes were evaluated by comparisons both to simulated truth and to existing results on common datasets. Our results show that the implementation of the proposed methodology can be used to improve the performance of existing models. This analysis could also find application in other classification problems with compositional data such as in the health sciences or geology.

Daniel T. Fuller
Department of Mathematics
Clarkson Unviversity
fullerdt@clarkson.edu

Shantanu Sur
Department of Biology
Clarkson University
ssur@clarkson.edu

Marko Budišic
Clarkson University
marko@clarkson.edu

Sumona Mondal
Department of Mathematics
Clarkson University
smondal@clarkson.edu

CP4

Scalable Uncertainty Quantification Framework for Spatio-Temporal Spread of Covid-19

Susceptible-Exposed-Infected-Recovered-Deceased (SEIRD) compartmental models received widespread attention in predicting the spatio-temporal spread of COVID-19. However, numerous coefficients of these coupled nonlinear PDEs are not precisely known, for instance, due to uncertain disease transmission mechanism, uncertain travel pattern of infected people. Hence we adopt a stochastic version of the SEIRD PDEs with spatio-temporally varying infection rates and population-dependent diffusion coefficients (along with other random system parameters). The numbers of the coupled PDEs and the associated random model parameters increase substantially due to stratification based on age-structure, vaccine types, variants of COVID-19, disease severity (hospitalization with/without ICU admission), co-morbidity and public health interventions (quarantine, social distancing, travel restrictions). Sampling based uncertainty quantification (UQ) methods become computationally intensive for this large number of coupled stochastic PDEs. Hence we exploit intrusive polynomial chaos based UQ method combined with the overlapping domain decomposition algorithm to tackle the large coupled stochastic PDE based SEIRD model. We demonstrate the numerical and parallel (weak and strong) scalabilities of one-level and two-level Schwarz preconditioners against the number of random input parameters, mesh resolutions and the order of stratification.

Sudhi Sharma P V
Student
sudhipv@cmail.carleton.ca

Victorita Dolean
University of Strathclyde, UK and Universite Coˆte dAzur
victorita.dolean@strath.ac.uk

Mohammad Khalil
Sandia National Laboratories
mkhalil@sandia.gov

Chris Pettit
United States Naval Academy, USA
pettitcl@usna.edu

Abhijit Sarkar
Associate Professor
carleton University, Ottawa, Canada

carleton@abhijitsarkar.net

CP4

Nonparametric Estimation of the Reproductive Rate in Generalized SEIR Models

We develop a non-parametric method to estimate the time-dependent reproductive rate in generalized Susceptible-Exposed-Infected-Recovered (SEIR) models. The reproductive rate plays a central role in tracking infectious disease outbreaks at real time. It also measures epidemic severity and effectiveness of virus control strategies. We apply our methodology to study the severity and evolution of the Coronavirus 2019 outbreak in various American states. Our analysis reveals that virus control policies, including travel restrictions and national lockdown orders that close public meeting places, substantially reduced disease prevalence.

Imelda Trejo
University of Texas at Arlington
imelda@lanl.gov

CP5

Bayesian Calibration of Hypersonic Turbulent Flow Models Using Machine Learning and Reduced Order Modeling

In this work we study the efficacy of combining machine-learning methods with projection-based model reduction techniques for creating data-driven surrogate models of computationally expensive, high-fidelity simulated physics models. Such surrogate models are essential for many-query applications e.g., engineering design optimization and inverse problems or parameter estimation, where it is necessary to invoke the high-fidelity model sequentially, many times. Surrogate models are usually constructed for individual scalar quantities. However, there are scenarios where a spatially variable model output, e.g., a field needs to be modeled as a function of the model's input parameters. We develop a method to do so, using projections to represent spatial variability while a machine-learned model captures the dependence of the model's response on the inputs. The method is demonstrated on modeling the heat flux and pressure on the surface of the HIFiRE geometry in a Mach 7.16 turbulent flow. The surrogate model is then used to perform Bayesian estimation of freestream conditions and parameters of the SST (Shear Stress Transport) turbulence model embedded in the high-fidelity flow simulator, using wind-tunnel data. The paper provides the first-ever Bayesian calibration of a turbulence model for complex hypersonic turbulent flows. We also provide an easy-to-use Python library called tesuract, built on top of the scikit-learn API, for building the surrogates used in this work.

Kenny Chowdhary
Sandia National Laboratories
kchowdh@sandia.gov

Jaideep Ray
Sandia National Laboratories, Livermore, CA
jairay@sandia.gov

Chi K. Hoang
Sandia National Laboratory
ckhoang@sandia.gov

Kookjin Lee
Arizona State University
klee263@asu.edu

CP5

Deep Learning-Based Surrogate Modeling for Uncertainty Quantification in Soil-Structure Interaction Problems

Direct modeling of soil-structure interaction (SSI) problems calls for modeling the infrastructure, near-field soil, and boundaries/interfaces to translate the far-field earthquake excitation and absorb the scattered outgoing waves. On the other hand, a rigorous seismic performance assessment of SSI problems calls for considering and propagating uncertainties associated with the soil and structure properties and the excitation field characteristics. Monte Carlo-based methods are the commonly used approach for this purpose, which require running many forward simulations and can become computationally expensive and not practical for large-scale SSI systems. A viable solution to this problem is to replace the high-fidelity forward simulator with a surrogate model to provide an accurate and fast approximation of the high-fidelity simulator to accelerate Monte Carlo-based uncertainty quantification tasks. In this research, our primary goal is to determine the promises of deep learning-based models as a surrogate to approximate the time-dependent response of SSI direct models. Since the quantities of interest are time-dependent, the recurrent neural network is expected to perform better than other architectures, and we have already tested this architecture's predictive capability for a single degree of freedom system to emulate the system's dynamic response. In this presentation, we will show the extension of this work to elastodynamics and, more specifically, to SSI problems.

Elnaz Seylabi
University of Nevada Reno
elnaze@unr.edu

Hamid Ganji
University of Nevada, Reno, U.S.
h.taghavi.g@gmail.com

CP5

Deep Learning Weather Uncertainty Quantification for Earth Observation Satellite Mission Planning

To make full profit of Agile Earth Observation Satellites (AEOS), it is important to acquire as many cloudless images as possible. However, nebulosity is uncertain and if this uncertainty is not taken into account while computing the mission plan of an AEOS (which is a scheduling problem) it can lead to large amounts of images taken that are useless because they are too cloudy. There are two ways to deal with this uncertainty: firstly, it can be reduced by decreasing the time delta between the moment the weather is predicted and the moment the image is taken. This requires the ability to predict nebulosity on a short horizon, which is challenging knowing that numerical weather forecasting is highly computationally demanding. Secondly, robust scheduling techniques can be used by taking into account the probability of validation of an image during the mission plan computation. However, this probability can be computationally demanding to know. The present work proposes to tackle those two problems using Deep Learning. Nebulosity data with 15 minutes time discretization are publicly available. It gives the possibility to build short term predictive models. Besides, DL models can also give a

measure of confidence on the output which can be a highly valuable information for planning. We demonstrate that it can directly be used to evaluate the expectation of the criteria to optimize the mission without the use of computationally demanding techniques such as Monte Carlo.

Jonathan Guerra, Mathieu Picard
Airbus Defence & Space
jonathan.guerra@airbus.com, mathieu.picard@airbus.com

CP5

Machine Learning-Based Conditional Mean Filter: a Generalization of the Ensemble Kalman Filter for Nonlinear Data Assimilation

Filtering is a data assimilation technique that performs the sequential inference of dynamical systems states from noisy observations. Herein, we propose a machine learning-based ensemble conditional mean filter (ML-EnCMF) for tracking possibly high-dimensional non-Gaussian state models with nonlinear dynamics based on sparse observations. The proposed filtering method is developed based on the conditional expectation and numerically implemented using machine learning (ML) techniques combined with the ensemble method. The contribution of this work is twofold. First, we demonstrate that the ensembles assimilated using the ensemble conditional mean filter (EnCMF) provide an unbiased estimator of the Bayesian posterior mean, and their variance matches the expected conditional variance. Second, we implement the EnCMF using artificial neural networks, which have a significant advantage in representing nonlinear functions over high-dimensional domains such as the conditional mean. Finally, we demonstrate the effectiveness of the ML-EnCMF for tracking the states of Lorenz-63 and Lorenz-96 systems under the chaotic regime. Numerical results show that the ML-EnCMF outperforms the ensemble Kalman filter.

Truong-Vinh Hoang
RWTH-Aachen University
RWTH-Aachen University
hoang.tr.vinh@gmail.com

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

Raúl Tempone
Chair of Mathematics for Uncertainty Quantification
RWTH-Aachen University, Germany
tempone@uq.rwth-aachen.de

CP6

New Methodology to Avoid Overfitting in Bayesian Neural Networks

Frequentist deep neural networks have two significant shortcomings: the large number of weight parameters render them prone to overfitting, and their inability to adequately handle uncertainty in the prediction. In recent years, various regularisation schemes, such as early stopping, weight decay, and dropout, have been proposed to address the former issue, but they are still inadequate to handle the latter. Conversely, Bayesian neural networks (BNN) introduce uncertainties in the weights of the network to address the latter concern, yet overfitting issue remains as a serious concern. To address the aforementioned challenges, we propose the nonlinear sparse Bayesian neural network (NSBNN) by exploiting the recently proposed nonlinear sparse Bayesian learning algorithm (NSBL). It aims to address the practical and computational challenges involved in BNNs by 1) employing a sparsity inducing prior based on the concept of automatic relevance determination (ARD) and Gaussian mixture model approximations which permit semi-analytical calculation of Bayesian entities. This algorithm provides optimal sparse BNN(s) nested under a fully connected network by pruning redundant weight parameters through the evidence optimization procedure. The benefits of the NSBNN algorithm will be highlighted in relation to BNN.

Abhijit Sarkar
Associate Professor
carleton University, Ottawa, Canada
carleton@abhijitsarkar.net

Nastaran Dabiran
Carleton University, Ottawa, ON, Canada
nastarandabiran@cmail.carleton.ca

Brandon Robinson
Carleton University
brandonrobinson@cmail.carleton.ca

Rimple Sandhu
National Renewable Energy Laboratory Golden, Colorado, Unite
rimple.sandhu@nrel.gov

Mohammad Khalil
Sandia National Laboratories
mkhalil@sandia.gov

CP6

Ptychographic Inversion and Uncertainty Quantification Using Invertible Neural Networks

Ptychography is an essential imaging technique for high-resolution and nondestructive material characterization; however, its reconstruction requires solving a challenging large-scale nonlinear and non-convex inverse problem, whose solution is nonunique. Therefore, a robust reconstruction method with the capability of quantifying solution quality is highly desirable. Motivated by the recent success of invertible neural networks in solving inverse problems, in this work, we explore its application to ptychographic reconstruction. Specifically, invertible neural networks are employed to construct normalizing flows that act as surrogates for the high dimensional posterior density. More importantly, this also allows for the uncertainty quantification of the reconstruction: a desired capability for judging the solution quality among many local minima in the absence of ground truth, spotting spurious artifacts and guiding future experiments using the returned uncertainty patterns. The performance of the proposed method is demonstrated on a synthetic sample with added noise and in various physical experimental settings. The reconstruction quality is found to be comparable with existing methods.

Agnimitra Dasgupta
University of Southern California
adasgupt@usc.edu

Zichao Wendy Di
Mathematics and Computer Science Division

Argonne National Laboratory
wendydi@mcs.anl.gov

CP6

Bayesian Learning Machines for Discovering Dynamical Model Functions

We utilize and extend our rigorous PDE-based Bayesian learning framework for simultaneous learning of state variables, parameters, parameterizations, constitutive relations, and differential equations of high-dimensional dynamical models. The Bayesian learning machines can discriminate among existing models and now also extrapolate into the space of models to discover newer ones. The extended framework combines our Gaussian Mixture Model (GMM) - dynamically orthogonal (DO) filter for nonlinear reduced-dimension Bayesian inference with novel schemes from approximation theory and statistical learning theory for discovering new terms and functional forms in model equations. We also develop theory and methodology for handling stochastic boundary conditions, and for performing data-driven subspace augmentation using machine learning methods to represent the missing uncertainty in our reduced-dimension Bayesian inference. Results are showcased for varied coupled fish-biogeochemical-physical ocean dynamics.

Pierre F. Lermusiaux
MIT
pierrel@mit.edu

Abhinav Gupta
Massachusetts Institute of Technology
guptaa@mit.edu

CP6

Multi-Fidelity Hamiltonian Monte Carlo Method with Deep Learning-Based Surrogate

In recent years, the Hamiltonian Monte Carlo (HMC) method has emerged as a state-of-the-art MCMC technique that exploits the geometry of the target distribution to generate samples in an efficient manner. Despite its increasing popularity and impressive empirical success, its wide-scale adoption is still limited due to the high computational cost associated with gradient calculation. Moreover, the application of this method is simply not possible in scenarios where the gradient of the target posterior distribution cannot be computed (for example, with black-box simulators and/or with non-differentiable priors). To overcome these challenges, we propose a novel two-stage Hamiltonian Monte Carlo algorithm with a deep learning (DL)-based surrogate forward model. Splitting the standard HMC algorithm into two stages allows for efficient and computationally inexpensive evaluation of the gradient of the posterior by leveraging automatic differentiation capabilities of the surrogate model (thus retaining the advantages of HMC, such as scalability to high dimensions and faster convergence) while producing accurate posterior samples by using HF numerical solvers in the second stage. We demonstrate the effectiveness of this algorithm on a range of physics-based linear and non-linear Bayesian inverse problems, where it outperforms the traditional HMC algorithm in computational efficiency while retaining similar accuracy.

Dhruv V. Patel
Stanford University
dvpatel@stanford.edu

Jonghyun Lee
University of Hawaii at Manoa
jonghyun.harry.lee@hawaii.edu

Mojtaba Forghani
Stanford University
mojtaba@stanford.edu

Matthew Farthing
US Army Engineer Research and Development Center
matthew.w.farthing@erdc.dren.mil

Tyler Hesser
Coastal and Hydraulics Laboratory
US Army Engineer Research and Development Center
tyler.hesser@usace.army.mil

Peter K. Kitanidis
Dept. of Civil and Environmental Engineering
Stanford University
peterk@stanford.edu

Eric F. Darve
Institute for Computational and Mathematical Engineering
Stanford University
darve@stanford.edu

CP6

Critical Analysis of Physics-Aware Deep Learning Surrogates for Reverse Time Migration

In seismic exploration, the interpretation of seismic images is key to decision-making. Seismic imaging is affected by the presence of multiple sources of uncertainty. Reverse time migration (RTM) is a high-resolution depth migration approach for extracting seismic imaging in complex geologic structures. RTM is time-consuming and data-intensive. Also, when embedded in an uncertainty quantification algorithm (like the Monte Carlo method), RTM shows a manifold increase in its complexity and cost due to the high input-output dimensionality. Hence, one of the main challenges facing uncertainty quantification in seismic imaging is reducing the computational cost of the analysis. This work evaluates physics-aware deep learning strategies to act as surrogate models for RTM under uncertainty. Inputs are an ensemble of velocity fields expressing the uncertainty and the outputs, the seismic images. Here, we evaluate the power of machine learning, particularly generative adversarial nets (GANs), to learn probability distributions from the data, potentially modeling uncertainties in the seismic images. Also, we explore the versatility of the deep learning models to build a surrogate from data generated by different levels of fidelity. We show by numerical experimentation that the surrogate models can reproduce the seismic images accurately and, more importantly, the uncertainty propagation from the input velocity fields to the image ensemble.

Fernando A. Rochinha
Federal University of Rio de Janeiro
Rua Nascimento Silva 100 401
faro@mecanica.coppe.ufrj.br

Rodolfo Freitas
Federal University of Rio de Janeiro
rodolfosmfreitas@gmail.com

Alvaro Coutinho
Dept. of Civil Engineering
Federal University of Rio de Janeiro
alvaro@nacad.ufrj.br

Charlan Alves
Fedral University of Rio de Janeiro
charlandellon@gmail.com

Carlos Barbosa
High Performance Computing Center
Federal University of Rio de Janeiro
c.barbosa@nacad.ufrj.br

Debora Pina
Federal University of Rio de Janeiro
dbpina@cos.ufrj.br

Liliane Kunstmann
Computer Science
Federal university of Rio de Janeiro
lneves@cos.ufrj.br

Gabriel Guerra
Fluminense Federal University
olimarenho@gmail.com

Marta Mattoso
Fedral University of Rio de Janeiro
marta@cos.ufrj.br

Romulo M. Silva
Federal University of Rio de Janeiro
romulo.silva@nacad.ufrj.br

Bruno Silva
Fedral University of Rio de Janeiro
b.silva@nacad.ufrj.br

CP7

A Posteriori Error Estimation for Stochastic Collocation Applied to Parametric Parabolic PDEs

In this talk, we investigate a posteriori error estimation for time-dependent parametric partial differential equations discretized using non-intrusive stochastic collocation finite element methods. In particular, we look to estimate the distinct contributions to the total approximation error stemming from both the parametric and time-stepping discretization schemes, in order to drive adaptive solution algorithms. The parametric error associated with the stochastic collocation method is estimated using a hierarchical method. The time-stepping algorithm is treated as a black-box with control imposed over the estimated local truncation error. Numerical results will be presented for a time-dependent advection-diffusion problem with uncertain wind field. The evolution of error in time will be analyzed and the challenges of driving an adaptive-in-time stochastic collocation algorithm will be discussed.

Benjamin Kent, Catherine Powell
University of Manchester
benjamin.kent@manchester.ac.uk,
catherine.powell@manchester.ac.uk

David Silvester
School of Mathematics
University of Manchester
david.silvester@manchester.ac.uk

Malgorzata Zimon
IBM Research UK
Daresbury Laboratory
malgorzata.zimon@uk.ibm.com

CP7

A Multilevel Intrusive Method for Parametric PDEs

In this talk we discuss a state-of-the-art multilevel intrusive method that facilitates forward UQ for PDE models with finitely or countably infinitely many uncertain inputs. It is well known that for some classes of parametric PDEs, multilevel approximation schemes, unlike their single-level counterparts, can combat the curse of dimensionality. However, from a computational point of view, intrusive methods, also known as stochastic Galerkin methods, are often viewed as less attractive than non-intrusive methods such as stochastic collocation schemes. We describe theoretical and computational aspects of an adaptive multilevel stochastic Galerkin finite element method (SGFEM). After discussing the hierarchical a posteriori error estimation strategy that drives the adaptive algorithm, and reviewing known theoretical results regarding optimal convergence rates for a class of parametric elliptic PDEs, we present numerical results showing that our algorithm achieves these rates for certain benchmark problems, outperforming adaptive single-level methods.

George Papanikos
Research associate The University of Manchester
george.papanikos@manchester.ac.uk

Catherine Powell
University of Manchester
catherine.powell@manchester.ac.uk

CP7

Sparse Learning of Nonlinear Stochastic Dynamical Systems

In many applications it may be desirable to model physical systems semi-empirically; in part based on a simplified description of the physics and complemented by a data-driven component. The physics-informed component of the model may be parameterized by imprecisely known parameters for which some prior knowledge is available. Conversely, the unknown coefficients in the data-driven component typically will not have any associated prior knowledge. It is possible to estimate these parameters using standard methods for Bayesian inversion. However, this approach can suffer from overfitting problem, leading to a large uncertainty in the prediction. To address this challenge, the nonlinear sparse Bayesian learning (NSBL) algorithm can be used to induce sparsity among the data-driven coefficients. NSBL leverages concepts from sparse Bayesian learning (SBL) and relevance vector machines (RVM) and extends these popular machine learning tools to nonlinear models and non-Gaussian likelihood functions. Using automatic relevance determination (ARD) priors and Gaussian mixture model approximations, NSBL offers a semi-analytical method for efficiently inducing sparsity among the model parameters. In this talk, the benefit of the NSBL algorithm will be illustrated for nonlinear dynamical systems for which we have data. We will also provide insights into the algorithm through parametric studies and by comparing its performance to standard methods for Bayesian

model selection.

Brandon Robinson
Carleton University
brandonrobinson@cmail.carleton.ca

Rimple Sandhu
National Renewable Energy Laboratory Golden, Colorado, Unite
rimple.sandhu@nrel.gov

Mohammad Khalil
Sandia National Laboratories
mkhalil@sandia.gov

Chris Pettit
United States Naval Academy, USA
pettitcl@usna.edu

Dominique Poirel
Royal Military College of Canada, Canada
poirel-d@rmc.ca

Abhijit Sarkar
Associate Professor
carleton University, Ottawa, Canada
carleton@abhijitsarkar.net

CP7

A Dimension-Adaptive Sparse Grid Combination Technique for High-Dimensional Parametric Elliptic PDEs

We present an adaptive algorithm for the computation of quantities of interest involving the solution of a stochastic elliptic PDE where the diffusion coefficient is parametrized by means of a Karhunen-Loeve expansion. The approximation of the equivalent parametric problem requires a restriction of the countably infinite dimensional parameter space to a finite-dimensional parameter set, a spatial discretization and an approximation in the parametric variables. We consider a sparse grid approach between these three approximation directions in order to reduce the computational effort and propose a dimension-adaptive combination technique. In addition, a sparse grid structure for the high-dimensional parametric approximation is considered and detected simultaneously with the spatial and stochastic approximation. The adaptive construction of the sparse grid is based on the benefit-cost ratio such that, different to a-priori approaches, regularity and decay of the Karhunen-Loeve coefficients are not required. The decay of the KL is detected and exploited as the algorithm adjusts to the anisotropy in the parametric variables. We include numerical examples for the Darcy problem with a lognormal permeability field, which illustrate a good performance of the algorithm: For sufficient smooth random fields, we essentially recover the rate of the spatial discretization as asymptotic convergence rate with respect to the computational cost.

Uta Seidler
University of Bonn
Institute for Numerical Simulation
seidler@ins.uni-bonn.de

Griebel Michael
University of Bonn, Institute for Numerical Simulation
Fraunhofer Institute SCAI
griebel@ins.uni-bonn.de

CP7

A Multi-Level Stochastic Collocation Method for Schrdinger Equations with a Random Potential

We propose and analyze a numerical method for time-dependent linear Schrdinger equations with uncertain parameters in both the potential and the initial data. The random parameters are discretized by stochastic collocation on a sparse grid, and the sample solutions in the nodes are approximated with the Strang splitting method. The computational work is reduced by a multi-level strategy, i.e. by combining information obtained from samples computed on different refinement levels of the discretization. We prove new error bounds for the time discretization which take the finite regularity in the stochastic variable into account, and which are crucial to obtain convergence of the multi-level approach. The predicted cost savings of the multi-level stochastic collocation method are verified by numerical examples.

Benny Stein
Karlsruhe Institute of Technology (KIT)
Institute for Applied and Numerical Mathematics
benny.stein@kit.edu

Tobias Jahnke
Karlsruhe Institute of Technology
Karlsruhe, Germany
tobias.jahnke@kit.edu

CP8

Mixed Effects State-Space Models Across Populations of Dynamic Systems

The rapid progression of sensing technology enables rich population-level data in practical applications; in particular, the work here is motivated by examples of engineering systems and infrastructure - from offshore wind farms to fleets of vehicles. We consider a set of differential equations that correspond to a population of engineering dynamical systems. We then propose an adapted mixed-effects linear Gaussian state-space model, to capture the variation across the population, alongside within-system correlations. The resultant model gives an interpretation of the global (population) dynamics as well as system-specific models. Most critically, the multi-level approach allows for knowledge transfer - since correlated latent variables allow systems with sparse information to borrow statistical strength from those that are data-rich. Firstly, the hierarchical Bayesian approach is constructed with a Gibbs procedure, and then with a novel sequential Monte Carlo method, to support online inference, which is advantageous in applications to streaming data. We demonstrate the mixed-effects state-space model in a simulated structural dynamics case study and show applications to measurements from an operational wind farm. The proposed population state-space model is general and should be widely applicable to longitudinal studies of dynamic systems.

Lawrence A. Bull
The Alan Turing Institute
lbull@turing.ac.uk

Mark Girolami
University of Cambridge
mag92@cam.ac.uk

Timothy Rogers
The University of Sheffield
tim.rogers@sheffield.ac.uk

CP8

Bayesian Lagrangian Data Assimilation and Learning

Dynamical transports a variety of natural quantities (e.g. aerosols, pathogens, water masses, plankton, sediments, etc.) and artificial materials (e.g. pollutants, floating debris, search and rescue, etc.). Lagrangian Coherent Structures (LCSs) or the most influential/persistent material lines in a flow, provide a robust approach to characterize such transports and organize classic trajectories. Using flow-map stochastic advection, dynamically-orthogonal (DO) decompositions, and Gaussian Mixture Model (GMM)-DO filtering, we derive uncertainty prediction and Bayesian learning schemes for both Eulerian and Lagrangian variables. The resulting nonlinear Eulerian-Lagrangian Bayesian data assimilation allows the simultaneous non-Gaussian estimation of Eulerian variables (e.g. velocity, tracers, etc.) and Lagrangian variables (e.g. positions, trajectories, LCSs, etc.), as well as the learning of model functions. Finally, we outline the deep machine learning of Lagrangian flow maps from snapshot and trajectory data. We showcase results in idealized and realistic ocean examples. We further show how our Bayesian mutual information and adaptive sampling equations provide a rigorous efficient methodology to plan optimal Lagrangian deployments.

Manan Doshi, Chinmay S. Kulkarni
Massachusetts Institute of Technology
mdoshi@mit.edu, chinmayk@mit.edu

Pierre F. Lermusiaux
MIT
pierrel@mit.edu

CP8

Uncertainty Quantification of Bifurcations in Random Ordinary Differential Equations

Subsystems of the earth might undergo critical transitions under sustained global warming with severe impacts on various ecosystems and human habitat. This phenomenon is not restricted to climate science but appears also in ecology and epidemiology [C. Kuehn, C. Bick, A universal route to explosive phenomena (2021), Science Advances, Vol. 7, No. 16]. Here, we approach these critical transitions mathematically in terms of bifurcation theory for nonlinear ordinary differential equations. It is well-known in the theory of dynamical systems that parameter variation can induce bifurcations. Our main question of interest is how uncertainties in system parameters propagate through the possibly highly nonlinear dynamical system and affect the system's bifurcation behavior. We analyze the effect of parametric uncertainty on the occurrence of different bifurcation types (sub- vs super-critical) along a given bifurcation curve. We come up with estimates for the bifurcation type probabilities to contribute to an uncertainty quantification of the exposure to critical transitions. In our methodology, we combine known statistical and probabilistic concepts with classical analysis and bifurcation theory. In a numerical case study, we illustrate the performance of the estimation procedure.

Kerstin Lux
Technische Universität München
kerstin.lux@tum.de

Christian Kuehn
Technical University of Munich
ckuehn@ma.tum.de

CP8

Physically Driven EV-GDEE for the Estimation of Time-Variant Failure Probability of High-Dimensional Stochastic Dynamical Systems

Time-variant reliability assessment of engineering systems subjected to stochastic excitations, especially for rare events, is of paramount importance for the performance-based decision-making of design, but is still of great challenge due to the nonlinear and random coupling in high-dimensional systems. For this purpose, an ensemble-evolving-based generalized density evolution equation (EV-GDEE) is established, as a one- or two-dimensional partial differential equation, with respect to the response of interest in a high-dimensional system. The equivalent drift coefficient in the EV-GDEE represents the physically driving force of evolution of the probability density function (PDF) in the ensemble sense, and is identified as the conditional expectation of the original drift function. In this sense, the proposed method can be called as the physically driven EV-GDEE. Some representative dynamical analyses of the underlying physical system are performed for the identification of the equivalent drift coefficient. Then, the EV-GDEE can be solved to capture the transient PDF of the response of interest. For the purpose of reliability, an absorbing boundary process is constructed. Its EV-GDEE can be established and solved to obtain time-variant first-passage reliability. The proposed method shows the high accuracy of the failure probability even in the order of magnitude 10-4 10-6 for rare events, which are achieved with only hundreds of dynamical analyses.

Meng-Ze Lyu, Jian-Bing Chen
Tongji University
lyumz@tongji.edu.cn, chenjb@tongji.edu.cn

CP8

Building and Solving Efficient Reduced Models for the Uncertain Boltzman Equation with MC-gPC: Applications to Neutronics (keff) and Photonics

Many physical applications rely on Monte-Carlo (MC) codes to solve deterministic partial differential equations. The simulations are costly but the MC resolution is competitive due to the high dimensional (7) deterministic problem. Propagating uncertainties with respect to d parameters is also of interest. Non-intrusive methods are usually applied. But they demand a high number of runs of the code. In our MC resolution context, each run is costly. One run can take several hours on hundreds of processors. For this reason, MC-gPC is introduced: it consists in building a Polynomial Chaos (gPC) based reduced model and in solving it with an MC scheme. The method is intrusive but fast convergence rates have been observed/proved: the method is efficient for the linear, nonlinear Boltzman equation and keff computations.

Gael Poette
CEA

gael.poette@gmail.com

CP9

A Probabilistic, Data-Driven Closure Model for Rans Simulations and Model Uncertainty

Despite significant improvements in high-performance computing, high-fidelity simulations of real-world, turbulent flows are largely infeasible. The Reynolds Averaged Navier-Stokes (RANS) model remains the most efficient alternative but it requires a Reynolds-stress (RS) closure model. Several such models have been developed but their accuracy is limited and they lack generality. In this work, we argue that the information loss that takes place when coarse-graining the original equations in order to come up with the RANS equations gives rise to model uncertainty which goes beyond the stochastic variability of the parameters of the closure model. The latter is written as the sum of a parameterized term dependent on the gradient of mean velocity and a stochastic, discrepancy tensor. We attempt to learn its statistical characteristics using a few high-fidelity simulations (e.g. LES or DNS) and a differentiable RANS solver. This in turn provides a stochastic correction to the RANS model which can now produce quantitative metrics of its predictive uncertainty. We demonstrate the efficacy of the proposed model and report on its performance as compared to classical turbulence models and high-fidelity solvers.

Atul Agrawal, Phaedon-Stelios Koutsourelakis
Technical University of Munich
atul.agrawal@tum.de, p.s.koutsourelakis@tum.de

CP9

Tree-Based Gaussian Process for Computer Experiments with Many-Category Qualitative Factors and Application to Cooling System Design

In computer experiments, Gaussian process models are commonly used for emulation. However, when both qualitative and quantitative factors in the experiments, emulation using Gaussian process models becomes challenging. In particular, when the qualitative factors contain many categories in the experiments, existing methods in the literature become cumbersome due to curse of dimensionality. Motivated by the computer experiments for the design of a cooling system, a new tree-based Gaussian process for emulating computer experiments with many-category qualitative factors is proposed. The proposed method incorporates a tree structure to split the categories in the qualitative factors, and Gaussian process models are employed for modeling the simulation outputs in the leaf nodes. The splitting rule takes into account the cross-correlations between the categories of the qualitative factors, which have been shown by a recent theoretical study to be a crucial element for improving the prediction accuracy. The application to the design of a cooling system indicates that the proposed method not only enjoys marked computational advantages and produces accurate predictions, but also provides valuable insights into the cooling system by discovering the tree structure.

Wei-Ann Lin
National Cheng Kung University
saga.lin.14@gmail.com

Chih-Li Sung
Michigan State University
Department of Statistics and Probability
sungchih@msu.edu

Ray-Bing Chen
Department of Statistics
National Cheng Kung University
rbchen@ncku.edu.tw

CP9

Automatic Dynamic Relevance Determination for Gaussian Process Regressions with Functional Inputs

Numerous geophysical parameters are retrieved from the spectral line radiance measured by NASA's Microwave Limb Sounder via a computer model (forward model). Its expensive evaluation time motivates the design of a statistical surrogator for uncertainty quantification. The forward model takes a finite realization of atmospheric state variables over a pressure grid characterizing vertical profiles over a continuous index. A common emulator for a computer experiment with functional inputs involves a Gaussian Process with a vector input resulting from the functional input pre-processing. Generalizing a framework originally proposed for time-varying inputs, we introduce the asymmetric Laplace functional weight (ALF): a flexible, parametric mapping between the output correlation and the input functional structure. Automatic Dynamic Relevance Determination (ADRD) is achieved with at most three unknowns per input variable. A simulation study is conducted to assess ADRD capability to recover the true model and produce accurate predictions. Additionally, ADRD is applied to predict water vapor spectra for different atmospheric configurations. The parameters are learned from data using a fully-Bayesian approach. The ALF posterior density is explored to identify the most relevant regions of the atmosphere and out-of-sample validation is performed to benchmark ADRD against vector-input models.

Luis Damiano
Iowa State University
ldamiano@iastate.edu

Joaquim Teixeira, Margaret Johnson
Jet Propulsion Laboratory
joaquim.p.teixeira@jpl.nasa.gov,
maggie.johnson@jpl.nasa.gov

Jarad Niemi
Iowa State University
niemi@iastate.edu

CP9

Sparse Tensor Product Approximation for a Class of Generalized Method of Moments Estimators

Generalized Method of Moments (GMM) estimators in their various forms, including the popular Maximum Likelihood (ML) estimator, are frequently applied for the evaluation of complex econometric models with not analytically computable moment or likelihood functions. As the objective functions of GMM- and ML-estimators themselves constitute the approximation of an integral, more precisely of the expected value over the real world data space, the question arises whether the approximation of the moment function and the simulation of the entire objective function can be combined. Motivated by the popular Probit and Mixed Logit models, we consider double integrals with a

linking function which stems from the considered estimator, e.g. the logarithm for Maximum Likelihood, and apply a sparse tensor product quadrature to reduce the computational effort for the approximation of the combined integral. Given Hlder continuity of the linking function, we prove that this approach can improve the order of the convergence rate of the classical GMM- and ML-estimator by a factor of two, even for integrands of low regularity or high dimensionality. This result is illustrated by numerical simulations of Mixed Logit and Multinomial Probit integrals which are estimated by ML- and GMM-estimators, respectively.

Michael Griebel
Institut für Numerische Simulation, Universität Bonn & and Fraunhofer-Institut für Algorithmen und Wissenschaftlich
griebel@ins.uni-bonn.de

Alexandros Gilch
Bonn Graduate School of Economics, Universität Bonn
alexandros.gilch@uni-bonn.de

Jens Oettershagen
Institut für Numerische Simulation, Universität Bonn
jens.oettershagen@gmx.de

CP9

Higher-Dimensional Deterministic Formulation of Hyperbolic Conservation Laws with Uncertain Initial Data

We discuss random hyperbolic conservation laws and introduce a novel formulation interpreting the stochastic variables as additional spatial dimensions with zero flux. The approach is compared with established non-intrusive approaches to random conservation laws. In the scalar case, an entropy solution is proven to exist if and only if a random entropy solution for the original problem exists. Furthermore, existence and numerical convergence of stochastic moments is established. Along with this, the boundedness of the L^1-error of the stochastic moments by the L^1-error of the approximation is proven. For the numerical approximation a Runge-Kutta discontinuous Galerkin method is employed and a multi-element stochastic collocation is used for the approximation of the stochastic moments. By means of grid adaptation the computational effort is reduced in the spatial as well as in the stochastic directions, simultaneously. Results on Burgers' and Euler equation are validated by several numerical examples and compared to Monte Carlo simulations.

Adrian J. Kolb, Michael Herty, Siegfried Mueller
RWTH Aachen University
kolb@eddy.rwth-aachen.de, herty@igpm.rwth-aachen.de, mueller@igpm.rwth-aachen.de

CP9

Continuous Space Representation of Geological Conceptual Models

Considering an ensemble of geological models is a necessity for uncertainty quantification in the exploration of natural resources such as groundwater, minerals or geothermal energy. However, it is often limited to a single concept or single geological interpretation, out of convenience or because of perception bias. Even though several conceptual models are retained, they might not be fit for real data application. Indeed, the lack of knowledge and data combined with human perception, interpretation and bias might lead to a set of erroneous or incomplete initial scenarios. Here, we propose to build a continuous parameter space representation of conceptual models from a discrete and finite set of initial geological scenarios. The approach relies on summary statistics dissimilarities between models, and scenario representatives. It allows us to explore additional locations of the conceptual model parameter space to build scenarios that are compatible with collected geological data and it allows us to quantify uncertainty around the initial scenarios. To illustrate our approach, we use an ensemble of one million synthetic geological models with complete tectonic histories (from https://github.com/Loop3D/noddyverse). We acknowledge the support from the ARC-funded Loop: Enabling Stochastic 3D Geological Modelling consortia (LP170100985) and the Mineral Exploration Cooperative Research Centre. This is MinEx CRC Document 2021/***.

Guillaume Pirot, Mark Jessell
The University of Western Australia
guillaume.pirot@uwa.edu.au, mark.jessell@uwa.edu.au

Mark Lindsay
CSIRO
mark.lindsay@csiro.au

CP10

Robust Bayesian Experimental Design

In the field of Bayesian Experimental Design (BED), there has been much recent work on algorithms for maximizing the expected information gain. This optimization problem is difficult due to the nested expectations over the prior distribution of outcomes and over the posterior distribution of parameters conditioned on each outcome that define the objective. Examples of recent approaches include the use of neural networks to estimate the KL divergence between prior and posterior distributions, and the use of nested Monte Carlo sampling with adaptive proposal distributions. These approaches, however, do not try to account for the robustness of their solutions with respect to variations in the prior distribution. In this work, we model uncertainty in the prior distribution with an ambiguity set of small divergence from a reference prior, and consider several BED formulations that seek to maximize the information gain over the ambiguity set as a whole. To address the issue related to the expensive computation, we build on those recent algorithms, but we explore further changes of this problem including dual formulations to reduce the dimensionality.

Jinwoo Go
Georgia Institute of Technology
jgo31@gatech.edu

Tobin Isaac
Georgia Tech
tisaac@cc.gatech.edu

CP10

Targeted Adaptive Design

We describe Targeted Adaptive Design a new data-driven algorithm for adaptively exploring an unknown multi-output response in a multivariate input design space, while searching for inputs that result in specific desired outputs, within specified tolerance. Design of products using advanced manufacturing equipment can frequently be de-

scribed by this kind of formalism. The algorithm bears a similar relationship to optimal sequential experimental design that root-finding bears to optimization. At each stage of the algorithm, a vector-valued Gaussian process (VVGP) surrogate of the response is constructed and optimized using the latest round of acquired samples, as well as all previously-acquired samples, after which N new sample locations and a target location are optimized to maximize the expected predictive log-likelihood of the target parameters at the target location, conditioned on the (known) acquired data and (unknown, marginalized) N future samples. A stopping decision (success or failure) is made based on the chosen tolerance, the current surrogate, and the current optimal target location. We describe the performance of the algorithm on a simulated design problem, and compare the performance of algorithms based on grid and random sampling data acquisition strategies.

Carlo Graziani
Argonne National Laboratory
cgraziani@anl.gov

Marieme Ngom
University of Illinois at Chicago
mngom@anl.gov

CP10

Turbulence Model Form Errors in Complex Flow Configurations

Direct Numerical Simulation of turbulence is a powerful tool that has the disadvantage of high computational cost. Models can be used to avoid this issue, however, these models introduce uncertainty, namely model form uncertainty, due to their inherent physical assumptions. Understanding these uncertainties and how they affect the quantities of interest is of the utmost importance when analyzing simulation results. In this work, an implied models approach is used to better understand the sources and dynamics of model form error in complex turbulent flows. The implied models approach is a physics-based uncertainty quantification (UQ) approach in which a transport equation is derived for the model error through the transport equation implied by the model for the quantity of interest. Multiple flow configurations have been analyzed with this UQ approach, including a boundary layer over a flat plate with a statistically stationary separation bubble. This flow is shown to have two error modes corresponding to the qualitative behavior of turbulent wall-bounded and free-shear model form errors, which have previously been studied. These results show a complex picture of model error that changes through the flow but also that calibration of turbulence models against simpler flows may capture the main modes of model failure. Preliminary analysis of multi-physics turbulent flows is conducted to understand the modes of model failure with the introduction of other physical processes.

Kerry S. Klemmer
Princeton University
kklemmer@princeton.edu

Wen Wu
University of Mississippi
wu@olemiss.edu

Michael E. Mueller
Princeton University
muellerm@princeton.edu

CP10

Bayesian Multi-Fidelity Inverse Analysis for Computationally Demanding Models in High Stochastic Dimensions

The biggest challenges in (Bayesian) inverse analysis of large-scale numerical models are posed by the high computational demands in combination with the high stochastic dimension. The solution process is further impeded when model derivatives are inaccessible as is often the case in legacy codes and coupled problems. We propose Bayesian multi-fidelity inverse analysis (BMFIA) which overcomes the aforementioned difficulties by employing computationally inexpensive, lower-fidelity models and constructing a multi-fidelity likelihood function. The latter is learned robustly from a small number of high-fidelity simulations and reflects the uncertainty, not only in the original inverse problem but also due to the (small) training data employed. BMFIA is independent of the problem's stochastic dimension as it primarily relies on the dependence between the outputs of models of varying fidelities and not on the input. Furthermore, improved efficiency is attained since the inference process, which can be performed using state-of-the-art sampling-based or variational methods, requires solely evaluations of the low-fidelity model(s). The latter can be chosen or constructed so that they provide model derivatives (e.g. from adjoint formulations) which further expedite inference. We demonstrate our approach on large-scale biomechanical problems and compare them with state-of-the-art single- and multi-fidelity methods.

Jonas Nitzler, Wolfgang Wall
Technical University of Munich
Institute for Computational Mechanics (LNM)
jonas.nitzler@tum.de, wolfgang.a.wall@tum.de

Phaedon S. Koutsourelakis
Technical University Munich
p.s.koutsourelakis@tum.de

CP10

Bayesian Calibration of Imperfect Computer Models Using Physics-Informed Priors

We introduce a computational efficient data-driven framework suitable for the quantification of the uncertainty in physical parameters of computer models, represented by differential equations. We construct physics-informed priors for time-dependent differential equations, which are multi-output Gaussian process priors that encode the model's structure in the covariance function. We extend this into a fully Bayesian framework which allows quantifying the uncertainty of physical parameters and model predictions. Since physical models are usually imperfect descriptions of the real process, we allow the model to deviate from the observed data by considering a discrepancy function. To obtain the posterior distributions we use Hamiltonian Monte Carlo sampling. To demonstrate our approach we use the arterial Windkessel model, which describes the hemodynamics of the heart through differential equation with physically interpretable parameters of medical interest. As most physical models, the Windkessel model is an imperfect description of the real process. In a synthetic case study, we simulate noisy data from a more complex physical model with known mathematical connections to our modelling choice. We show that without accounting for discrepancy, the posterior of the physical parameters devi-

ates from the true value while accounting for discrepancy gives reasonable quantification of physical parameters uncertainty and reduces the uncertainty in subsequent model predictions.

Michail Spitieris
NTNU (Norwegian University of Science and Technology)
NTNU
michail.spitieris@ntnu.no

Ingelin Steinsland
NTNU (Norwegian University of Science and Technology)
ingelin.steinsland@ntnu.no

CP11

Optimization under Epistemic Uncertainty Using Bayesian Hybrid Models

We propose Bayesian hybrid models (BHMs) to simultaneously account for epistemic and aleatory uncertainty. Model inadequacy or neglected control variables lead to epistemic uncertainty, for example, the use of low fidelity surrogate models in multiscale engineering frameworks. Aleatory uncertainty arises due to random phenomena such as experimental variability. The proposed BHMs extend Kennedy and OHagans statistical framework and takes the form: y=?(.)+d(.)+? The glass-box component ?(.) is the mechanistic model with epistemic uncertainty. A Gaussian process (GP) models the discrepancy d(.) which is the black-box component of the model. The aleatory uncertainty is modeled by the normally distributed random noise ?. The proposed hybrid model combines the benefits of black-box models that leverage statistics and machine learning to reveal trends in data and glass-box models that are developed from scientific theories. The probabilistic GP enables Bayesian model calibration and provides readily interpretable uncertainty information as opposed to other hybrid modeling constructs such as neural differential equations. Using ballistic firing and reaction engineering case studies, we show the superior performance of the BHMs over simplified glass-box and GP regression only models for optimal decision-making under aleatory and epistemic uncertainties.

Elvis Eugene, Jialu Wang
University of Notre Dame
eeugene@nd.edu, jwang44@nd.edu

Xian Gao
Notre Dame
xgao1@nd.edu

Alexander Dowling
University of Notre Dame
adowling@nd.edu

CP11

Lookahead Bayesian Optimization and Applications

We propose a novel Bayesian optimization framework that introduces a *lookahead* acquisition principle. This includes a generalized framework which can be used to construct lookahead versions of existing (myopic) acquisition principles. Furthermore, we show that our framework can be leveraged to solve problems in a multifidelity setting, i.e., when multiple models that trade accuracy for computational expense are available. We demonstrate our method on application problems in optimization, surrogate modeling, and uncertainty quantification.

Ashwin Renganathan
Argonne National Laboratory
ashwin.renganathan@utah.edu

CP11

Batch-Based Bayesian Optimization of Stochastic Functions Through Uncertainty Quantification

In this talk, we will introduce an uncertainty quantification-based approach for the optimization of stochastic functions with heteroscedastic noise (e.g., agent-based models, molecular dynamics simulations, and physical experiments). The advantage of the proposed approach is twofold: (i) reduced computational cost through carefully selected batches of samples, and (ii) tractability by considering the tradeoff between exploration and replication. The benefit of replicating at previously observed designs is that it gives direct insight into the noise of our data source and it improves scalability by reducing the size of the Gaussian process models covariance matrix. The main concern when deciding where to sample next is to minimize the uncertainty in the optimal design. Consequently, the exploration versus replication decision is made based on what alternative minimizes the posterior predictive uncertainty in the regions of the design spaces that are expected to contain the global optimum. Finally, the identification of batches is achieved through the introduction of a preposterior analysis that is compatible with stochastic data. Through the optimization of eight test functions and one engineering problem, we demonstrate that the proposed framework has superior performance compared to available methods, the validity of the introduced preposterior analysis, and its practical utility in design.

Anton Van Beek, Wei Chen
Northwestern University
antonvanbeek2022@u.northwstern.edu, weichen@northwestern.edu

CP11

Efficient Importance Sampling via Optimal Control for Stochastic Reaction Networks

We are interested in the efficient estimation of statistical quantities, particularly rare event probabilities, for stochastic reaction networks (SRNs). To this aim, we propose a novel importance sampling (IS) approach to improve the efficiency of Monte Carlo (MC) estimators when based on an approximate tau-leap (TL) scheme. In IS, the crucial step is to choose an appropriate change of probability measure to achieve a substantial variance reduction. Based on an original connection between finding the optimal IS parameters, within a class of probability measures, and a stochastic optimal control (SOC) formulation, we propose an automated approach to derive a highly efficient path-dependent measure change. Given that it is challenging to solve this SOC problem analytically, we propose a numerical dynamic programming algorithm to approximate the optimal control parameters. In the one-dimensional case, our numerical results show that the variance of our proposed estimator decays with rate $\mathcal{O}(\Delta t)$ for a step size of Δt, compared to being $\mathcal{O}(1)$ for the standard MC estimator. To mitigate the curse of dimensionality in the multi-dimensional case, we propose an alternative learning-based method that approximates the value function by a neural

network whose parameters are determined via a stochastic optimization algorithm. Our numerical experiments demonstrate that our learning-based IS approach substantially reduces the MC estimator's variance.

Sophia F. Wiechert, Chiheb Ben Hammouda
RWTH Aachen University
wiechert@uq.rwth-aachen.de,
benhammouda@uq.rwth-aachen.de

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

Raul F. Tempone
Mathematics, Computational Sciences & Engineering
King Abdullah University of Science and Technology
raul.tempone@kaust.edu.sa

CP11

Stochastic Momentum Methods for Optimal Control Problems Containing Uncertain Coefficients

In this talk, we investigate a numerical analysis of a strongly convex and smooth optimization problem governed by a convection diffusion equation with uncertain coefficients. To solve the underlying optimization problem finite element approximation in the spatial domain and Monte Carlo approximation in the risk measure are employed. On the other hand, stochastic approximation where true gradient is replaced by a stochastic ones is used to minimize the objective functional containing random terms. However, naive use of the stochastic gradient algorithm in many instances suffers difficult tuning of parameters and extremely slow convergence rate due to the noisy nature of stochastic gradient iteration. Therefore, we add a momentum term to accelerate the convergence behavior. Efficiency of the proposed methodology is illustrated by numerical experiments on the benchmark problems. It has been shown in the numerical simulations that stochastic approximation methods can be an alternative to solve PDE–constrained optimization problems containing uncertain terms.

Hamdullah Yücel
Institute of Applied Mathematics
Middle East Technical University
yucelh@metu.edu.tr

Sitki Can Toraman
Middle East Technical University
storaman@metu.edu.tr

CP11

Stochastic Discontinuous Galerkin Methods for Robust Deterministic Control of Convection Diffusion Equations with Uncertain Coefficients

This talk aims to investigate a numerical behaviour of a robust deterministic optimal control problem subject to a convection diffusion equation containing uncertain inputs. Stochastic Galerkin approach, turning the original optimization problem containing uncertainties into a large system of deterministic problems, is applied to discretize the stochastic domain, while a discontinuous Galerkin method is preferred for the spatial discretization due to its better convergence behaviour for optimization problems governed by convection dominated PDEs. Large matrix system emerging from the stochastic Galerkin method is addressed by the low–rank version of GMRES method, which reduces both the computational complexity and the memory requirements by employing the Kronecker–product structure of the obtained linear system. Benchmark examples with and without control constraints are presented to illustrate the efficiency of the proposed methodology.

Pelin iloglu
Middle East Technical University
pciloglu@metu.edu.tr

Hamdullah Yücel
Institute of Applied Mathematics
Middle East Technical University
yucelh@metu.edu.tr

CP12

Sar Image Formation Using Empirical Bayesian Inference With Joint Sparsity

Synthetic aperture radar (SAR) is a day and night, all weather imaging modality that utilizes a satellite-mounted radar to send and receive a complex signal. Multiple data acquisitions are often made over the same scene. These measurements are corrupted by noise as well as speckle, a phenomenon inherent to coherent imaging systems. Since the intensity of the underlying image is presumably sparse in some domain, e.g. the gradient or edge domain, compressive sensing may be used to obtain a point estimate. There is growing interest when considering the problem of SAR image reconstruction to quantify the uncertainty of the reconstruction. In this regard, by assuming that the true signal is represented as a random variable, one can attempt to recover its associated probability density function. We propose a novel method of SAR image reconstruction using empirical Bayesian inference that recovers both a point estimate and samples of a posterior distribution on the image pixels, providing this uncertainty quantification. Based on an a-priori estimate of the support in the edge domain which is obtained from the multiple data acquisitions, we define the support informed sparse prior which is then used in turn to generate samples for the posterior density function. We demonstrate the efficacy of our method on SAR phase history data.

Dylan Green
Dartmouth College
dylan.p.green.gr@dartmouth.edu

Anne Gelb
Dartmouth College
Department of Mathematics
annegelb@math.dartmouth.edu

Theresa Scarnati
Qualis Corporation
tscarnati@qualis-corp.com

CP12

Multiplicative Denoising with Uncertainty Quantification for Synthetic Aperture Radar Imaging

Traditional methods for removing the multiplicative noise inherent in synthetic aperture radar (SAR) images produce single point-estimate images resulting from the maximum a posteriori optimization of an objective functional or some filtering procedure. We build on previous work introducing a sampling-based approach to SAR image uncertainty quantification by considering the performance of several

difference priors for despeckling piecewise-constant images. Although a Laplace difference prior corresponding to TV regularization of pixel edge differences yields a MAP estimate that is sparse in the edge domain, it fails to similarly promote piecewise constant image samples in the posterior. Furthermore, uncertainty quantification under such a prior is unreliable. Thus we present several choices of priors that promote edge-sparsity in the posterior and permit reliable uncertainty quantification for SAR images. We discuss efficient sampling approaches for this high-dimensional problem, as well as hyperpriors that allow our speckle model to better handle images with edge-domain heterogeneity.

Jonathan Lindbloom
Dartmouth College
jonathan.t.lindbloom.gr@dartmouth.edu

Anne Gelb
Dartmouth College
Department of Mathematics
annegelb@math.dartmouth.edu

Matthew Parno
Dartmouth College
matthew.d.parno@dartmouth.edu

CP12

Uncertainty Quantification in 3D Imaging of Atmospheric Dispersion Processes with Dial

We consider the problem of fitting atmospheric dispersion parameters from time-resolved back-scattered differential absorption Lidar (DIAL) measurements. A clear advantage of optical remote sensing modalities is an extended range which makes them less sensitive to strictly local modelling errors or the distance to a plume source. In contrast to other state-of-the-art DIAL methods, we dont make a single scattering assumption but propose the collection of multiply scattered photons from wider/multiple fields-of-view that can aid in the reconstruction of certain image features. The scattering of photons in heterogeneous media is modelled through the time dependent Radiative Transfer Equation (RTE) which drastically increases the computational complexity compared current DIAL based approaches. Motivated by environmental emergency response applications and the need to solve the problem in nearly real-time, we address this issue by reconstructing an image representation based on dispersion parameters which avoids a high-dimensional inverse problem and regularises the otherwise ill-posed problem. The obtained parameters are directly related to a dispersion model and any point estimate or UQ can be associated with meaningful physical units. This approach not only achieves a high degree of interpretability but has the potential to naturally incorporate the effect of uncertainties in the optical measurements as well as atmospheric quantities on the reconstructed gas concentration.

Robert Lung
University of Edinburgh
robert.lung@ed.ac.uk

Nick Polydorides
Institute for Digital Communications
University of Edinburgh
n.polydorides@ed.ac.uk

CP12

Near-Field Microwave Mapping by Bayesian Inversion-Segmentation

An experimental set-up is dedicated to the microwave control at full scale of the homogeneity of electromagnetic (EM) properties, such as permeability and permittivity. By moving an antenna close to the material, it is possible to measure near-field reflection coefficients for various locations. One consider that the EM properties can be reduced to a surface impedance with possible discontinuities related to material inhomogeneities. It can be linked to the reflection coefficients by a linearized forward model, derived from Maxwell's equations. It includes an additive term corresponding to various errors: measurement, linearization and modeling. Since the number of data available is small compared to the number of unknowns, the direct model is not invertible. Therefore, we have developed a Bayesian inversion method that provides a segmented impedance map. A Gauss-Potts prior relies on a region division by a set of labels, modelled by a Potts field and on the impedance, modeled by piecewise Gaussian fields conditionally to the labels. Estimation is very tricky due to the large dimension and the non-standard form of the posterior. Consequently, we have turned towards a Gibbs algorithm that, in the stationary phase, provides random samples of the a posteriori law, that are used to compute the estimate of the impedance map and labels, and to quantify uncertainties. It is applied to simulated data based on the linearized forward model, derived from a 3D Maxwell solver.

Pierre Minvielle, Julien Boissy
CEA
pierre.minvielle@cea.fr, julien.boissy@cea.fr

Jean-François Giovannelli
Université de Bordeaux, CNRS, IMS
jean-francois.giovannelli@u-bordeaux.fr

CP12

Quantization Applied to the Visualization of Low-Probability Flooding Events

Visualization is essential in the risk assessment of coastal or river flooding. In this work, we deal with expensive-to-evaluate hydraulic simulators, taking as random scalar inputs offshore meteo-oceanic conditions and dyke breach parameters, whereas the output is a flooding map. The challenge is to display a few prototype maps representing at best the probability law of the flooding event, which is a typical quantization problem. The K-Means algorithm classically serves to minimize the expected squared distance between samples of the random event and their representatives. This clustering technique is adapted to handle three key specificities of our context. First, the quantization is done in the specific space of flooding maps, which requires to define an appropriate distance measure. Second, because of the time-consuming simulators, a Gaussian process based approach adapted to spatial outputs is proposed. Third, flooding being a low-probability event, traditional Monte Carlo approaches are inefficient and an Importance Sampling scheme is introduced to generate the maps. The prototype maps represent the distribution of the floodings and are each is associated to a probability mass. To discriminate between the scenarii leading to similar floodings, a second quantization procedure is carried

out in the input space within each cluster. The method is evaluated on an analytical example before being extended to the coastal case of Les Boucholeurs on the French Atlantic coast.

Charlie Sire
Institute for Radiological Protection and Nuclear Safety
charlie.sire@irsn.fr

Le Riche Rodolphe
CNRS LIMOS at Ecole des Mines de St-Etienne
leriche@emse.fr

rohmer Jérémy
BRGM
j.rohmer@brgm.fr

rulliere Didier
CNRS, LIMOS
drulliere@emse.fr

richet Yann, pheulpin Lucie
Institute for Radiological Protection and Nuclear Safety
yann.richet@irsn.fr, lucie.pheulpin@irsn.fr

CP13

Rare Events in Heavy-Tailed Distributions through Large Deviation Theory

The fact that the least unlikely realization of a set of rare events is the minimizer of a function called the *rate function* is at the heart of large deviation theory. The probability of this set of events is dominated by this realization. In heavy-tailed distributions, standard algorithms for computing the minimizer (aka. *instanton*) fail to converge. This failure is caused by the divergence of the *scaled cumulant generating function* as a result of a non-convex large deviation rate function. In this talk, we will present a solution to this problem by 'convexifying' the rate function via *nonlinear reparametrization* of the observable, which allows us to compute instantons even in the presence of super-exponential or algebraic tail decay. We will show how the new formalism works by applying it to rare events in a variety of stochastic systems with fat-tails, such as high power spikes in fibre optics caused by soliton formation.

Mnerh Alqahtani
University of Warwick
UK
M.Alqahtani@warwick.ac.uk

Tobias Grafke
Warwick Mathematics Institute
University of Warwick
T.Grafke@warwick.ac.uk

CP13

Correlated Bernoulli Processes Using De Bruijn Graphs

Some numerical models have two distinct regions in output space where classification is required. E.g. a computer model may fail to complete for specific input regions, and we would like to predict where to avoid running the model, or incorrectly running an emulator. A widely used method for classification is logistic regression, which produces a distribution for the predictive class membership of being in either region. When sampling from this to make predictions, current practice is to draw from an independent Bernoulli distribution; drawing marginally means that any correlation between data is lost and can result in large numbers of misclassifications. If simulating chains/fields of 0s and 1s, it is hard to control the stickiness of like symbols. In this paper, we present a novel approach to generating a correlated Bernoulli process to create chains of 0s and 1s, for which like symbols cluster together. We use the structure from de Bruijn Graphs - a directed graph, where given a set of symbols V and a word length m, the nodes of the graph consist of all possible sequences of V of length m. De Bruijn Graphs are a generalisation of Markov chains, where the word length controls the number of states that each individual state is dependent on. This increases correlation over a wider area. Properties of De Bruijn graphs including expected run length and inference will be presented, as well as an application of modelling the Oxford and Cambridge university boat race.

Louise Kimpton
University of Exeter
l.m.kimpton@exeter.ac.uk

Peter Challenor
College of Engineering, Mathematics and Physical Sciences
University of Exeter
p.g.challenor@exeter.ac.uk

Henry P. Wynn
London School of Economics
h.wynn@lse.ac.uk

Daniel Williamson
University of Exeter
d.williamson@exeter.ac.uk

CP13

Anova Decomposition of Functions with Non-Independent Variables

Hoeffding decomposition of functions (i.e., ANOVA) is widely used in statistical modeling and uncertainty quantification such as variance-based sensitivity analysis. In this abstract, we extend functional ANOVA to cope with non-independent variables by making use of dependency functions of any non-independent variables. The proposed decomposition of functions variances or covariances inherits the properties of the well-known ANOVA, giving us the ability to define new sensitivity indices for models with non-independent variables in such a way that the main-effects and interactions indices sum up to one. Dependent generalized sensitivity indices recently introduced are good approximations of the new sensitivity indices we proposed.

Matieyendou Lamboni
University of Guyane
matieyendou.lamboni@gmail.com

CP13

Quantifying Surrogate Trustworthiness: A Principled Uncertainty Gauge through a Bayesian Approach

Surrogate models are deemed trustworthy, or not, based on heuristic diagnostics. A limitation of this binary view is: i) if the surrogate is not trustworthy, it is obsolete. ii) if the surrogate is trustworthy, then the contribution

of the uncertainty of the surrogate itself to the inferential uncertainty on the quantity of interest (QoI) is neglected. Here, we quantify surrogate uncertainty as a continuous random variable. Then, we can also investigate the inferential uncertainty caused by the surrogate uncertainty itself. We discovered that, most curiously, from this assumption emerges naturally a generalized measure of surrogate trustworthiness that is gauged to an objective scale. This suggests a completely new interpretation of convergence of and uncertainties obtained by surrogates. For generalized linear surrogate models and a Student-t likelihood for the simulation data, we find simple Bayesian estimates for surrogate uncertainty and its effect on QoI uncertainty. Terms are identified as input-parametric uncertainty and inferential uncertainty [Ranftl & von der Linden 2021: Bayesian Surrogate Analysis and Uncertainty Propagation. DOI: 10.3390/psf2021003006]. We discuss the special cases of Polynomial Chaos or Gaussian Processes, and demonstrate a numerical example where surrogate uncertainties are in part negligible and in part non-negligible [Ranftl et al. 2022: A Bayesian approach to Blood Rheological Uncertainties in Aortic Hemodynamics. DOI: 10.1002/cnm.3576].

Sascha Ranftl, Wolfgang Von Der Linden
Graz University of Technology
Institute of Theoretical Physics-Computational Physics
ranftl@tugraz.at, vonderlinden@tugraz.at

CP13

Bayesian Finite Mixture Regression Models with Cluster-Specific Variable Selection

Heterogeneous data are ubiquitous in scientific studies. In regression problems, different subpopulations may differ not only in the effect size of covariates on the response, but also in the subset of covariates that are useful predictors. We propose using Bayesian finite mixture models (FMM) to address the heterogeneity in data, where the number of subpopulations, M, is modeled as a random variable. Notable features of our models are (1) the adoption of a class of priors based on the Normalized Independent Finite Point Process (NIFPP) introduced by Argiento and De Iorio (2019), (2) the inclusion of spike-and-slab components in generating NIFPP priors to achieve variable selection that is specific to each cluster, and (3) the joint modelling of the covariate variables to enable straightforward posterior predictions. We demonstrate improved performance of our model over classical ones, thanks to the more flexible priors and the variable selection feature. For the computation of the proposed Bayesian models, we extend existing MCMC algorithms for NIFPP to perform versatile posterior inferences, such as clustering, individual profiling, and predictions.

Zhen Wang, Aixin Tan
Department of Statistics & Actuarial Science
University of Iowa
zwangiowa@gmail.com, aixin-tan@uiowa.edu

CP14

A Rank and Dimension-Adaptive Integration Scheme with Robustness to Overapproximation for the Dynamical Low-Rank Approximation

Scientific computing is often limited by the available processing power and storage space due to the fine resolution, high dimensionality, or stochastic dynamics of the system. Reduced-order models (ROMs) aim to approximate the full system by optimally truncating the degrees of freedom to both speed up computation and reduce storage space. Dynamical ROMs, such as the dynamical low-rank approximation (DLRA), automatically adjust the low-dimensional representation according to the system dynamics, and the DLRA may be applied to deterministic or stochastic PDEs by assuming a low-rank decomposition in the spatial dimensions, the stochastic space, or a combination of the two. To numerically integrate the DLRA, we propose a high-order alternating least-squares (ALS) integrator with two interpretations: first as a high-order retraction which approximates the singular value decomposition and second as a projector-splitting integrator. Due to the lack of matrix inversions, the ALS scheme is robust to overapproximation; i.e. it is stable when we overestimate the rank of the solution. Furthermore, this scheme may be adapted so that we can dynamically increase and reduce the rank of the solution during computation. The deterministic and stochastic dimensions of the problem may also be changed mid-simulation to accommodate a time-dependent spatial domain or a varying number of stochastic realizations. We illustrate results with examples from computational acoustics and path planning.

Aaron Charous, Pierre F. Lermusiaux
MIT
acharous@mit.edu, pierrel@mit.edu

CP14

Approximate Bayesian Computation with Path Signatures

Simulation models of scientific interest often lack a tractable likelihood function, precluding standard likelihood-based statistical inference. A popular likelihood-free method for inferring simulator parameters is approximate Bayesian computation, where an approximate posterior is sampled by comparing simulator output and observed data. However, effective measures of closeness between simulated and observed data are generally difficult to construct, particularly for time series data which are often high-dimensional and structurally complex. Existing approaches typically involve manually constructing summary statistics, requiring substantial domain expertise and experimentation, or rely on unrealistic assumptions such as independent and identically distributed data. Others are inappropriate in more complex settings like multivariate or irregularly sampled time series data. In this paper, we introduce the use of path signatures as a natural candidate feature set for constructing distances between time series data for use in approximate Bayesian computation algorithms. Our experiments show that such an approach can generate more accurate approximate Bayesian posteriors than existing techniques for time series models.

Joel Dyer
University of Oxford
Alan Turing Institute
joel.dyer@maths.ox.ac.uk

Patrick Cannon
Improbable
cannonpw1@gmail.com

Sebastian Schmon
Durham University
Improbable

sebastian.schmon@durham.ac.uk

CP14

A Solver for Stochastic Galerkin Matrix Equations Associated with Linear Elasticity Problems

We consider the three-field model for linear elasticity with uncertain Youngs modulus. To perform forward uncertainty quantification (UQ), we apply a stochastic Galerkin mixed finite element method which yields stable numerical approximations even in the nearly incompressible case. The associated discrete problems are extremely large and can be formulated in two ways: (i) as a large linear system with saddle-point structure (the so-called Kronecker formulation) and (ii) as a linear multi-term matrix equation. We focus on the multi-term matrix equation formulation, and describe a so-called multi-term reduced basis solver. For the linear elasticity problem, the matrix equation is ill-conditioned not only with respect to the discretization parameters but also with respect to the Poisson ratio. The proposed solution strategy has two components. First, we present two preconditioning strategies and apply shifts to modify the preconditioned matrix equations so that they have more desirable properties. Then, we iteratively construct approximation spaces and apply a projection method to solve a problem of reduced size at each step. Numerical results will also be presented.

Ying Liu
Department of Mathematics
The University of Manchester
ying.liu-2@manchester.ac.uk

Catherine Powell
University of Manchester
catherine.powell@manchester.ac.uk

Valeria Simoncini
Universita' di Bologna
valeria.simoncini@unibo.it

CP14

$\mathcal{L}_2$-Optimal Model Order Reduction for Parametric Stationary Problems

There is a variety of system-theoretic model order reduction (MOR) methods for non-parametric non-stationary systems. For linear time-invariant (LTI) systems, one is the iterative rational Krylov algorithm (IRKA) for $\mathcal{H}_2$-optimal MOR. In this presentation, we discuss adapting it to parametric stationary problems, and in particular, when the parameter is a random variable with a given distribution and the goal is estimating the statistics of the quantities of interest quickly using a reduced-order model.

Petar Mlinaric
Max Planck Institute for Dynamics of Complex Technical Syste
mlinaric@vt.edu

Serkan Gugercin
Virginia Tech
Department of Mathematics
gugercin@vt.edu

CP14

Multi-Index Ensemble Kalman Filtering

In this work we marry multi-index Monte Carlo with ensemble Kalman filtering (EnKF) to produce the multi-index EnKF method (MIEnKF). The MIEnKF method is based on independent samples of four-coupled EnKF estimators on a multi-index hierarchy of resolution levels, and it may be viewed as an extension of the multilevel EnKF (MLEnKF) method developed by the same authors in 2020. Multi-index here refers to a two-index method, consisting of a hierarchy of EnKF estimators that are coupled in two degrees of freedom: time discretization and ensemble size. Under certain assumptions, the MIEnKF method is proven to be more tractable than EnKF and MLEnKF, and this is also verified in numerical examples.

Gaukhar Shaimerdenova
King Abdullah University of Science and Technology
SRI Center for Uncertainty Quantification in Computati
gaukhar.shaimerdenova@kaust.edu.sa

Hakon Hoel
RWTH Aachen Univeristy
haakonah1@gmail.com

Raul F. Tempone
Mathematics, Computational Sciences & Engineering
King Abdullah University of Science and Technology
raul.tempone@kaust.edu.sa

CP15

Robust Decision-Making under Risk and Ambiguity

Economists often estimate economic models on data and use the point estimates as a stand-in for the truth when studying the models implications for optimal decision making. This practice ignores model ambiguity, opens the door for misspecification of the decision problem, and leads to post-decision disappointment. We develop a framework to explore, evaluate, and optimize robust decision rules that explicitly account for the uncertainty in the estimation using statistical decision theory. We show how to operationalize our analysis by studying robust decisions in a stochastic dynamic investment model in which a decision-maker directly accounts for uncertainty in the models transition dynamics.

Max Blesch
Berlin School of Ecoomics
maximilianblesch@web.de

Philipp Eisenhauer
University of Bonn
peisenha@protonmail.com

CP15

Challenges in Industrial Applications of Value of Information Analysis for Risk Optimal Data Collection

Risk management of the built environment is transitioning towards methods of data-centric engineering. The emergence of structural health monitoring systems, and scalable Bayesian inference will allow for more sophisticated uncer-

tainty quantification of damage and degradation. This information could then inform digital twin representations of structural systems, which would provide improved decision support to engineers. Understanding the required quantity and quality of data will continue to be a challenge to engineers. Should sensing systems be retrofit to existing structures? If so, how precise and how reliable do they need to be? Does a malfunctioning sensor require replacement? What supplementary inspection or testing data is required? This lecture discusses how methods with a basis in Bayesian experimental design can quantify the expected value of such (imperfect) data collection activities, in the context of supporting risk management decisions. Computational challenges associated with this analysis include the requirement to consider multivariate prior domains of possible measurement outcomes, where decision boundaries complicate the possible use of variance reduction methods (such as importance sampling). Conceptual challenges include the need to define the sources of utility associated with the reduction in epistemic uncertainty, some of which may not be immediately apparent. These key barriers to wider industrial application are critically considered.

Domenic J. Di Francesco
The Alan Turing Insititute
ddifrancesco@turing.ac.uk

Mark Girolami
Imperial College London
mag92@eng.cam.ac.uk

CP15

On the Design of Outlier Exposure-Based Out-of-Distribution Detectors

Deep learning methods are routinely adopted to build predictive models with complex data and to guide crucial decision-making in several real-world applications. However, in practice, predictions from AI models can be poorly calibrated thus producing unreliable decisions. This emphasizes the need to ensure that models are calibrated to detect OOD samples while not trading off on in-distribution performance. Outlier Exposure (OE) is a popular strategy for building well-calibrated models by using an auxiliary dataset of outliers during training. Conceptually, OE encourages the model to produce low-confidence predictions for OOD data, wherein the confidence is characterized based on suitable uncertainty estimates. In this work, we systematically study the design of deep models with OE, namely, (i) the confidence estimator choice such as prediction entropy, loss estimates, energy etc.; (ii) the OOD dataset; and (iii) input augmentations such as mixup, semantics-preserving augmentations etc. Using case studies in dermatology and histopathology applications, we will present a rigorous evaluation suite and benchmark the performance of the different design choices for OE-based out-of-distribution detection.

Vivek Narayanaswamy
Arizona State University
vnaray29@asu.edu

Jayaraman J. Thiagarajan
LLNL
jjayaram@llnl.gov

Yamen Mubarka, Rushil Anirudh
Lawrence Livermore National Laboratory
mubarka1@llnl.gov, anirudh1@llnl.gov

CP15

Multi-Marginal Optimal Transport and Multi-Population Matching Beyond Discrete Measures

We consider the multi-marginal optimal transport problem with general measures that are not necessarily discrete. We develop a relaxation scheme accompanied by duality results to approximate such a problem by a linear semi-infinite optimization problem, where the approximation error can be controlled. The developed relaxation scheme leads to a numerical framework for solving multi-marginal optimal transport problems. Specifically, we are able to compute both an upper bound and a lower bound on the multi-marginal optimal transport problem, and the gap between the bounds provides an explicit estimate of the approximation error. Using this numerical framework, we develop cutting-plane algorithms for the class of multi-population matching problems introduced by [Carlier and Ekeland, Matching for teams, 2010], which contains the well-known Wasserstein barycenter problem as a special case. Numerical experiments demonstrate that the proposed algorithms are capable of computing high-quality solutions of multi-population matching problems along with explicit estimates of the approximation errors that are much less conservative compared to the theoretical estimates. Finally, we discuss how the relaxation scheme and the numerical framework can be applied to distributionally robust optimization problems with dependence uncertainty.

Qikun Xiang
School of Physical and Mathematical Sciences
Nanyang Technological University
qikun001@e.ntu.edu.sg

Ariel Neufeld
Nanyang Technological University
Singapore
ariel.neufeld@ntu.edu.sg

CP15

The Back-and-Forth Method on Multi-Marginal Optimal Transport with Applications in Sea Ice Prediction

Climate change is real and is causing fundamental changes in the seasonal behavior of sea ice in the Arctic. Given a temporal sequence of images on sea ices with partial information, we want to recover the underlying dynamics. The use of multi-marginal optimal transport (MMOT) is the key in our approach and is motivated by: first, the dynamics following optimal transport minimize the average kinetic energy in the Brenier's theory; second, MMOT can fuse information from the whole temporal sequence rather than from each single image in pixel-wise sense. However the computation of OT can be expensive and the popular regularized Wasserstein metric may introduce additional blurring. We generalize the back-and-forth method on computing the accurate OT for two marginals to multiple marginals, and illustrate by examples the faithful joint recover of images with sharp boundaries, with the applications of sea ices.

Bohan Zhou, Matthew Parno
Dartmouth College
Bohan.Zhou@dartmouth.edu,

matthew.d.parno@dartmouth.edu

CP16

Performant Global Sensitivity Analysis Using GlobalSensitivity.jl

With increasing adoption of Global Sensitivity Analysis as part of modelling workflow in various domains with large models such as climate science or quantitative systems pharmacology, there is an urgent need for optimized and scale able GSA implementations. Such optimization can come through both algorithmic improvements, and through utilization of modern and extensible scientific computing stack. GlobalSensitivity.jl is a generalized GSA package written in the Julia programming language which makes it capable of handling varied problems due to composability offered by julia. The built-in support for parallelism allows analysis of large models with significant simulation overhead with ease for domain scientists looking to use GSA. Currently GlobalSensitivity.jl supports the Sobol, Morris, eFAST, Regression based, DGSM, Delta Moment, EASI, Fractional Factorial and RBD-FAST GSA methods. This talk will cover a comprehensive tutorial of running various different GSA methods and analysing their results using visualizations on the Lotka-Volterra differential equation model using SciML's DifferentialEquantions.jl package. Further, there will be a focus on demonstrating use of GSA in Pharmacometrics by analyzing some example PK/PD, PBPK and QsP models. For this purpose the talk will include tutorials focused on using Pumas for running GSA and post processing results to derive insights on a PK/PD model of Hepatitis-C Virus (HCV) and a PBPK model for Voriconazole.

Vaibhav K. Dixit
Julia Computing
vaibhavyashdixit@gmail.com

Vijay Ivaturi
Pumas-AI and University of Maryland, Baltimore
vijay@pumas.ai

Chris Rackauckas
Julia Computing and MIT
chris.rackauckas@juliacomputing.com

CP16

High Dimensional Sensitivity Analysis Method for a Computationally Expensive Code

Calculation tools are commonly used to understand and predict physical phenomena. Often, these simulation codes uses many input parameters to characterize the studied phenomenon. In this context, a sensitivity analysis can be performed in order to study the behavior of these codes with respect to the variations of the uncertain inputs. However, the large CPU-time cost of some codes limits the sample size of simulations and requires the use of appropriate strategies to perform this sensitivity analysis. To comply with these constraints, a specific methodology has been developed, built upon two main steps. First, a screening of influential inputs is performed using two recent global sensitivity tools: first order Sobol indices with new estimators based on rank statistics, and the Hilbert-Schmidt Independence Criterion which relies on covariance operators in reproducing kernel Hilbert spaces. Then, from screening results, the outputs of interest are approximated by Gaussian process metamodel. Higher order and total Sobol indices can therefore be estimated with the obtained metamodels for a more accurate sensitivity analysis. The development of this methodology is carried out in the context of severe nuclear accidents. Tools are applied on a simulator of fuel-coolant interaction that can lead to the steam explosion phenomenon. The complexity of the studied phenomenon leads to many uncertain inputs (60 parameters), and a long CPU-time(12 hours) for each code run.

Faouzi Hakimi
Paul Sabatier University-Toulouse III
CEA-DES/DTN/SMTA/LMAG
faouzi.hakimi@cea.fr

Claude Brayer, Benot Habert
CEA-DES/DTN/SMTA/LMAG
claude.brayer@cea.fr, benoit.habert@cea.fr

Fabrice Gamboa
Institut de Mathématiques de Toulouse. AOC Project
University of Toulouse
fabrice.gamboa@math.univ-toulouse.fr

Amandine Marrel
CEA Cadarache, France
amandine.marrel@cea.fr

CP16

Dynamic Mode Decomposition for Flow and Transport Problems

Dynamic mode decomposition (DMD) is a powerful data-driven technique for construction of reduced-order models (ROMs) of complex dynamical systems. Despite its popularity, DMD and other singular-value decomposition (SVD) based techniques (e.g., POD) struggle to formulate accurate ROMs for advection-dominated problems because of the nature of SVD-based methods. We investigate this shortcoming of conventional POD and DMD methods formulated within the Eulerian framework. Then we propose a Lagrangian-based DMD method to overcome this so-called translational problem. Our approach is consistent with the spirit of physics-informed DMD since it accounts for the evolution of characteristic lines. Furthermore, we address the limitation of Lagrangian DMD in hyperbolic problems with shocks and propose a physics-informed DMD based on hodograph transformation. This strategy is consistent with the spirit of physics-aware DMDs in that it retains information about shock dynamics. Several numerical tests are presented to demonstrate the accuracy and efficiency of the proposed methods.

Hannah Lu
Stanford University
hannahlu@stanford.edu

CP16

Polynomial Chaos Methods for Nonlinear Partial Differential Equations with Correlated Gaussians Parameters

In uncertainty quantification, polynomial chaos is a non-sampling technique used to approximate the solution of a stochastic partial differential equation. The solution is estimated by a polynomial expansion truncated to finitely many terms, whose deterministic coefficient functions are recovered through Galerkin projections. In the presence of multiple uncertainties, the projection step introduces products (nth order moments) of the basis polynomials, where n is the degree of nonlinearity of the governing PDE. When

these uncertainties are represented by correlated random variables, there is no closed-form expression for these products, even when the uncertainties have a joint Gaussian distribution. Consequently, the products are typically computed via sampling methods, which can: (a) become computationally expensive to implement, and (b) introduce errors if the sample count is insufficient. In recent work, a new expression was found for the simple and efficient evaluation of the double products (second order moments) of the basis polynomials (multivariate Hermite) with correlated Gaussian inputs. We present a formula for the nth order product in terms of the double product computations. This formula will allow polynomial chaos methods to be more readily applied to solving stochastic nonlinear PDEs with correlated Gaussian parameters.

Laura A. Lyman
Stanford University
lymanla@stanford.edu

Gianluca Iaccarino
Stanford University
Mechanical Engineering
jops@stanford.edu

CP16

Multifidelity Polynomial Chaos Expansion Using Leja Grid Points

The high computational cost of the models is one of the major hurdles for performing the forward uncertainty quantification (UQ) analysis. To overcome this problem, we use the multifidelity modelling within the forward UQ setup. We combine the information from different hierarchies of fidelities to obtain the statistical moments of the highest fidelity model by consuming low computational resources. We use polynomial chaos expansion (PCE) to obtain the statistical moment. In order to build the PCE, we use the low fidelity function and a discrepancy function. The discrepancy function models the transformation done on the low-fidelity model to convert it to a high fidelity model. We model the transformation in the form of a linear combination of an addition term and a multiplication term. We perform this operation recursively on each level, to finally obtain the PCE of the highest fidelity model. We use the sparse combination technique to combine polynomials of different orders. We adaptively choose the model evaluation points based on the surplus of the variance. We also use the leja grid points to further decrease the model evaluations. Finally, we show using different examples that we can efficiently obtain the statistical moments using our proposed method.

Kislaya Ravi
MSc. Computational Science and Enginnering,
Currently, doing PhD. at Techinal Univ. Munich
kislayaravi@tum.de

Tobias Neckel
Technical University of Munich
neckel@in.tum.de

Hans Joachim Bungartz
Technical University Munich
bungartz@tum.de

CP16

Likelihood Free SAMC

Approximate Bayesian Computation (ABC) has become a valuable tool for Bayesian Uncertainty Quantification, as it enables inference to be made even when the likelihood is intractable. ABC methods can produce unreliable inference when they introduce high approximation bias into the posterior through careless specification of the ABC kernel. Additionally, MCMC-ABC methods often suffer from the local trapping problem which causes poor mixing when the tolerance parameter is low. We introduce a new ABC algorithm, the Stochastic Approximation Monte Carlo ABC (SAMC-ABC), which enables Bayesian Uncertainty Quantification in increasingly complex systems where inference was previously unreliable. SAMC-ABC adaptively constructs the ABC likelihood kernel, both reducing the approximation bias and providing immunity to the local trapping problem. We demonstrate the performance of the proposed algorithm with some benchmark examples and find that the method outperforms its competitors. We use our algorithm to analyze a computer model which describes the transmission of the Ebola virus against data from the 2014-15 Ebola outbreak in Liberia.

Kieran Richards, Georgios Karagiannis
University of Durham
kieran.96@outlook.com, georgios.stats@gmail.com

Guang Lin
Purdue University
guanglin@purdue.edu

CP17

Data-Driven Model Correction for Complex Dynamical Systems with Application in Sea Ice Modeling

We propose a numerical method to correct existing models from measurement data using deep neural networks. Existing models such as PDEs are often used to predict system dynamics, but they may be inaccurate and have discrepancies with measurement data. Deep neural networks correct the existing system and compensate for the model error. We validate the proposed method on a set of numerical tests including linear and nonlinear hyperbolic PDEs in addition to an application in sea ice.

Zhen Chen
Dartmouth College
zhen.chen@dartmouth.edu

Yoonsang Lee, Anne Gelb
Dartmouth College
Department of Mathematics
yoonsang.lee@dartmouth.edu, annegelb@math.dartmouth.edu

CP17

Expert Elicitation and Opinion Aggregation for Bayesian Prior Construction with Application to the U.S. Steel Flow Analysis

The elicitation of expert knowledge in the form of the subjective probability distribution is an important aspect of prior construction in Bayesian statistical inference.

However, phycologists have identified several issues that may cause systematic errors during elicitation. Individual knowledge may be influenced by heuristics and biases including anchoring, availability, and overconfidence. Combining multiple experts judgments through behavioral aggregation could be affected by pressure for consensus and strong personality within the group of experts. A systematic and mathematical approach to aggregating experts information can avoid these potential challenges, where experts are also scored by performance on empirical data such as seeding questions. However, these performance scores also need to be validated, which we achieve through out-of-sample validation. This talk describes a recent application of expert elicitation for Bayesian prior construction in the U.S. steel flow analysis.

Jiayuan Dong, Jiankan Liao, Daniel Cooper, Xun Huan
University of Michigan
jiayuand@umich.edu, jkliao@umich.edu, drcooper@umich.edu, xhuan@umich.edu

CP17

Numerical Studies of Bayesian Quadrature Applied to Offshore Wind Turbine Load Estimation

Offshore wind turbine (OWT) new technologies tend to reach for more difficult and uncertain environmental conditions. This industry needs probabilistic tools to manage risks associated to OWT operation and maintenance. The OWT behavior is computed by a transient costly multiphysics numerical simulation code developed by EDF R&D and deployed on a high performance computer. To propagate the various sources of uncertainties through such numerical models, identifying the best sample to estimate a specific quantity of interest is of prime importance. Historically, sampling methods such as low-discrepancy sequences were proven to improve the Monte Carlo reference convergence rate. An alternative strategy is to emulate the costly function by a regression model. For instance, using a Gaussian process regression model provides an estimation of the regression error represented by the variance of the Gaussian process conditioned to the learning sample. This property is widely exploited by adaptive methods for optimization, rare event and quadrature estimation to iteratively pick samples with respect to a specific goal. The aim of this work is to perform a numerical comparison between various adaptive quadrature methods to estimate the expected value of the mechanical loads of an OWT over environmental random variables. Additionally, theoretical equivalences between Bayesian quadrature and Kernel Herding using Maximum Mean Discrepancy shall be verified on an industrial use case.

Elias Fekhari, Matteo Capaldo, Vincent Chabridon, Bertrand Iooss
EDF R&D
elias.fekhari@edf.fr, matteo.capaldo@edf.fr, vincent.chabridon@edf.fr, bertrand.iooss@edf.fr

CP17

Uncertainty Quantification in Hierarchical Vehicular Flow Models

Although traffic models have been extensively studied, obtaining reliable forecast from these models is still challenging, since the evolution of traffic is also exposed to the presence of uncertainties. In this talk, we will investigate the propagation of uncertainties in traffic flow models, especially we focus on kinetic models of BGK type, which allow to derive a hierarchy of traffic models including a hydrodynamic description by considering different spatial and temporal scales. The kinetic BGK–model is extended by introducing a parametric stochastic variable to describe possible uncertainty in traffic. The interplay of uncertainty with the given model hierarchy is studied in detail. Theoretical results on consistent formulations of the stochastic differential equations on the hydrodynamic level are given. The effect of the possibly negative diffusion in the stochastic hydrodynamic model is studied and numerical simulations of uncertain traffic situations are presented.

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

CP17

Optimization Under Uncertainty of a 3-D Magnetohydrodynamic Generator

Magnetohydrodynamics (MHD) is the study of an electrically-conductive medium flowing through a magnetic field. By artificially generating and accelerating a plasma under a strong magnet, an electric field and current spontaneously arise. Doing so within a channel, with electrodes strategically placed and loaded, the power within the channel can be harvested. The governing equations for this multi-physics problem consists of Navier-Stokes, Maxwell's equations, and the generalized Ohm's law. We choose to focus on the kinematic MHD equations for optimization purposes, and thus neglect the fluid equations, and assume a stationary system. Within MHD generators, the angle between electrodes, and the load being placed on the channel, dramatically affect performance. The optimal choice of these parameters depend upon the given state of the system, implying that the optimal parameter set may change based on current conditions. Thus, choosing these optimal parameters within an operational generator must be done from estimates of the state parameters, which must include uncertainty due to aleatoric noise and the indirect measurements of the state parameters. With this in mind, we propose an optimization scheme that utilizes the stochastic collocation method to maximize the expected power of the MHD generator. Numerical implementation of this method is discussed, concluding with demonstrations of the feasibility of the optimization under uncertainty scheme.

Evan Rajbhandari, Nathan L. Gibson
Oregon State University
rajbhane@oregonstate.edu, gibsonn@oregonstate.edu

Rigel Woodside
National Energy Technology Laboratory
rigel.woodside@netl.doe.gov

CP18

Surrogate Modeling for High-Dimensional Engineering Problems

In engineering analysis, complex physics models are often replaced with surrogate models in order to achieve computational efficiency when the physics models need to run multiple times. The quality and quantity of data collected from the expensive physics model is crucial to the accuracy of the surrogate models. We present a novel approach to construct surrogate models for high-dimensional engineering problems. Methods for dimension reduction for both the input and output are investigated: variance-based sen-

sitivity analysis and active subspace discovery for the input space; singular value decomposition (SVD), random projection, randomized SVD, and diffusion map for the output space. The most effective combination of options for input and output dimension reduction is identified, and Gaussian process surrogate models are constructed in the low-dimensional space. The predictions of the quantities of interest in the original high-dimensional space are obtained using the surrogate models. The errors associated with the predictions consist of surrogate errors and reconstruction errors, and a systematic approach is developed to quantify and compare the relative contributions of the two types of errors. An analysis on an aircraft fuselage panel is used to compare various dimension reduction techniques for surrogate model construction for high-dimensional problems, in terms of both accuracy and computational effort.

Yulin Guo, Sankaran Mahadevan
Vanderbilt University
yulin.guo@vanderbilt.edu,
sankaran.mahadevan@vanderbilt.edu

Shunsaku Matsumoto, Shunsuke Taba, Daigo Watanabe
Mitsubishi Heavy Industries
shunsaku.matsumoto.zj@mhi.com, shunsuke.taba.du@mhi.com, daigo.watanabe.d5@mhi.com

CP18

Multi-Fidelity Uncertainty Quantification in Additive Manufacturing

Sophisticated additive manufacturing (AM) models are accurate but computationally expensive, whereas analytical models with simplified physics are fast but have significant prediction errors. This work presents a Bayesian approach for constructing a multi-fidelity prediction and uncertainty quantification model by fusing information from physics-based models of different fidelity and experimental data. A surrogate model is first constructed to replace the LF model. Variance-based sensitivity analysis is then carried out using the surrogate model to identify the model parameters to be calibrated. The lower fidelity (LF) model is corrected in two stages: first using the higher fidelity (HF) model simulation results and then the experimental data. Bayesian calibration is used to calibrate the correction factors and the LF model parameters to account for and quantify the uncertainty in the process. The proposed methodology is demonstrated for the laser powder bed fusion AM process. The HF multiphysics model, which accounts for heat transfer, fluid flow, phase change, laser-material interactions, etc., and the LF approximate heat conduction model are combined to predict the lack-of-fusion porosity in an additively manufactured part.

Paromita Nath, Matthew Sato, Sankaran Mahadevan
Vanderbilt University
paromita.nath@vanderbilt.edu,
matthew.m.sato@vanderbilt.edu,
sankaran.mahadevan@vanderbilt.edu

CP18

Title of the Paper Sequential Experimental Design for Materials Strength Model Calibration

Due to the time and expense associated with physical experiments, the next experimental conditions should be chosen optimally to inform parameter estimation. For lower-strain-rate experiments related to material strength, stress-strain curves are obtained as experimental data. We seek to calibrate the material strength models, and we employ mutual information, based on Shannon entropy, to select which experiment should be performed to achieve the greatest reduction in strength model parameter uncertainty. Moreover, we adapt the existing framework to handle the functional output of the strength experiments. LLNL-ABS- 827713 This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Kathleen Schmidt
Lawrence Livermore National Lab
schmidt41@llnl.gov

CP18

Reliability and Risk Metrics to Assess Adequacy and Flexibility of Bulk Power Systems

Bulk power systems (BPS) need to deal with supply and demand uncertainty, as well as the unplanned and sudden loss of system elements (generators, transmission lines, etc.). In recent years, increasing participation of renewable energy sources (RES) has significantly increased the supply side uncertainty and therefore increased BPS vulnerability to inadequate and inflexible power supply. In this work, we present reliability and risk metrics to evaluate the adequacy and flexibility of BPS market clearing solutions for unit commitment and economic dispatch. The proposed approach could either support the evaluation of solutions from existing deterministic approaches - which do not explicitly consider uncertainty or risk - or be incorporated in a fully stochastic optimization for market clearing. We define three levels of system assessment metrics that consider conditional expectation, probability of failure, and risk. The first two metrics can be used for reliability assessment, whereas the third metric considers the (monetary) consequence to evaluate the risk. We demonstrate the computation of these metrics to assess reliability and risk for a 200-bus synthetic grid.

Oliver T. Stover
Vanderbilt University
Civil Engineering
oliver.stover@vanderbilt.edu

Pranav Karve
vanderbilt
pranav.m.karve@vanderbilt.edu

Sankaran Mahadevan
Vanderbilt University
sankaran.mahadevan@vanderbilt.edu

CP18

Uncertainty Quantification for PDEs on Graphs and Applications to Simulations of Gas Networks

In this talk we will present a new uncertainty quantification (UQ) method that enables accurate uncertainty quantification (UQ) paradigm for energy networks to characterize input/output relationships of energy supply and delivery variability in both time and space. Our approach is based on semi-intrusive Stochastic Finite Volume (SFV) method to quantify the uncertainties arising due to random model coefficients in our underlying hyperbolic PDE that models the gas flow. The SFV method requires some modifications of the deterministic code which however only involve additional integration of the numerical fluxes over the cells

in the stochastic space and are therefore considered mild. This approach preserves the hyperbolicity of the model, and at the same time is more computationally efficient than e.g. Monte Carlo method. We then extend the SFV method to perform uncertainty quantification on a graph of PDEs and apply the method to gas networks. A crucial ingredient for the numerical modelling of gas pipe networks is an accurate treatment of physical constraints at pipe junctions. These constraints include, for example continuity of pressure and conservation of flow. The numerical fluxes at junctions are obtained by solving a Junction-Generalized Riemann problem. We demonstrate the results of our SFV approach first for a single pipe and then for a test network of gas pipes.

Svetlana Tokareva
Los Alamos National Laboratory
Applied Mathematics and Plasma Physics Group (T-5)
tokareva@lanl.gov

Anatoly Zlotnik, Vitaly Gyrya
Los Alamos National Laboratory
azlotnik@lanl.gov, vitaliy_gyrya@lanl.gov

MS1

Hypothesis Testing for Spatial Fields Generated by Climate Models

Climate model evaluation and diagnosis often involves comparing spatial fields. The fields might be two different model runs, as is the case in various model intercomparison projects (MIPs), or one model-generated field and one derived from observational data. Traditionally, such comparisons are based on simple descriptive statistics such as root-mean-squared error, and are only relative: there is no probability model that can be used to conduct a hypothesis test for determining whether the differences are significant. To accomplish that, one needs to generate ensembles of these statistics, under a suitably defined null-hypothesis on spatial structure. Here, we explore the use of Kernel Flows (KF; Owhadi and Yoo, 2019) to fit a spatial model to a climate-model-generated spatial field (the parent), and simulate statistical replicates from it. KF is a computationally efficient algorithm for estimating parameters from large data sets. In effect, we exploit internal variability of the parent to create new, random realizations that preserve spatial structure in a statistical sense. A test statistic that quantifies the difference between two fields can then be computed over the ensemble to yield a null distribution. In this talk, we demonstrate the idea using climate model runs from the CMIP6 archive.

Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology
Amy.Braverman@jpl.nasa.gov

Emily L. Kang
Department of Mathematical Sciences
University of Cincinnati
kangel@ucmail.uc.edu

Pulong Ma
Clemson University

Houman Owhadi
Applied Mathematics
Caltech
owhadi@caltech.edu

Jouni Susiluoto
Jet Propulsion Laboratory
jouni.i.susiluoto@jpl.nasa.gov

Snigdhansu Chatterjee
University of Minnesota
Department of Statistics
chatt019@umn.edu

MS1

In-Situ Spatial Inference and Extreme Value Modeling of Climate Data in E3SM

As extreme-scale climate simulation becomes increasingly memory and storage expensive, the ability to access full simulation data for statistical analysis is becoming increasingly limited. The capability to perform in-situ statistical inference of state variables is becoming critical for leveraging these huge amounts of data. In this work, we report the results fitting scalable Gaussian process regression to the state information of the E3SM climate model in-situ. We present both spatial regression of near-surface temperature data and modeling of temperature extremes was performed using the Julia programming language coupled to the E3SM simulation. The resulting inference model uses distributed, sparse Gaussian processes for capturing spatial variation, showing strong predictive performance using a small number of representative observations. These results provide the backbone for more general in-situ spatial inference with Gaussian process models in complex physics simulations.

Michael Grosskopf, Earl Lawrence, Ayan Biswas, Soumya Dutta, Luke Van Roekel
Los Alamos National Laboratory
mikegros@lanl.gov, earl@lanl.gov, ayan@lanl.gov, sdutta@lanl.gov, lvanroekel@lanl.gov

Nathan Urban
Brookhaven National Laboratory
nurban@bnl.gov

MS1

On the Extreme Event Probability Estimation of Land Ice Melting

We explore forward uncertainty quantification methods that rely on optimization to estimate the probability of certain levels of sea level change due to ice mass loss/gain. The uncertainty is due to our imperfect knowledge of the present state of ice sheets, the governing equations and future climate forcing. The proposed UQ methods build on ideas from PDE-constrained optimization and the probability estimation of tail probabilities. Standard Monte Carlo methods may require an infeasible large number of expensive forward simulations to obtain useful estimates. Being able to replace such sampling with an optimization problem is likely to be computationally favorable.

Kim Liegeois
Sandia National Laboratories
knliege@sandia.gov

Mauro Perego
CCR Sandia National Laboratories
mperego@sandia.gov

Georg Stadler

Courant Institute for Mathematical Sciences
New York University
stadler@cims.nyu.edu

MS1

Spatio-Temporal Land Model Calibration Using Karhunen-Loeve Expansions

Computationally expensive climate models challenge ensemble-intensive studies such as parameter estimation, uncertainty quantification, and experimental design. To make such studies tractable, we build accurate surrogate approximations to map from input parameters to output quantities of interest. The model of interest is the land component of the Energy Exascale Earth System Model (E3SM). Besides the large number of input parameters, calibrating the model is challenged by the high-dimensional, spatio-temporal output fields of model evaluations and observational data. We employ Karhunen-Loeve (KL) expansions to represent the spatio-temporal outputs in terms of a small number of eigenmodes. Each KL coefficient is approximated by parametric surrogate models such as neural networks. The resulting spatio-temporal surrogate is then employed in Bayesian calibration of E3SM land model (ELM) parameters using gridded observation datasets. The observational data is also projected to the KL eigenmodes to construct likelihood functions with respect to the de-correlated features. These likelihood functions are further enhanced by embedded statistical error terms representing model structural errors. We will demonstrate the full workflow of ELM surrogate construction, calibration and uncertainty attribution to components due to surrogate errors, parametric uncertainty and model structural errors, and how this information advances predictive understanding of the Earth system.

Daniel Ricciuto
Oak Ridge National Laboratory
ricciutodm@ornl.gov

Khachik Sargsyan, Cosmin Safta
Sandia National Laboratories
ksargsy@sandia.gov, csafta@sandia.gov

MS2

Multi Agent Reinforcement Learning for Complex Systems

The modeling of turbulent flows is critical to scientific and engineering problems ranging from aircraft design to weather forecasting and climate prediction. Over the last sixty years numerous turbulence models have been proposed, largely based on physical insight and engineering intuition. Recent advances in machine learning and data science have incited new efforts to complement these approaches. To date, all such efforts have focused on supervised learning which, despite demonstrated promise, encounters difficulties in generalizing beyond the distributions of the training data. In this work we introduce multi-agent reinforcement learning (MARL) as an automated discovery tool of turbulence models. We demonstrate the potential of this approach on Large Eddy Simulations of homogeneous and isotropic turbulence using as reward the recovery of the statistical properties of Direct Numerical Simulations. Here, the closure model is formulated as a control policy enacted by cooperating agents, which detect critical spatio-temporal patterns in the flow field to estimate the unresolved sub-grid scale (SGS) physics. The present results are obtained with state-of-the-art algorithms based on experience replay and compare favorably with established dynamic SGS modeling approaches. Moreover, we show that the present turbulence models generalize across grid sizes and flow conditions as expressed by the Reynolds numbers.

Petros Koumoutsakos
Harvard University
petros@seas.harvard.edu

MS2

Kronecker Product Dual Layers for Deep Learning

We introduce Kronecker dual layers (KDL) as a novel class of architectures for deep neural networks. Our approach differs from existing deep learning paradigms that exercise Kronecker product structure: we formulate new types of architectures that implicitly compress weight and bias matrices using Kronecker product approximations. Training due to backpropagation with these architectures is substantially accelerated using a KDL framework compared to analogous fully connected layers, and we demonstrate on real datasets that KDL neural networks are more efficient at prediction as well.

Akil Narayan, Jarom Hogue, Robert Kirby
University of Utah
akil@sci.utah.edu, jdhogue@sci.utah.edu, mkirby@cs.utah.edu

MS2

Active Neuron Least Squares: A Training Method for Rectified Neural Networks

In this talk, we will present the Active Neuron Least Squares (ANLS), an efficient training algorithm for neural networks (NNs). ANLS is designed from the insight gained from the analysis of gradient descent training of NNs, particularly, the analysis of Plateau Phenomenon. The core mechanism is the option to perform the explicit adjustment of the activation pattern at each step, which is designed to enable a quick exit from a plateau. The performance of ANLS will be demonstrated and compared with existing popular methods in various learning tasks ranging from function approximation to solving PDEs.

Yeonjong Shin
Brown University
yeonjong_shin@brown.edu

Mark Ainsworth
Division of Applied Mathematics
Brown University
Mark_Ainsworth@brown.edu

MS2

The Approximation Theory of Shallow Neural Networks

A shallow neural network is a linear combination of ridge functions whose profile is determined by a fixed activation function. We will introduce spaces of functions which can be efficiently approximated by shallow neural networks for a wide variety of activation functions and analyze their properties. Specifically, we will compute their metric entropy and n-widths, which are fundamental quantities in approximation theory that control the limits of linear and non-linear approximation and statistical estimation for a

given class of functions. Consequences of these results include: the optimal approximation rates which can be attained for shallow neural networks, that shallow neural networks dramatically outperform linear methods of approximation, and even that shallow neural networks are optimal among all non-linear methods on these spaces, if continuity of the non-linear method is required. Finally, we discuss algorithmic aspects of approximation by shallow networks. Specifically, we analyze a class of greedy algorithms and show that they can attain the theoretically optimal approximation rates.

Jonathan Siegel
Penn State University
jus1949@psu.edu

MS3

On Graphical Gaussian Processes for Highly Multivariate Spatial Data

For multivariate spatial (Gaussian) process models, common cross-covariance functions do not exploit graphical models to ensure process-level conditional independence among the variables. This is undesirable, especially for highly multivariate settings, where popular cross-covariance functions such as the multivariate Matrn suffer from a "curse of dimensionality" as the number of parameters and floating point operations scale up in quadratic and cubic order, respectively, in the number of variables. We propose a class of multivariate "graphical Gaussian Processes" using a general construction called "stitching" that crafts cross-covariance functions from graphs and ensure process-level conditional independence among variables. For the Matrn family of functions, stitching yields a multivariate GP whose univariate components are exactly Matrn GPs, and conforms to process-level conditional independence as specified by the graphical model. For highly multivariate settings and decomposable graphical models, stitching offers massive computational gains and parameter dimension reduction. We demonstrate the utility of the graphical Matrn GP to jointly model highly multivariate spatial data using simulation examples and an application to air-pollution modelling. This is a joint work with Debangan Dey and Abhirup Datta (Johns Hopkins University).

Sudipto Banerjee
Department of Biostatistics
UCLA
sudipto@ucla.edu

MS3

Efficient Uncertainty Propagation and Estimation in Microscopic Imaging Analysis

In the first half of the talk, we will discuss uncertainty quantification and estimation in differential dynamic microscopy (DDM-UQ), a Fourier-based image analysis tool to extract physical information of dynamical properties, including the intermediate scattering function, mean squared displacement and bulk modulus. Despite its straightforward analysis, DDM has not been fully adopted as a routine characterization tool, largely due to computational cost and lack of algorithmic robustness. We present statistical analysis that propagates the uncertainty of the imaging noise through the Fourier analysis, reduces the computational order by Gaussian process regression, and enhances the robustness of the analysis. In the second part of the talk, we will introduce our recent experimental study of dynamic processes of cellular alignment in a system of human dermal fibroblasts, when placed on a liquid crystal elastomer with a molecularly aligned, crosslinked, orientational field.

Mengyang Gu
Department of Statistics and Applied Probability
University of California, Santa Barbara
mengyang@pstat.ucsb.edu

MS3

In Situ Uncertainty Quantification

The Department of Energys investment in exascale computing will enable simulations with unprecedented resolution. This will let scientists investigate fine-scale behavior in areas of interest to DOE such as climate and space physics. However, the computational power of exascale machines ha outstripped their I/O and storage capacity which will make some forms of post hoc analysis impossible. To address this, we are working on methods for in situ uncertainty quantification, that is analysis done inside the simulations as they are running. I will provide an overview of the problem and describe some of the work that we are doing at LANL to fit Bayesian hierarchical models to data inside of simulations of climate and space weather.

Earl Lawrence
Los Alamos National Laboratory
earl@lanl.gov

MS3

Deep Gaussian Processes for Uncertainty Quantification in Non-Stationary Computer Simulations

Deep Gaussian processes (DGPs) are increasingly popular as predictive models in machine learning for their non-stationary flexibility and ability to cope with abrupt regime changes in training data. The layered structure of the DGP likelihood makes direct inference impossible. Existing variational inference methods offer thrifty predictions but oversimplify uncertainties, particularly in the low data settings common in computer experiments. To achieve full uncertainty quantification, we present a novel elliptical slice sampling Bayesian posterior inferential scheme for DGP surrogates. Elliptical slice sampling is particularly suited for sampling latent Gaussian layers as it is free of tuning parameters and is able to bounce readily between multiple modes. Efficient computation relies on parsimonious layouts of latent layers, effective mixing of MCMC chains, and careful utilization of parallel processing. Our methods are illustrated on simulation data and real computer experiments of varying input dimensionality. We provide an open source implementation in the "deepgp package on CRAN.

Annie Sauer
Virginia Tech
anniees@vt.edu

MS4

A Splitting Algorithm for Dynamical Low-Rank Approximation Motivated by the Fibre Bundle Structure of Matrix Manifolds

We present a new splitting algorithm for dynamical low-rank approximation motivated by the fibre bundle structure of the set of fixed rank matrices. We first introduce

a geometric description of the set of fixed rank matrices which relies on a natural parametrization of matrices. More precisely, it is endowed with the structure of analytic principal bundle, with an explicit description of local charts. For matrix differential equations, we introduce a first order numerical integrator working in local coordinates. The resulting algorithm can be interpreted as a particular splitting of the projection operator onto the tangent space of the low-rank matrix manifold. Numerical experiments for the resolution of dynamical systems with uncertain parameters are presented.

Marie Billaud Friess
Centrale Nantes
Laboratoire de Mathématiques Jean Leray
marie.billaud-friess@ec-nantes.fr

Antonio Falco Montesinos
Universidad CEU Cardenal Herrera
afalco@uchceu.es

Anthony Nouy
Ecole Centrale Nantes
anthony.nouy@ec-nantes.fr

MS4

Dynamical Low-Rank Approximation to Treat Uncertainty in Hyperbolic Problems

Quantifying uncertainties in hyperbolic equations is a source of several challenges. First, the solution forms shocks, which can lead to oscillatory behavior in the numerical approximation. Second, the number of unknowns required for an effective discretization of the solution grows exponentially with the dimension of the uncertainties, yielding high computational costs and large memory requirements. An efficient representation of the solution via adequate basis functions permits tackling these difficulties. The generalized polynomial chaos (gPC) polynomials allow such an efficient representation when the distribution of the uncertainties is known. These distributions are usually only available for input uncertainties such as initial conditions, therefore the efficiency of this ansatz can get lost during runtime. In this talk, we will make use of the dynamical low-rank approximation (DLRA). This guarantees an efficient approximation of the solution even if the underlying probability distributions change over time. Furthermore, filters to mitigate the appearance of spurious oscillations are implemented, and a strategy to enforce boundary conditions is introduced.

Lukas Einkemmer
University of Innsbruck
lukas.einkemmer@gmail.com

Jonas Kusch
Karlsruhe Institute of Technology (KIT)
jonas.kusch@kit.edu

Gianluca Ceruti
Universität Tübingen
ceruti@na.uni-tuebingen.de

Martin Frank
Karlsruhe Institute of Technology
martin.frank@kit.edu

MS4

Rank-Adaptive Tensor Methods for High-Dimensional Nonlinear PDEs

We present a new rank-adaptive algorithm for temporal integration of high-dimensional nonlinear PDEs on tensor manifolds. The new algorithm combines functional tensor train (FTT) series expansions, operator splitting time integration, and a new criterion to add or remove tensor tensor modes from the PDE solution as time integration proceeds. This allows us to overcome well-known computational challenges associated with dynamic tensor approximation, including low-rank modeling errors and the need to invert the covariance matrix of the FTT modes at each time step. Numerical applications are presented and discussed for linear and nonlinear advection problems, and for the Fokker-Planck equation.

Daniele Venturi
Department of Applied Mathematics
University of California Santa Cruz
venturi@ucsc.edu

MS4

Variational Formulation and Stability Properties of a Projector-Splitting Scheme for Dynamical Low Rank Approximation of Random Parabolic Equations

Dynamical Low Rank (DLR) approximation for time-dependent problems with random parameters can be seen as a reduced basis method, in which the solution is expanded as a linear combination of few deterministic functions with random coefficients. The spatial basis is free to evolve in time, thus adjusting at each time to the current structure of the solution. In this talk we consider the DLR approximation of random parabolic equations and propose a class of fully discrete numerical schemes. Similarly to the continuous DLR approximation, our schemes are shown to satisfy a discrete variational formulation. By exploiting this property, we establish stability of our schemes: we show that our explicit and semi-implicit versions are conditionally stable under a parabolic type CFL condition which does not depend on the smallest singular value of the DLR solution; whereas our implicit scheme is unconditionally stable. Moreover, we show that, in certain cases, the semi-implicit scheme can be unconditionally stable if the randomness in the system is sufficiently small. Furthermore, we show that these schemes can be interpreted as projector-splitting integrators and are strongly related to the scheme proposed in [C. Lubich, I. V. Oseledets: A projector-splitting integrator for dynamical low-rank approximation. Bit Numer Math, 2014], to which our stability analysis applies as well. The analysis is supported by numerical results showing the sharpness of the obtained stability conditions.

Eva Vidlickova
EPFL, Lausanne
eva.vidlickova@epfl.ch

Fabio Nobile
EPFL, Switzerland
fabio.nobile@epfl.ch

Yoshihito Kazashi
École polytechnique fédérale de Lausanne

y.kazashi@uni-heidelberg.de

MS5

Hierarchical Methods for Bayesian Experimental Design

This talk presents a multilevel double-loop Monte Carlo (MLDLMC) method for efficiently computing the Expected Information Gain (EIG) criterion used in simulation-based Bayesian optimal experimental design for nonlinear models. The multilevel approach uses the simulation model on a hierarchy of meshes with different resolutions as generalized control variates for the EIG with a model on a fine enough mesh given the desired accuracy. We use Laplace approximations to construct importance-sampling measures to reduce the computational work for the inner-loop sample averaging. We determine the values for the method parameters by minimizing the average computational work, subject to satisfying the desired error tolerance with a specified probability of success. The computational efficiency of MLDLMC for computing EIG is demonstrated for an electrical impedance tomography experiment where the goal is to infer the fiber orientation in composite laminate materials from experimental data.

Joakim Beck
King Abdullah University of Science and Technology (KAUST)
joakim.beck@kaust.edu.sa

Ben Mansour Dia
CPG, King Fahd University of Petroleum and Minerals
diabenmansour@gmail.com

Luis Espath
RWTH Aachen University
espath@uq.rwth-aachen.de

Raul F. Tempone
Mathematics, Computational Sciences & Engineering
King Abdullah University of Science and Technology
raul.tempone@kaust.edu.sa

MS5

Bayesian Learning of Multifidelity Surrogate Networks: Mcmc and Variational Inference Approaches

This talk discusses Bayesian inference methodologies for MFNets – a paradigm for multifidelity information fusion via directed acyclic graphs. MFNets provide a flexible approach to modeling the relationships between unstructured ensembles of models and information sources by linking the outputs of each information source though a network of models. As a result, data on a high-fidelity information source informs a full cascade of it's ancestral information sources. In this talk we discuss Bayesian learning of the parameters of this network. We discuss both sampling-based methodologies targeting the full posterior, as well as variational approaches for its approximation. We demonstrate the complexity of the shape of the posterior and its resulting challenges for sampling methods. We also demonstrate the performance of variational inference methodologies and discuss their challenges for representing highly complex multi-modal behavior. Examples from both synthetic and physical models are provided.

Alex Gorodetsky, Trung Pham
University of Michigan
goroda@umich.edu, trungp@umich.edu

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

MS5

Parameter Estimation for the McKean-Vlasov Stochastic Differential Equation

We consider the problem of parameter estimation for a stochastic McKean-Vlasov equation, and the associated system of weakly interacting particles. We first establish consistency and asymptotic normality of the offline maximum likelihood estimator for the interacting particle system in the limit as the number of particles . We then propose an online estimator for the parameters of the McKean-Vlasov SDE, which evolves according to a continuous-time stochastic gradient descent algorithm on the asymptotic log-likelihood of the interacting particle system. We prove that this estimator converges in to the stationary points of the asymptotic log-likelihood of the McKean-Vlasov SDE in the joint limit as and , under suitable assumptions which guarantee ergodicity and uniform-in-time propagation of chaos. We then demonstrate, under the additional assumption of global strong concavity, that our estimator converges in to the unique maximiser of this asymptotic log-likelihood function, and establish an convergence rate. We also obtain analogous results under the assumption that, rather than observing multiple trajectories of the interacting particle system, we instead observe multiple independent replicates of the McKean-Vlasov SDE itself or, less realistically, a single sample path of the McKean-Vlasov SDE and its law. Our theoretical results are demonstrated via two numerical examples, a linear mean field model and a stochastic opinion dynamics model.

Nikolas Kantas
Imperial College London
n.kantas@imperial.ac.uk

Louis Sharrock
Bristol university
louis.sharrock@bristol.ac.uk

Grigorios Pavliotis
Imperial College London
Department of Mathematics
g.pavliotis@imperial.ac.uk

Panos Parpas
Imperial College London
panos.parpas@imperial.ac.uk

MS5

Multilevel Estimation of Normalization Constants Using the Ensemble Kalman-Bucy Filter

In this article we consider the application of multilevel Monte Carlo for the estimation of normalizing constants. In particular we make use of the filtering algorithm, the ensemble Kalman-Bucy filter (EnKBF), which is an N-particle representation of the Kalman-Bucy filter (KBF). The EnKBF is of interest as it coincides with the optimal filter in the continuous-linear setting, i.e. the KBF. This motivates our particular setup in the linear setting. The resulting methodology we use is the multilevel ensemble Kalman-Bucy filter (MLEnKBF). We provide an analysis

based on deriving L_q-bounds for the normalizing constants using both the single-level, and the multilevel algorithms.

Hamza Ruzayqat, Neil Chada, Ajay Jasra
King Abdullah University of Science and Technology
hamza.ruzayqat@kaust.edu.sa, neilchada123@gmail.com, ajay.jasra@kaust.edu.sa

MS6

Low-Rank Bures-Wasserstein Barycenters

We study the problem of estimation and computation of Bures-Wasserstein barycenters from rank-one measurements. In particular, we prove convexity properties of the energy function, which ensure fast statistical rates as well as efficient computation. We highlight an intriguing connection with low-rank matrix recovery problems that yield efficient and novel algorithms for this task.

Tyler Maunu
Brandeis University
maunu@brandeis.edu

MS6

Low-Rank Regularized Optimal Transport Meets the Generalized Bayesian Inversion

I will present a class of hierarchically low-rank optimal transport (OT) dissimilarity measures that overcome the computational limitations of the celebrated Wasserstein metric. The new class of measures are obtained through a combination of the entropic regularization of OT problem, the hierarchical matrix technique, and data normalization. Motivated by recent improvements put forth by Cuturi and others, I will analyze and implement a new algorithm that computes a hierarchically low-rank approximation of the entropic regularized OT problem. The algorithm exploits the Kronecker structure and the asymptotic regularity of the underlying kernel matrix in Sinkhorns algorithm of regularized OT to achieve quasi-linear time and memory complexity. The application of the new OT measure to problems involving signed data, where positivity and mass balance are not expected, will be addressed by data normalization. The main challenge with data normalization is the possible loss of convexity. Employing a one-to-one and smooth normalization with appropriate pre-selected hyperparameters, I will show that the proposed normalized OT measure preserves strong convexity as well as Lipschitz continuity. Relevant real-world applications include 2D quantum data, 3D seismic data (e.g. the USArray), and 4D functional magnetic resonance imaging (fMRI) data. I will also discuss a general Bayesian framework based on the proposed measure for structured, signed data.

Mohammad Motamed
University of New Mexico
motamed@math.unm.edu

MS6

Linearised Optimal Transport Distances

Optimal transport is a powerful tool for measuring the distances between signals. A common choice is to use the Wasserstein distance where one is required to treat the signal as a probability measure. This places restrictive conditions on the signals and although ad-hoc renormalisation can be applied to sets of unnormalised measures this can often dampen features of the signal. The second disadvantage is that despite recent advances, computing optimal transport distances for large sets is still difficult. In this talk I will focus on the Hellinger–Kantorovich distance, which can be applied between any pair of non-negative measures. I will describe how the distance can be linearised and embedded into a Euclidean space (the analogue of the linear optimal transport framework for Hellinger–Kantorovich). The Euclidean distance in the embedded space is approximately the Hellinger–Kantorovich distance in the original space. This method, in particular, allows for the application of off-the-shelf data analysis tools such as principal component analysis.

Matthew Thorpe
University of Manchester
matthew.thorpe-2@manchester.ac.uk

MS6

Inverse Optimal Transport

Discrete optimal transportation problems arise in various contexts in engineering, the sciences and the social sciences. Often the underlying cost criterion is unknown, or only partly known, and the observed optimal solutions are corrupted by noise. In this talk I will present a systematic approach to infer unknown costs from noisy observations of optimal transportation plans. It is based on the Bayesian formulation of the inverse optimal transportation problem and allows for the quantification of uncertainty. I will then illustrate the developed methodologies using the example of international migration flows.

Marie-Therese Wolfram
Mathematics Department, University of Warwick
m.wolfram@warwick.ac.uk

Andrew Stuart
Computing + Mathematical Sciences
California Institute of Technology
astuart@caltech.edu

MS7

Learning Elliptic Partial Differential Equations with Randomized Linear Algebra

Can one learn a differential operator from pairs of solutions and righthand sides? If so, how many pairs are required? These two questions have received significant research attention in differential equation learning. Given input-output pairs from an unknown elliptic partial differential equation in three dimensions, we will derive a theoretically rigorous scheme for learning the associated Green's function G. By exploiting the hierarchical low-rank structure of Greens functions and extending the randomized SVD algorithm to Hilbert-Schmidt operators, we will identify a learning rate associated with elliptic partial differential operators in three dimensions and bound the number of input-output training pairs required to recover a Greens function approximately.

Nicolas Boulle
Mathematical Institute
University of Oxford
boulle@maths.ox.ac.uk

Alex J. Townsend
Cornell University

townsend@cornell.edu

MS7

Convergence Rates for Learning Linear Operators from Noisy Data

Motivated by the learnability of compact, bounded, and unbounded operators that define forward and inverse problems, in this talk we analyze the Bayesian inverse problem of recovering an unknown linear operator on an infinite-dimensional Hilbert space from noisy input-output data. Under our imposed assumptions, this problem is explicitly solvable and we establish convergence rates of the Bayesian posterior solution in the large data limit. Connections to statistical learning theory will also be discussed. Our numerical results for learning differential (unbounded), identity (bounded), and inverse differential (compact) operators exhibit excellent agreement with our theory, suggesting the sharpness of the rates, and also exhibit consistent behavior even when our idealized theoretical assumptions are violated.

Nicholas H. Nelsen
California Institute of Technology
nnelsen@caltech.edu

MS7

Physics-Constrained Bayesian Inference of an Uncertain Operator in the Sparse-Data Regime

We present a novel Bayesian inverse problem to infer an infinite-dimensional uncertain operator appearing in a differential equation, whose action on an observable state variable affects its dynamics. Inference is made tractable by parametrizing the operator using its eigendecomposition. The plausibility of operator inference in the sparse data regime is explored in terms of an uncertain, generalized diffusion operator appearing in an evolution equation for a contaminant's transport through a heterogeneous porous medium. Sparse data are augmented with prior information through the imposition of deterministic constraints on the eigendecomposition and the use of qualitative information about the system in the definition of the prior distribution. Limited observations of the state variable's evolution are used as data for inference, and the dependence on the solution of the inverse problem is studied as a function of the frequency of observations, as well as on whether the data is collected as a spatial or time series. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energys National Nuclear Security Administration under contract DE-NA0003525.

Teresa Portone
Sandia National Laboratories
tporton@sandia.gov

Robert Moser
UT Austin
Oden Institute
rmoser@oden.utexas.edu

MS7

Learning the Optimal Tikhonov Regularization Operator for Inverse Problems

A common strategy to tackle the instability of inverse problems is the use of regularizers, which are (families of) operators providing a stable approximation of the solution map. The application of machine learning techniques has recently made it possible to design data-driven reconstruction algorithms, including the possibility to learn regularization operators through sample data. In this talk, I will consider a linear inverse problem defined on Hilbert spaces and tackle the problem of learning the optimal operator among the family of generalized Tikhonov regularizers. After setting a statistical framework for the proposed learning problem, allowing to consider a sufficiently large class of noise models, I will characterize the optimal regularizer, showing that it is completely independent of the forward operator. Then, I will propose two strategies to learn the regularizer from a finite training set: a supervised one, based on samples of both inputs and outputs, and an unsupervised one, based only on a sample of outputs. In both cases, I will prove asymptotic bounds on the excess risk, comparing the performances of the sample-based regularizers with the optimal one. I will finally provide a validation of the discussed results by means of a set of numerical examples, both related to a denoising problem and an ill-posed problem. This is based on joint work with G. S. Alberti, E. De Vito, M. Santacesaria (University of Genoa), and M. Lassas (University of Helsinki)

Luca Ratti
University of Genoa
luca.ratti@unige.it

MS8

Layer Adaptive Node Selection in Bayesian Neural Networks: Statistical Guarantees and Implementation Details

Sparse deep neural networks have proven to be efficient for predictive model building in large- scale studies. Although several works have studied sparse neural architectures, they have primarily focused on the edge selection. Sparsity through edge selection might be intuitively appealing; however, it does not necessarily reduce the structural complexity of a network. Instead pruning excessive nodes in each layer leads to a structurally sparse network which has lower computational complexity and memory footprint. We propose a Bayesian sparse solution using spike-and-slab Gaussian priors to allow for node selection during training. The use of spike-and-slab prior alleviates the need of an ad-hoc thresholding rule for pruning redundant nodes. We adopt a variational Bayes approach to circumvent the computational challenges of Markov Chain Monte Carlo implementation. In the context of node selection, we establish the fundamental result of variational posterior consistency together with the characterization of prior parameters. Contrary to the previous works, our theoretical development accommodates sparse networks with layer-dependent node structures. With a layer-wise characterization of prior inclusion probabilities, we also discuss optimal contraction rates of the variational posterior. Finally, we provide empirical evidence to substantiate that our theoretical work facilitates layer-wise optimal node recovery together with competitive predictive performance.

Shrijita Bhattacharya
Michigan State University

bhatta61@msu.edu

MS8

Information Theoretic Objectives, Generalization, and Robustness

I will give an overview of a number of information theoretic objective functions and simple techniques for getting correct bounds on information theoretic quantities. I will show that careful application of these bounds can result in substantial empirical improvements to classical generalization as well as to robustness to a wide variety of distributional shifts, even on larger-scale problems like ImageNet.

Ian Fischer
Google Research
iansf@google.com

MS8

Bayesian Principles for Learning-Machines

Humans and animals have a natural ability to autonomously learn and quickly adapt to their surroundings. How can we design machines that do the same? In this talk, I will present Bayesian principles to bridge such gaps between humans and machines. I will show that a wide-variety of machine-learning algorithms are instances of a single learning-rule derived from Bayesian principles. The rule unravels a dual perspective yielding new mechanisms for knowledge transfer in learning machines. My hope is to convince the audience that Bayesian principles are indispensable for an AI that learns as efficiently as we do.

Emtiyaz Khan
RIKEN center for Advanced Intelligence Project
Tokyo
emtiyaz@gmail.com

MS8

How Do We Build Models That Learn and Generalize?

To answer scientific questions, and reason about data, we must build models and perform inference within those models. But how should we approach model construction and inference to make the most successful predictions? How do we represent uncertainty and prior knowledge? How flexible should our models be? Should we use a single model, or multiple different models? Should we follow a different procedure depending on how much data are available? How do we learn desirable constraints, such as rotation, translation, or reflection symmetries, when they don't improve standard training loss? In this talk I will present a philosophy for model construction, grounded in probability theory. I will exemplify this approach with methods that exploit loss surface geometry for scalable and practical Bayesian deep learning, and resolutions to seemingly mysterious generalization behaviour such as double descent. I will also consider prior specification, generalized Bayesian inference, and automatic symmetry learning. The talk will primarily be based on https://arxiv.org/abs/2002.08791 and https://arxiv.org/abs/2104.14421.

Andrew Wilson
New York University
andrewgw@cims.nyu.edu

MS9

Uncertainty Challenges in Energy Systems

This talk will discuss key challenges in uncertainty quantification associated with energy systems analysis and decision support. Current energy planning and policy questions are among the most important issues for society, and decision support inevitably leads to construction of large scale, computationally intensive computer models. Examples of current decision support applications and progress in uncertainty treatment will be presented based on the author's experience, including from overall energy policy, policy for a specific sector such as heat, network planning, national level generation capacity planning, and local area climate resilience. This will be followed by discussion of future requirements and directions for uncertainty work, including issues specifically associated with energy planning against a background of uncertain climate.

Chris Dent
University of Edinburgh
chris.dent@ed.ac.uk

MS9

Emulation Methodology to Perform History Matching of Energy Models under Uncertainty

Computer simulators are used in the energy sector to address problems at different scales: from modelling buildings' energy consumption, to optimising costs and carbon emissions of nationwide networks. In all cases, an accurate calibration of the model is crucial to ensure that robust decisions are taken on the basis of the simulated results. In this talk we review the main sources of uncertainties in calibrating energy simulators. We discuss and illustrate *History Matching* (HM), a process to carry out calibration that accounts for the uncertainties affecting the model and the data to which model outputs are compared. This feature differentiates HM from the use of measures commonly employed in the energy literature to assess the adequacy of a calibrated model. Moreover, by using Bayes Linear updates, we construct a statistical surrogate of the simulator and are therefore able to perform HM over remarkably large samples of the parameter space - again, as opposed to classical methods in the energy community. If time permits, we will also look into the mathematically interesting problem of history matching parameters, in the case where available data and model outputs take the form of time series.

Dario Domingo
Durham University
dario.domingo@durham.ac.uk

MS9

Bayesian Decision Support Tool for Energy Planning

Gaussian Process (GP) emulators are the principal tool in UQ to approximate complex computer model behaviour across the input space. In recent years, motivated by the problem of coupling computer models, progress has been made in the theory of the analysis of the network of connected GP emulators. In this talk, we combine these recent methodological advances with classical state-space models to construct a Bayesian decision support system. This ap-

proach gives a coherent probability model that produces predictions with the measure of uncertainty in terms of two first moments and enables the propagation of uncertainty from individual decision components. The developed methodology was used to produce a decision support tool for the Northumberland County Council to consider the low carbon technologies to transform its infrastructure to reach a net-zero carbon target in 2050. In particular, we demonstrate how to couple information from energy models and time-series data to quantitatively assess the impact on operational costs and carbon emissions of various policy choices.

Victoria Volodina
The Alan Turing Institute
vicvolodina@gmail.com

MS10

Safe Set Estimation of Expensive to Evaluate Functions with Gaussian Processes

In many reliability engineering problems we are interested in studying functions that can only be evaluated on a safe set. That is, a function returns a real value when evaluated on a set of parameters and a non-valid flag when evaluated outside this safe set. In this work we consider the case where such functions are expensive to evaluate, e.g. they are evaluated by numerically solving complex equations, and the safe set is unknown. We study how to estimate the function and the safe set given a dataset of mixed evaluations. We take a Bayesian approach and put a Gaussian process (GP) prior on the function. The function is evaluated on a set of inputs which provide a mixed dataset of real values and binary values (valid, non-valid). We model the observations with a mixed likelihood. By considering a probit likelihood for the binary observations and an independent Gaussian likelihood for the regression values, we show that the exact posterior is a skewGP. SkewGPs are a generalization of GPs which allows for a fast computation of the mean values and its uncertainty. We then provide a technique to estimate the safe set and to evaluate its uncertainty. In the particular case where the safe set is an excursion set of the function we use both the function and the set uncertainty to define an active learning technique for the set estimation problem built with the objective of obtaining better set estimates with a small number of non-valid evaluations.

Dario Azzimonti
Dalle Molle Institute for Atificial Intelligence USI-SUPSI
CH-6962 Lugano-Viganello
dario.azzimonti@idsia.ch

MS10

Agnostic Multi-Fidelity Regression for Aerospace Design Applications

Multi-fidelity regression models bring substantial advantages to the aircraft preliminary design phase, when data from computer simulations and preliminary experimental tests are exploited to define the best feasible configuration. In the multi-fidelity approach to the modeling issue, pieces of information of diverse fidelity and complexity are leveraged concurrently, to improve estimate accuracy and to reduce the burden associated to parametrization. In such setting, it is fundamental to establish the correct hierarchy in terms of data credibility w.r.t. the targeted application. Unfortunately, the complexity challenging aerospace design problems generally makes the direct estimation of data fidelity difficult, if not intractable. Therefore, the fidelity hierarchy is often established according to a hypotheses-driven process, hence leaving ground to modeling biases. We aim at developing an agnostic multi-fidelity regression framework robust to possible modeling biases due to a corrupted Modelers prior belief concerning the alleged fidelity of the available data sets. In particular, we focus on multi-fidelity co-kriging methods, proposing an extended formulation capable of mitigating the drawbacks of an ill-chosen fidelity hierarchy.

Giulio Gori
Politecnico di Milano
giulio.gori@polimi.it

Olivier Le Matre
CNRS/INRIA/CMAP, École Polytechnique, IPP, Route de Saclay
91128, Palaiseau, France
olivier.le-maitre@polytechnique.edu

Pietro M. Congedo
INRIA Bordeaux Sud-Ouest (FRANCE)
pietro.congedo@inria.fr

MS10

Set Inversion under Functional Uncertainties with Gaussian Process Regression Defined in the Joint Space of Control and Uncertain Variables

During this talk we will present an efficient sampling strategy to solve an inversion problem subjected to functional uncertainties. More precisely, we aim at characterizing a control variable region defined by exceedance above a prescribed threshold of specific Quantities of Interest. Our study is motivated by an automotive industrial application consisting in the identification of the set of values of control variables of a gas after-treatment device, in line with pollutant emission standards of a vehicle under driving profile uncertainties. In that context, driving profile uncertainties are modelled by a functional random variable and the constrained response in the inversion problem is formulated as the expectation over this functional random variable only known through a set of realizations. The unknown set is then associated only with control variables. As often in industrial applications, this problem involves high-fidelity and time-consuming computational models. We thus propose an approach that makes use of Gaussian Process metamodels built on the joint space of control and uncertain input variables. Specifically, we define a learning criterion based on uncertainty in the excursion of the Gaussian Process and derive tractable expressions for variance reduction in such a framework. We will present applications to analytical examples, then to the automotive industrial test case, demonstrating the accuracy and the efficiency brought by the procedure we propose.

Reda El Amri
Formerly IFP Energies Nouvelles
elamrireda.pro@gmail.com

Celine Helbert
Institut Camille Jordan
celine.helbert@ec-lyon.fr

Miguel Munoz Zuniga
IFP Energies Nouvelles
miguel.munoz-zuniga@ifpen.fr

Clémentine Prieur
Grenoble Alpes University
Jean Kunzmann Lab, INRIA project/team AIRSEA
clementine.prieur@univ-grenoble-alpes.fr

Delphine Sinoquet
IFP Energies nouvelles
delphine.sinoquet@ifpen.fr

MS10

Robust Design for Civil Flood Defense: Methodology, Algorithm and Application to the Loire River in France

The design of civil flood defenses aims to define combinations of defenses enough robust to support penalizing conditions of flooding. Nevertheless, the underlying numerical model of the river is often calibrated with a lone value of its environment parameters (say its frictions coefficients), which is sometimes a very strong hypothesis. For instance, the Loire river is known to have a quite changing minor bed due to high sediment transport, so that its friction coefficients is usually considered as a non-negligible uncertainty source. The purpose of this study is to investigate the consequences of such uncertainties taken as epistemic sources on a robust flood defenses design. The mathematical formulation of the problem (and so its solving) is now the merge of three canonical issues: optimization, inversion, and uncertainties propagation. It remains an open problem to support all these questions in one learning process, but a brute force integration of uncertainties is still possible. This example of a perturbed design will give some insights to support deeper uncertainties, possibly induced non stationary environment over time.

Yann Richet
Institut de Radioprotection et de Sreté Nucléaire
yann.richet@irsn.fr

Lucie Pheulpin
IRSN
lucie.pheulpin@irsn.fr

Charlie Sire
Institute for Radiological Protection and Nuclear Safety
charlie.sire@irsn.fr

MS12

Optimal Design of Infinite-Dimensional Bayesian Inverse Problems under Uncertainty

We consider optimal design of inverse problems governed by PDEs under uncertainty. In many application problems, the governing PDEs, in addition to the parameters being estimated, one also has additional model parameters that are not known exactly and are uncertain. Inversion and design of experiments must account for these additional model uncertainties. This talk considers methods for optimal experiment design for such inverse problems that are robust with respect to model uncertainties. We will discuss recent work on OED for linear inverse problems under uncertainty and then discuss extensions to nonlinear inverse problems.

Alen Alexandarian
Department of Mathematics
NC State University
alexanderian@ncsu.edu

Ruanui Nicholson
Department of Engineering Science
University of Auckland
ruanui.nicholson@auckland.ac.nz

Noemi Petra
University of California, Merced
npetra@ucmerced.edu

Georg Stadler
Courant Institute for Mathematical Sciences
New York University
stadler@cims.nyu.edu

MS12

A New Stochastic Learning Approach for Binary Optimization: Application to Bayesian OED

We present a novel stochastic approach to binary optimization ideally suited for optimal experimental design (OED) for Bayesian inverse problems governed by mathematical models such as partial differential equations. The OED utility function, namely, the regularized optimality criterion, is cast into a stochastic objective function in the form of an expectation over a multivariate Bernoulli distribution. The probabilistic objective is then solved by using a stochastic optimization routine to find an optimal observational policy. This formulation: a) is generally applicable to binary optimization problems with soft constraints, and is ideal for OED and sensor placement problems; b) does not require differentiability of the original objective function (e.g., utility function in OED applications) with respect to the design variable thus it enables direct employment of sparsity-enforcing penalty functions such as ℓ_0, without needing to utilize a continuation procedure or apply a rounding technique; c) when applied to OED problems, this framework exhibits much lower computational cost than traditional gradient-based relaxation approaches; and d) it can be applied to both linear and nonlinear OED problems with proper choice of the utility function. In this talk, we discuss the approach, briefly describe convergence guarantees, and demonstrate its power numerically by using a sensor placement for parameter identification.

Ahmed Attia
Argonne National Laboratory
attia@anl.gov

Sven Leyffer
Argonne National Laboratory
leyffer@mcs.anl.gov

Todd Munson
Argonne National Laboratory
Mathematics and Computer Science Division
tmunson@mcs.anl.gov

MS12

Laplace Approximation for Bayesian Optimal Experimental Design with Nuisance Uncertainty

Bayesian optimal experimental design requires the evaluation of the expected information gain functional. This computation usually requires using nested Monte Carlo sampling. The introduction of nuisance parameters to the model would introduce a second inner loop, further increasing the computational cost. We propose to use a small noise approximation for the nuisance uncertainty and a

Laplace approximation for the uncertainty stemming from the parameters of interest. We demonstrate the degradation of this method for large nuisance error in a numerical example showcasing a physical application. Additionally, we show the impact of nuisance error on the optimal design.

Arved Bartuska
RWTH Aachen
Germany
bartuska@uq.rwth-aachen.de

Luis Espath
RWTH Aachen University
espath@uq.rwth-aachen.de

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

MS12

A Bilevel Learning Approach for Optimal Observation Placement in Variational Data Assimilation

We propose a bilevel optimization approach for the placement of space and time observations in variational data assimilation problem 4D-VAR. Within the supervised learning framework, we consider a bilevel problem where the lower-level task is the variational reconstruction of the initial condition of a semilinear system, and the upper-level problem solves the optimal placement with the help of sparsity inducing function. Additionally, some regularity aspects of the variational data assimilation problems in the infinite-dimensional setting are discussed. Due to the pointwise nature of the observations, an optimality system with regular Borel measures is obtained as necessary optimality condition for the lower-level problem. The latter is then considered as constraint for the upper-level instance, yielding in an optimization problem constrained by a multistate system with measures. We demonstrate Lagrange multipliers' existence and derive necessary optimality conditions characterizing the optimal solutions of the bilevel problem. The numerical solution is carried out also on two levels. The lower-level problem is solved using a standard BFGS method, while the upper-level one is solved using a projected BFGS algorithm. Finally, some numerical experiments are presented to illustrate the main features of our approach.

Paula M. Castro C.
Escuela Politécnica Nacional
Ecuador
paula.castro@epn.edu.ec

Juan Carlos De los Reyes
MODEMAT, Escuela Politecnica Nacional
Quito, Ecuador
juan.delosreyes@epn.edu.ec

MS13

Efficient Importance Sampling for Large Sums of Iid Random Variables

We estimate the probability that the sum of nonnegative i.i.d. RVs falls below a given threshold, i.e., $\mathbb{P}(\sum_{i=1}^{N} X_i \leq \gamma)$, via importance sampling (IS). We are interested in the rare event regime when N is large and/or γ is small. The exponential twisting is a popular technique that, in most cases, compares favorably to other estimators. However, it has some limitations: i) it assumes the knowledge of the MGF of X_i and ii) sampling under the new IS PDF might be expensive. The aim is to propose an alternative IS PDF that yields, for certain classes of distributions and in the rare event regime, at least the same performance as the exponential twisting technique and, at the same time, does not introduce serious limitations. The first class includes distributions whose PDFs are asymptotically equivalent, as $x \to 0$, to bx^p, for $p > -1$ and $b > 0$. For this class of distributions, the Gamma IS PDF with appropriately chosen parameters retrieves approximately the same performance of the exponential twisting estimator. In the second class, we consider the Log-normal setting, whose PDF at zero vanishes faster than any polynomial, and we show numerically that a Gamma IS PDF with optimized parameters clearly outperforms the exponential twisting IS PDF. Numerical experiments validate the efficiency of the proposed estimator in delivering a highly accurate estimate in the regime of large N and/or small γ.

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

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

Gerardo Rubino
INRIA Rennes
gerardo.rubino@inria.fr

Raul F. Tempone
Mathematics, Computational Sciences & Engineering
King Abdullah University of Science and Technology
raul.tempone@kaust.edu.sa

MS13

Sequential Active Learning of Low-Dimensional Model Representations for Rare Event Estimation

To date, the reliability analysis of high-dimensional, computationally expensive engineering models remains a difficult challenge in risk and reliability engineering. We use a combination of dimensionality reduction and surrogate modelling termed partial least squares-driven polynomial chaos expansion $PLS{-}PCE$ to render such problems feasible. Standalone surrogate models typically perform poorly for reliability analysis. Therefore, in a previous work, we have used PLS-PCEs to reconstruct the intermediate densities of a sequential importance sampling approach to reliability analysis. Here, we extend this approach with an active learning procedure that allows for improved error control at each importance sampling level. To this end, we formulate an asymptotic estimate of the combined estimation error for both subspace and surrogate model. With this, it is possible to adapt the experimental design so as to optimally learn the subspace representation and the surrogate model constructed therein. The approach is gradient-free and thus works well with black box-type models. We demonstrate the performance of this approach with a series of low- $2D$ to high- $869D$ dimensional example problems featuring a number of well-known caveats for reliability methods besides high dimensions and expensive computational models: strongly nonlinear limit-state functions, multiple relevant failure regions and extremely small prob-

abilities of failure.

Max Ehre
TU München
max.ehre@tum.de

Iason Papaioannou
Engineering Risk Analysis Group
TU M{ü}nchen
iason.papaioannou@tum.de

Bruno Sudret
ETH Zurich
sudret@ibk.baug.ethz.ch

Daniel Straub
Engineering Risk Analysis Group
TU München
straub@tum.de

MS13

Bayesian-Network-Based Modeling and Inference for Failure Events of Building Structures under Sequence of Main and Aftershocks

To ensure the disaster resilience of our society, it is important to evaluate the risk induced by complex natural hazards such as the sequence of main and aftershocks. Mun and Song (2020) recently proposed a probabilistic framework using Bayesian network to describe general structural systems under the sequential earthquakes [Mun, C. and Song, J., Probabilistic Modeling and Inference for Structures under Sequence of Hazardous Events Using Matrix-based Bayesian Network, Proc. of IABSE Conference Seoul 2020, November 9-10, 2020]. In this paper, the BN-based framework is applied to a five-story reinforced concrete (RC) flat slab building constructed in the early 1980s in the Central United States (US). Especially, in modeling the structural responses to ground motions as a part of the BN model, conditional probabilities must be evaluated in the domain of a multi-dimensional feature space, which leads to a cumbersome process and exceedingly large costs of computational simulations. This study utilizes classification tree analysis to discretize the multi-dimensional space and estimates conditional probabilities for each discretized space of the domain. Finally, through the probabilistic inference using the developed BN model, complex relationships between the sequential earthquakes and the structural system are described. In particular, fragilities for main and aftershocks are evaluated in accordance with the realistic conditions encountered under sequential earthquake events.

Changuk Mun
Seoul National University
Department of Civil and Environmental Engineering
changwook80@snu.ac.kr

Jong-Wha Bai
California Baptist University
Department of Civil Engineering and Construction Management
jbai@calbaptist.edu

Junho Song
Seul National University
junhosong@snu.ac.kr

MS13

Computation of Failure Probabilities by Multifidelity Models

Multifidelity models attempt to reduce the computational effort by combining simulation models of different approximation quality and from different sources. They rely on a model hierarchy that is established either by means of an approximation parameter or the Pearson correlation of the model outputs. Information fusion combines outputs from a model hierarchy in order to obtain efficient estimators for a quantity of interest. In the case of additive information fusion, the focus is on differences of the quantity of interest, while multiplicative information fusion takes conditional probabilities into account and is generally based on maximum likelihood or maximum a posteriori estimators. In this presentation, additive and multiplicative information fusion is combined with importance sampling and importance splitting (notably the moving particles method) in order to yield efficient estimators for the probability of failure. Importance sampling and importance splitting based multifidelity reliability estimators are developed and critically assessed by means of numerical examples.

Carsten Proppe
Karlsruhe Institute of Technology
carsten.proppe@kit.edu

Jonas Kaupp
Karlsruhe Technology
jonas.kaupp@kit.edu

MS14

Learning Uncertain Quantities of Interest from Dynamical Systems for Data-Consistent Inversion

A common challenge is to quantify uncertainties on model inputs corresponding to a quantitative characterization of uncertainties on observable Quantities of Interest (QoI). We therefore consider a stochastic inverse problem (SIP) with a solution described by a pullback probability measure. This is referred to as a data-consistent solution since its subsequent push-forward through the QoI map matches the observed probability distribution constructed from data on model outputs. A distinction is made between QoI useful for solving the SIP and arbitrary model output data. Model output data are often given as a (noisy) series of state variables recorded in time resulting in data dimensions that can exceed $\mathcal{O}(1E4)$, and the identification of QoI from this data is not self-evident. We present a new framework, Learning Uncertain Quantities (LUQ), that facilitates the tractable solution of SIPs for dynamical systems. LUQ provides routines for filtering data, learning the underlying dynamics in an unsupervised manner, classifying the observations, and performing feature extraction to learn the QoI map. Subsequently, time series data are transformed into samples coming from the underlying predicted and observed distributions associated with the QoI so that solutions to the SIP are computable. Following an intro & demo of LUQ, numerical results from several SIPs are presented for a variety of dynamical systems arising in the life and physical sciences.

Troy Butler
University of Colorado Denver

troy.butler@ucdenver.edu

MS14

Model-Constrained Approach for Neural Architecture Design with Application to UQ

Machine learning assisted UQ methods for large scale problems and the development of UQ methods for machine learning has received considerable interest in the past few years. However, it is generally unclear on the neural network architecture to be adopted for such applications. In particular, training deep neural networks suffers from common problems such as the need for gigantic training data set to confront over fitting issue. Therefore, there is a need to develop an efficient procedure for training a DNN without making any compromise on the accuracy achieved by traditional back propagation algorithm. This work develops a framework for automatically determining neural architectures to generalize well on the given data set. Numerical investigation on regression problems, classification problems, and UQ problems suggest that our proposed approach outperforms adhoc baseline networks.

Krishnanunni C G, Tan Bui-Thanh
University of Texas at Austin
krishnanunni@utexas.edu, tanbui@oden.utexas.edu

MS14

Multi-Output Surrogate Construction for Fusion Simulations

The ability to perform quality Bayesian and uncertainty analyses is often limited by the computational expense of first-principles physics models. In the absence of a reliable low-fidelity physics model, phenomenological surrogate models can be used to mitigate this expense; however, such models may not adhere to known physics or properties. Furthermore, the interactions of complex physics in high-fidelity codes lead to dependencies between quantities of interest (QoIs) that are difficult to capture when individual surrogates are used for each observable. In applications that consider multiple QoIs simultaneously, separate Gaussian Processes (GPs) are constructed for each QoI. This results in a loss of valuable information regarding the correlated behavior of QoIs. Predicting multiple QoIs with a single GP preserves valuable insights regarding the correlated behavior of the target observables and maximizes the information gained from available data. We present a method of constructing GPs that emulate multiple QoIs simultaneously. As an exemplar, we consider Magnetized Linear Inertial Fusion, a fusion concept that relies on the direct compression of magnetized, laser-heated fuel by a metal liner to achieve thermonuclear ignition. SNL is managed and operated by NTESS under DOE NNSA contract DE-NA0003525.

Kathryn Maupin, Anh Tran
Sandia National Laboratories
kmaupin@sandia.gov, anhtran@sandia.gov

MS14

Stationary Density Estimation of Ito Diffusions Using Deep Learning

We consider the density estimation problem associated with the stationary measure of ergodic Itô diffusions from a discrete-time series that approximate the solutions of SDEs. To take advantage of the characterization of density function through the stationary solution of a parabolic-type Fokker-Planck PDE, we employ deep neural networks to learn the drift and diffusion terms of the SDE and solve a steady-state Fokker-Plank equation associated with the estimated drift and diffusion coefficients with a neural-network-based least-squares method. We establish the convergence of the proposed scheme under appropriate mathematical assumptions, accounting for the generalization errors induced by regressing the drift and diffusion coefficients, and the PDE solvers. This theoretical study relies on a recent perturbation theory of Markov chain result that shows a linear dependence of the density estimation to the error in estimating the drift term, and generalization error results of nonparametric regression and of PDE regression solution obtained with neural-network models.

Yiqi Gu
Purdue University
yiqigu1989@gmail.com

John Harlim
Pennsylvania State University
jharlim@psu.edu

Senwei Liang, Haizhao Yang
Purdue University
liang339@purdue.edu, haizhaoyang@yahoo.com

MS15

Predicting the Effects of Multi-Physics Ensemble of Atmospheric Simulations

Uncertainty in atmospheric and climate models can be estimated using ensembles of simulations that sample uncertainty in physical processes, parameterizations, boundary conditions, initial conditions, and other unconstrained factors. Multi-physics ensembles, which estimate one important component of model uncertainty, are constructed by running a numerical model multiple times over the same simulation period varying categorical physics packages. Large multi-physics ensembles may require thousands of simulations for accurate uncertainty estimates but may not be feasible due to the high computational cost of running a single simulation, even using large supercomputers. Using machine leaning methods, we demonstrate a capability to predict the outcomes of a large multi-physics weather forecast ensemble without having to run the full ensemble. Adaptive sampling methods are used to intelligently select ensemble members, while regression algorithms enable predictions of high-dimensional spatial effects. Our method is validated on a 1,200-member ensemble of the Weather Research and Forecasting model, where we show that the full ensemble variance is well approximated with relatively few training samples and achieve order-of-magnitude speedups needed for real-time forecasting applications.

Don Lucas, Nipun Gunawardena, Giuliana Pallotta
Lawrence Livermore National Laboratory
ddlucas@llnl.gov, gunawardena1@llnl.gov,
pallottagold1@llnl.gov

Matthew Simpson
University of California San Diego
mdsimpson@ucsd.edu

MS15

A Fresh Look at Variography for Global Sensitivity

Analysis of Models with Correlated Inputs: Implications for Climate Models

Variogram Analysis of Response Surfaces (VARS) is a recently developed approach for Sensitivity Analysis (SA). VARS offers three unique features. It (1) provides a more comprehensive characterization of sensitivities including the response surface structure; (2) bridges derivative-based and variance-based approaches to SA, the two most common in the field; and (3) affords significant computational efficiency compared with alternative approaches. This presentation provides an overview of VARS theory and its more recent extension to account for problems with non-uniform and correlated input variables. An example application of VARS is presented in the context of hydroclimatic modelling, where the common, but invalid assumption that the inputs are independent can lead to misleading results.

Saman Razavi
University of Saskatchewan Canada
saman.razavi@usask.ca

MS15

Conditional Polynomial Chaos for Data-Driven Uncertainty Quantification and Reduction

We use conditional Karhunen-Love (KL) expansions to quantify and reduce uncertainty in the coupled linear diffusion equation (a combination of the Darcy law and continuity equations for average pore velocity and hydraulic head) and the advection-dispersion partial differential equation (PDE) with an unknown sparsely sampled space-dependent hydraulic conductivity coefficient $k(x) = \exp(g(x))$ describing fluid flow and transport of a conservative tracer in fully saturated heterogeneous porous media. We achieve reduction in uncertainty by modeling $g(x)$ as a random field with a KL expansion conditioned on g measurements. We employ the conditional KL expansion together with the sparse grid collocation method to compute the moments of the resulting stochastic PDEs. The accuracy of the moment solutions is verified against the Monte Carlo solution of the stochastic PDEs. Uncertainty in the flow problem (the variance of hydraulic head) is further reduced by adaptively selecting additional observation locations of g. We demonstrate that conditioning leads to dimension reduction of the KL representation of $g(x)$ and the uncertainty reduction in the SPDEs states.

Alex Tartakovsky
University of Illinois Urbana-Champaign
amt1998@illinois.edu

MS15

Surrogate-Aided Global Sensitivity Analysis and Uncertainty Characterization in Earth System Model Components

I discuss the use of statistical emulators to perform Sobol variance-based global sensitivity analysis of perturbed-parameter ensembles of Earth system model simulations, with applications to a coupled climate model and its standalone sea ice component, including an analysis of spatially-varying model outputs over the Arctic polar region. As time permits, I will also discuss broader questions of model-form / model structural uncertainties in Earth system modeling.

Nathan Urban
Brookhaven National Laboratory
nurban@bnl.gov

MS16

The Necessity of Uncertainty Quantification in Artificial Intelligence Techniques for Medical Data Analysis

In this talk, we briefly describe the importance of coupling of uncertainty quantification (UQ) with two important artificial intelligence (AI) technologies, i.e., machine learning (ML) and deep learning (DL) methods. A wide variety of ML and DL methods are mainly used in the clinical decision support system (CDSSs). Over the last few years, the importance of quantifying uncertainties of ML and DL methods has been seriously debated. This issue is more prominent in the field of medicine to use these intelligent systems due to the importance of patients' lives. However, there are many practical approaches which can be taken to improve results and measure the uncertainty by adding to the classical ML and DL methods. In other words, using this technique, intelligent CDSSs can be developed in such a way that can say "I DO NOT KNOW" or "I AM NOT SURE". By including uncertainties in designing ML and DL methods as well as their predictions can be fundamentally trusted by the medical community. This allows clinicians to quickly calibrate their trust on CDSS system's outputs, possible flaws in their reasoning, or even uncertainty. Crucially, we end by offering several open research recommendations regarding application of UQ approaches with ML and DL methods which can result in significant improvements to the presentation of trustworthiness results.

Moloud Abdar
Deakin University, Geelong, Australia
mabdar@deakin.edu.au

MS16

Deep Learning, Shearlets, and Uncertainty Quantification: On the Path Towards Interpretable Inverse Problems

In this talk, we will develop a conceptual approach towards inverse problems in imaging sciences by combining the model-based method of sparse regularization by shearlets with the data-driven method of deep learning. Our solvers pay particular attention to the singularity structures of the data. Focussing then on the inverse problem of (limited-angle) computed tomography, we will show that our algorithms significantly outperform previous methodologies, including methods entirely based on deep learning. Finally, we will also touch upon the issue of how to reliable the results of such algorithms, and present a novel, state-of-the-art explainability method based on information theory.

Gitta Kutyniok
Ludwig-Maximilians-Universität München
kutyniok@math.lmu.de

MS16

Sparse X-Ray Tomography with Data-Driven Priors for Sawmill Log Imaging

In high-dimensional inverse problems that are severely ill-posed, the choice of the prior distribution has a significant impact on the result of Bayesian inference. Since in many applications, there exists abundant prior informa-

tion in the form of training examples, the solution accuracy in the Bayesian approach can be improved by harnessing this information to construct a data-driven prior through some generative models such as generative adversarial networks (GANs). We discuss methods for generating data-driven priors in application to sparse X-ray tomography for sawmill log imaging and the stability or well-posedness of Bayesian inverse problems, i.e. the perturbations of GAN-generated prior and posterior distributions.

Angelina Senchukova
Lappeenranta-Lahti University of Technology
LUT, Finland
angelina.senchukova@lut.fi

Tapio Helin
LUT University
Tapio.Helin@lut.fi

Lassi Roininen
Lappeenranta-Lahti University of Technology
lassi.roininen@lut.fi

MS17

A Linearized Learning with Multiscale Deep Neural Network for Stationary Navier-Stokes Equations with Oscillatory Solutions

We present linearized learning schemes to accelerate the convergence of training for stationary nonlinear Navier-Stokes equations. To solve the stationary nonlinear Navier-Stokes (NS) equation, we integrate the procedure of linearization of the nonlinear convection term in the NS equation into the training process of multi-scale deep neural network approximation of the NS solution. Three forms of linearizations are considered and the highly oscillating stationary flows in Complex domains are solved using the proposed linearized learning with multi-scale neural networ. The results show that multiscale deep neural network combining with the linearized schemes can be trained fast and accurately. On the other hand, to learn the operator of dynamic systems with causality, which mathematically can be described by a retarded Greens function, we propose a framework to handle the causality and extend the universal operator approximations theory to operators with causalities. The proposed procedure learns the operator only requires very few training data, but could predict the testing cases accurately and quickly.

Lizuo Liu
Southern Methodist University
lizuol@mail.smu.edu

Bo Wang
School of Mathematics and Statistics
Hunan Normal University
bowang@hunnu.edu.cn

Wei Cai
Department of Mathematics
Southern Methodist University
cai@smu.edu

MS17

Constrained Bayesian Inversion for Transpiration Cooling Problems

Transpiration cooling is a promising cooling technique, aiming to protect materials from extreme temperatures in different applications including rocket combustion chambers, where the material will not withhold the temperatures by itself. Numerous studies investigate possibilities to simulate and evaluate the complex cooling mechanism. In this talk, an approach is introduced that solves an inverse problem while constraining the maximum temperature of the system under parameter uncertainties. Mathematically, this chance inequality constraint is dealt with by a generalized Polynomial Chaos expansion of the system. Utilizing its statistical information, a new approach for quantile estimation is presented. The posterior distribution will be evaluated by different Markov Chain Monte Carlo based methods. A novel method for the constrained case is proposed and tested among others on two-dimensional transpiration cooling models.

Ella Steins, Michael Herty
RWTH Aachen University
steins@aices.rwth-aachen.de, herty@igpm.rwth-aachen.de

MS17

A Probabilistic Characterization of Aleatoric and Epistemic Uncertainty in Solutions to Stochastic Inverse Problems Using Machine Learning Surrogate Models

Stochastic inverse problems have been attracting increasing attention in recent years due to recent advances in data acquisition techniques which enable utilizing tremendous amounts of data to construct data-informed physics-based computational models. At the same time, machine learning methods have also become much more prevalent in computational science, largely due to their ability to learn and exploit low-dimensional structure in high-dimensional data. In this presentation, we will describe some of our recent work on using machine learning surrogate models in the context of solving a particular class of stochastic inverse problems where a data-consistent probability measure on model inputs is sought such that the push-forward of this probability measure matches a given target measure on observations. On the theoretical side, we show that the universal approximation theorem allows for the use of such machine learning surrogate models in this context. However, on the practical side, we demonstrate that the errors and uncertainties in these surrogate models can significantly impact the accuracy of the inferred probability measure, and therefore on any subsequent predictions. This naturally leads to a probabilistic characterization of the solution to the stochastic inverse problem that appropriately characterizes the error/uncertainty in the solution and predictions due to the use of the surrogate model.

Tim Wildey
Optimization and Uncertainty Quantification Dept.
Sandia National Laboratories
tmwilde@sandia.gov

Tian Yu Yen, Troy Butler
University of Colorado Denver
tyen@sandia.gov, butler.troy.d@gmail.com

MS17

Solving Bayesian Inverse Problems via Variational Autoencoders

In recent years, the field of machine learning has made phenomenal progress in the pursuit of simulating real-world

data generation processes. One notable example of such success is the variational autoencoder (VAE). In this work, with a small shift in perspective, we leverage and adapt VAEs for a different purpose: uncertainty quantification in scientific inverse problems. We introduce UQ-VAE: a flexible, adaptive, hybrid data/model-informed framework for training neural networks capable of rapid modelling of the posterior distribution representing the unknown parameter of interest. Specifically, from divergence-based variational inference, our framework is derived such that most of the information usually present in scientific inverse problems is fully utilized in the training procedure. Additionally, this framework includes an adjustable hyperparameter that allows selection of the notion of distance between the posterior model and the target distribution. This introduces more flexibility in controlling how optimization directs the learning of the posterior model. Further, this framework possesses an inherent adaptive optimization property that emerges through the learning of the posterior uncertainty.

Jonathan Wittmer
University of Texas at Austin, U.S.
jonathan.wittmer@utexas.edu

Sheroze Sheriffdeen
University of Texas at Austin
sheroze@oden.utexas.edu

Hwan Goh, Tan Bui-Thanh
Oden Institute of Computational Sciences and Engineering
Hwan.Goh@utexas.edu, tanbui@oden.utexas.edu

MS18

Inference of the Erythrocytes Mechanical Properties Through a Hierarchical Bayesian Framework

In their journey through the human body, the red blood cells (RBCs) experience highly non-linear deformations to flow in arteries and micro-capillaries. These deformations depend on the bending resistance of the lipid-bilayer of the membrane and the shear elasticity of the cytoskeleton. The latter component has zero energy in the so-called stress-free state (SFS) of the cytoskeleton, which has not yet been determined with the current experimental data. Different shapes have been suggested in the literature, such as a biconcave shape, a sphere or an oblate. We infer the SFS (parameterized by its reduced volume) along the other mechanical properties of the RBC by using multiple experimental datasets (equilibrium shape of the RBC, cell stretching and shape relaxation). The inference of the parameters is performed through Bayesian uncertainty quantification, with a hierarchical generative model of the data, capturing the heterogeneity of the RBCs across different experiments

Lucas Amoudruz, Athena Economides
ETH Zuerich
amlucas@ethz.ch, eceva@ethz.ch

Petros Koumoutsakos
Harvard University
petros@seas.harvard.edu

MS18

Uncertainty Quantification in Fluid-Flow Simulations Using Bi-Fidelity Ensembles

Computer simulations are pervasive in science and engineering. Computations, together with theoretical analysis and experiments, constitute the foundation for building knowledge, whether it be to investigate a new physical phenomenon or to assess the performance of an innovative device. However, a single computation, despite its sophistication and complexity, can rarely provide sufficient insights and credible evidence to support critical decisions. To build confidence in computed outcomes, one typically conducts sensitivity analyses, investigates the effect of uncertainties, and explores design variations. All these approaches require an ensemble of computations whose results are rigorously combined using statistical analysis and/or optimization methodologies. In this talk we discuss the use of simulation ensembles for uncertainty quantification. Two different bi-fidelity approaches are compared. The first uses a classical control variate formulation while the second relies on an interpolative decomposition strategy. Applications range from canonical computational fluid dynamics problems to large-scale multi-physics simulations. I will also describe a novel algorithm to handle ensemble with different fidelity implemented in a new programming environment that enables the efficient use of next generation supercomputers.

Gianluca Iaccarino
Stanford University
Mechanical Engineering
jops@stanford.edu

MS18

Calibration and Uncertainty Quantification of Free Parameters in a Model for Turbulent Mixing in the Ocean Surface

The ocean surface boundary layer (OSBL) is a thin, turbulent interface between the atmosphere and ocean interior. Turbulent mixing of heat, momentum, and trace gases in the OSBL is a crucial aspect of ocean circulation and the evolution of Earths climate. Parameterizations of OSBL turbulent fluxes, usually associated with 1-10 m scale turbulent motions, are therefore important components of regional- to planetary-scale ocean models that employ grid resolutions between 100 m and 100 km. In this presentation, we describe the calibration of free parameters in a model for convective and shear-driven OSBL turbulent fluxes based on a prognostic turbulent kinetic energy variable. Our calibration method uses Ensemble Kalman Inversion to minimize the discrepancy between synthetic observations generated by large eddy simulation and multiresolution forward evaluations of the one-dimensional parameterized model. Two key advantages of Ensemble Kalman Inversion are flexibility and a small computational footprint, which proves crucial even for this one-dimensional problem due to the wide range of physical and numerical scenarios we consider. We finish with a discussion of next steps, which include uncertainty quantification via Markov Chain, Monte Carlo sampling (either brute force, or using an emulator) and further calibration or parameters against true observational data using a three-dimensional, realistic ocean regional model with 1 km resolution.

Gregory Wagner
Massachusetts Institute of Technology
glwagner@mit.edu

Adline Hillier
Massachuetts Institute of Technology
ahillier@mit.edu

Navid Constantinuou

Australian National University
navid.constantinou@anu.edu.au

Andre Souza
Massachusetts Institute of Technology
sandre@mit.edu

Raffaele Ferrari
Dept. Earth, Atmospheric, Planetary Sciences
Massachusetts Institute of Technology
rferrari@mit.edu

MS18

Multifidelity Composite Bayesian Optimization for the Inverse Stefan Problem

Process parameters in metal-based additive manufacturing are directly correlated to the microstructure and, therefore, to the properties of the manufactured parts. The evolution of the microstructure resulting from the solidification of the metal can be described by the direct Stefan problem. Conversely, the efficient design of new materials with targeted properties requires solving the inverse problem. The inverse Stefan problem aims to determine process parameters that yield a specific solidification behavior. In this work, we employ a multi-fidelity Bayesian optimization approach to cost-efficiently solve the inverse Stefan problem. We construct a multi-fidelity Gaussian process surrogate model by combining many low-fidelity estimates of a solidification problem with only a few high-fidelity measurements. Both are obtained using the simulation framework ALPACA, applying a conservative level-set model to simulate crystal growth. To solve the inverse problem, we employ the Gaussian process model in a Bayesian optimization approach based on the multi-fidelity knowledge gradient acquisition function. We apply this framework to identify process parameters that allow to observe targeted crystal-growth velocities during dendritic solidification. The target velocity can be easily switched by reusing previously obtained samples and surrogate models. We highlight the cost-efficiency of the multi-fidelity modelling by comparison with single-fidelity Bayesian optimization approaches.

Josef Winter, Rim Abaidi, Jakob Kaiser, Stefan Adami, Nikolaus Adams
Technical University of Munich
josef.winter@tum.de, rim.abaidi@tum.de, jakob.kaiser@tum.de, stefan.adami@tum.de, nikolaus.adams@tum.de

MS19

MCMC Methods on Manifolds Inspired by Geometric Mechanics

In this talk we will discuss the construction of Markov Chain Monte Carlo algorithms from dynamical systems and diffusions via geometric integration and mechanics. In particular, we shall derive the family of Hamiltonian Monte Carlo methods from measure-preserving diffusions, explain their implementation on various classes of manifolds, and discuss recent advances.

Alessandro Barp
University of Cambridge
ab2286@cam.ac.uk

MS19

Non-Reversible Guided Metropolis Kernel

We construct a class of non-reversible Metropolis kernels as a multivariate extension of the guided-walk kernel proposed by Gustafson 1998. The main idea of our method is to introduce a projection that maps a state space to a totally ordered group. By using Haar measure, we construct a novel Markov kernel termed Haar-mixture kernel, which is of interest in its own right. This is achieved by inducing a topological structure to the totally ordered group. Our proposed method, the Δ-guided Metropolis–Haar kernel, is constructed by using the Haar-mixture kernel as a proposal kernel. The proposed non-reversible kernel is at least 10 times better than the random-walk Metropolis kernel and Hamiltonian Monte Carlo kernel for the logistic regression and a discretely observed stochastic process in terms of effective sample size per second.

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

Xiaolin Song
Osaka University
Graduate School of Engineering Science
songxl@sigmath.es.osaka-u.ac.jp

MS19

Geometric Convergence of Polar Slice Sampling

Roberts and Rosenthal introduced and analyzed 1999 the so called polar slice sampler for approximate sampling w.r.t. a posterior target distribution on $\mathbb{R}^d$. They showed that it performs, in contrast to other sampling methods, dimension independent, at least if suitably initialized and if the posterior density satisfies some structural properties. By extending arguments of [Natarovskii, Rudolf, Sprungk, Quantitative spectral gap estimate and Wasserstein contraction of simple slice sampling, 2021] we prove that it has a particularly simple, explicit and dimension-independent spectral gap for strictly increasing, convex and twice differentiable negative log density functions.

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

MS19

Creating Manifold Structures to Accelerate MCMC Sampling

Consider the noisy and incomplete observation y of a quantity of interest x. In Bayesian inverse problems, the vector x typically represents the high-dimensional discretization of a continuous and unobserved field. When the observation process is very "informative", the posterior distribution concentrates in the neighborhood of a nonlinear manifold. As a result, the efficiency of standard MCMC algorithms deteriorates due to the need to take increasingly smaller steps. In this work, we describe how to express this posterior distribution as the marginal of an auxiliary distribution defined on an extended space. This generalizes previous work that only considered the case of an additive Gaussian noise model of the type $y = F(x) + (\text{noise})$. Our

proposed approach allows us to leverage constrained HMC methods that are robust to the low-noise regime.

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

MS20

Integration in Reproducing Kernel Hilbert Spaces of Gaussian Kernels

The Gaussian kernel plays a central role in machine learning, uncertainty quantification and scattered data approximation. However, the basic problem of finding an algorithm for efficient numerical integration of functions reproduced by Gaussian kernels has not been fully solved. In this talk we construct two classes of algorithms that use N evaluations to integrate d-variate functions reproduced by Gaussian kernels and prove the exponential or super-algebraic decay of their worst-case errors. In contrast to earlier work, no constraints are placed on the length-scale parameter of the Gaussian kernel. The first class of algorithms is obtained via an appropriate scaling of classical Gauss-Hermite rules. For these algorithms we derive lower and upper bounds on the worst-case error of the forms $\exp(-c_1 N^{1/d})N^{1/(4d)}$ and $\exp(-c_2 N^{1/d})N^{-1/(4d)}$, respectively, for positive constants $c_1 > c_2$. The second class of algorithms we construct is more flexible and uses worst-case optimal weights for points that may be taken as a nested sequence. For these algorithms we only derive upper bounds, which are of the form $\exp(-c_3 N^{1/(2d)})$ for a positive constant c_3. The talk is based on "Karvonen, Oates & Girolami (2021). Integration in reproducing kernel Hilbert spaces of Gaussian kernels. *Mathematics of Computation*, 90(331):22092233."

Toni Karvonen
Aalto University
Department of Electrical Engineering and Automation
toni.karvonen@helsinki.fi

MS20

Nonlinear Reduced Modelling and State Estimation of Parametric PDEs

We examine the problem of state estimation, that is, reconstructing the solution of a known parametric PDE from m linear measurements. When linear reduced models are used, well known results in reconstruction stability and approximation errors can be used to give bounds of overall error of state estimation. We present some new results and schemes for the deployment of nonlinear reduced models for this task, specifically models that are locally linear for disjoint partitions of the parameter domain. One challenge in this task is sensing which locally linear model to apply, given some specific measurements. Our strategy for this is to consider the residuals, and chose local linear models according to which minimizes the residual associated with the PDE. We discuss results and some interesting dual-minimization strategies for parameter estimation that arise, and present a numerical study of this strategy

James Nichols
Australian National University
BDSI
james.nichols@anu.edu.au

Albert Cohen
Université Pierre et Marie Curie, Paris, France
cohen@ann.jussieu.fr

Wolfgang Dahmen
RWTH Aachen
IGPM
dahmen@igpm.rwth-aachen.de

Olga Mula
CEREMADE (Paris Dauphine University)
mula@ceremade.dauphine.fr

MS20

Enhancing Accuracy of Deep Learning Algorithms by Training on Low-Discrepancy Sequences and Its Higher-Order Extension

We propose a deep supervised learning algorithm based on low-discrepancy sequences as the training set. By a combination of theoretical arguments and extensive numerical experiments we demonstrate that the proposed algorithm significantly outperforms standard deep learning algorithms that are based on randomly chosen training data, for problems in moderately high dimensions. We further extend the theory to higher-order Quasi-Monte Carlo points, which are proved to facilitate higher-order decay (in terms of the number of training samples) of the underlying generalization error, with consistency error bounds that are free from the curse of dimensionality in the input data space, provided that deep neural network weights in hidden layers satisfy certain summability conditions. We provide numerical experiments on elliptic and parabolic PDEs with uncertain inputs that confirm the theoretical analysis.

T. Konstantin Rusch
Seminar for Applied Mathematics
ETH Zurich
konstantin.rusch@sam.math.ethz.ch

MS20

Fast Approximation by Periodic Kernel-Based Lattice-Point Interpolation

In this talk I will discuss the kernel-based approximation of a multivariate periodic function by interpolation at the points of an integration lattice. This combination allows fast evaluation by fast Fourier transform, so avoiding the need for a linear solver. Our main contribution is the application to the approximation problem for uncertainty quantification of elliptic partial differential equations, with the diffusion coefficient given by a random field that is periodic in the stochastic variables. We have a full error analysis, and full details of the construction of lattices needed to ensure a good rate of convergence and an error bound independent of dimension, with numerical experiments supporting the theory.

Vesa Kaarnioja
Aalto University
Department of Mathematics and Systems Analysis
vesa.kaarnioja@iki.fi

Yoshihito Kazashi
École polytechnique fédérale de Lausanne
y.kazashi@uni-heidelberg.de

Frances Y. Kuo
School of Mathematics and Statistics
University of New South Wales

f.kuo@unsw.edu.au

Fabio Nobile
EPFL, Switzerland
fabio.nobile@epfl.ch

Ian H. Sloan
University of New South Wales
School of Mathematics and Stat
i.sloan@unsw.edu.au

MS21

A Bayesian Level Set Method for An Inverse Medium Scattering Problem in Acoustics

In this talk, we are interested in the determination of the shape of the scatterer for the two-dimensional time harmonic inverse medium scattering problems in acoustics. The scatterer is assumed to be a piecewise constant function with a known value inside inhomogeneities, and its shape is represented by the level set functions for which we investigate the information using the Bayesian method. In the Bayesian framework, the solution of the geometric inverse problem is defined as a posterior probability distribution. The well-posedness of the posterior distribution is discussed, and the Markov chain Monte Carlo (MCMC) method is applied to generate samples from the posterior distribution. Numerical experiments are presented to demonstrate the effectiveness of the proposed method.

Zhiliang Deng
University of Electronic Science and Technology of China
dengzhl@uestc.edu.cn

MS21

Level-Set Parameterisations for Ensemble Kalman Inversion

In this talk I discuss recent advances in the implementation of the level-set approach to parameterise unknown interfaces and discontinuous properties with the Ensemble Kalman Inversion (EKI) framework for inverse problems. The proposed approach uses an underlying level-set function parameterised via the SPDE formulation for Whittle-Matern (WM) random fields. We compose the level-set-based WM parameterisation with the forward map and, using the capability of EKI to handle parameter-to-output maps in a blackbox fashion, we infer all relevant (hyper-)parameters of the level-set-based WM parameterisation within the EKI algorithm. We demonstrate the applicability of this approach to solve various inverse problems where the unknown is a discontinuous property arising from the presence of an anomalous material/tissue. We will present numerical examples with applications to (i) non-destructive testing of composite materials, (ii) ground penetrating radar and (iii) magnetic resonance elastography.

Marco Iglesias
University of Nottingham
marco.iglesias@nottingham.ac.uk

MS21

Segmentation, Level Sets, and Uncertainty

The shapes of objects in an image can be recovered by processes of segmentation, however, uncertainties remain. These uncertainties are often expressed as pixel- or voxel-wise probabilities on class labels. I will first discuss Active Mean Fields, a method that imparts a specific probabilistic meaning to level sets used in Chan-Vese style image segmentation. I will describe a second method that essentially uses stochastic level sets to characterize uncertainties in shapes – shapes may be sampled by thresholding random draws from a specific Gaussian process. This sampling approach can be used to estimate arbitrary statistics on the modeled shapes.

William M. Wells
Harvard Medical School, M.I.T.
sw@bwh.harvard.edu

MS21

Bifidelity Data-Assisted Neural Networks in Nonintrusive Reduced-Order Modeling and Its Applications in Uncertainty Quantification

In this talk, I will introduce a new nonintrusive reduced basis method when a cheap low-fidelity model and an expensive high-fidelity model are available. The method employs proper orthogonal decomposition method (POD) to generate the high-fidelity reduced basis and a shallow multilayer perceptron to learn the high-fidelity reduced coefficients. In contrast to previously proposed methods, besides the model parameters, we augmented the features extracted from the data generated by an efficient bi-fidelity surrogate as the input feature of the proposed neural network. By incorporating relevant bi-fidelity features, we demonstrate that such an approach can improve the predictive capability and robustness of the neural network via several benchmark examples. Due to its nonintrusive nature, it is also applicable to general parameterized problems, such as uncertainty quantification.

Xueyu Zhu
University of Iowa
Department of Mathematics
xueyu-zhu@uiowa.edu

MS22

Functional-Input Gaussian Processes with Applications to Inverse Scattering Problems

Surrogate modeling by Gaussian process models has received increasing attention in the analysis of complex problems in science and engineering. Despite extensive studies on Gaussian process modeling, the developments for functional inputs are scarce. Motivated by an inverse scattering problem, a new class of kernel functions is introduced for Gaussian process models with functional inputs. The asymptotic convergence properties of the proposed Gaussian process models are derived and the finite sample performance is demonstrated by numerical examples. In the application to inverse scattering problem, the functional input which is associated with the support of the scattering region of interest is identified, given a measured far-field pattern.

Ying Hung
Department of Statistics and Biostatistics
Rutgers University
yhung@stat.rutgers.edu

Chih-Li Sung
Michigan State University
Department of Statistics and Probability

sungchih@msu.edu

MS22

Gaussian Process Assisted Active Learning of Physical Laws

In many areas of science and engineering, discovering the governing differential equations from the noisy experimental data is an essential challenge. It is also a critical step in understanding the physical phenomena and prediction of the future behaviors of the systems. However, in many cases, it is expensive or time-consuming to collect experimental data. This article provides an active learning approach to estimate the unknown differential equations accurately with reduced experimental data size. We propose an adaptive design criterion combining the D-optimality and the maximin space-filling criterion. In contrast to active learning for other regression models, the D-optimality here requires the unknown solution of the differential equations and derivatives of the solution. We estimate the Gaussian process (GP) regression models from the available experimental data and use them as the surrogates of these unknown solution functions. The derivatives of the estimated GP models are derived and used to substitute the derivatives of the solution. Variable selection-based regression methods are used to learn the differential equations from the experimental data. Through multiple case studies, we demonstrate the proposed approach outperforms the D-optimality and the maximin space-filling design alone in terms of model accuracy and data economy.

Lulu Kang
Department of Applied Mathematics
Illinois Institute of Technology
lkang2@iit.edu

Guang Lin
Purdue University
guanglin@purdue.edu

MS22

A New Uncertainty Quantification Method based on Koopman Operator

We propose a new uncertainty quantification (UQ) method for dynamical systems with random parameters. Unlike conventional UQ approaches in which a set of basis functions associated with random variables are used to construct a surrogate model of the quantity of interest (QoI), our new method uses the eigenfunction and eigenvalues of the Koopman operator of the system to construct the surrogate. The advantage of this new approach is that the eigenpairs of the Koopman operator incorporate key features of the dynamics, hence it can be more efficient in capturing behavior of the system. Empirically, our approach exhibits exponential convergence if the solution is smooth. Moreover, the computation of this new approach is very efficient as it doesn't need time integration like Runge-Kutta scheme, and the statistics of the QoI relies on the moment-generating function of the random variable in the system.

Bian Li, Xiu Yang
Lehigh University
bil215@lehigh.edu, xiy518@lehigh.edu

MS22

Deep Gaussian Process Emulation Using Stochastic Imputation

We introduce a novel inference method for deep Gaussian process (DGP) emulation using stochastic imputation. By stochastically imputing the latent layers, the approach converts the DGP to the linked GP, a state-of-the-art surrogate model formed by a feed-forward network of GPs. This transformation renders a simple while efficient DGP training procedure that only involves optimizations of conventional stationary GPs. In addition, the analytically tractable mean and variance of the linked GP allow one to make predictions from DGP emulators in a fast and accurate manner. We demonstrate the method in a synthetic example and a real-world application, and show that it is a competitive candidate for efficient DGP surrogate modeling in comparison to the Doubly Stochastic Variational Inference (DSVI) and the fully-Bayesian approach. A Python package, called `dgpsi`, implementing the method is produced and publicly available on GitHub.

Deyu Ming
University College London, UK
deyu.ming.16@ucl.ac.uk

Daniel Williamson
University of Exeter
d.williamson@exeter.ac.uk

Serge Guillas
University College London
s.guillas@ucl.ac.uk

MS23

Fast Approximation of High-Rank Hessians to Accelerate MCMC for Bayesian Seismic Inversion

Markov chain Monte Carlo (MCMC) has been used to quantify uncertainties in seismic full-waveform inverse problems. However, MCMC is usually prohibitive for such problems, due to the large number of samples required. The generalized preconditioned CrankNicolson method uses proposals with covariance operators based on the (Gauss-Newton) Hessian of the negative log likelihood at the MAP point to accelerate the sampling process. Computing the exact Hessian is intractable for large scale problems. Moreover, slow decay of its eigenvalues means that low-rank approximation is not effective. We propose a Hessian approximation method based on a local translation invariance approximation. This is achieved through a product-convolution scheme followed by a low-rank correction. The approximation can be expressed in the form of an H-matrix, which facilitates fast matrix computations needed for drawing samples. The results show that the approximation speeds up MCMC significantly compared to low-rank approximation.

Mathew Hu
Oden Institute for Computational Engineering & Sciences
University of Texas at Austin
mathewhu@utexas.edu

Nick Alger
The University of Texas at Austin
Center for Computational Geosciences and Optimization
nalger225@gmail.com

Longfei Gao
The Oden Institute for Computational Engineering and Science
The University of Texas at Austin
longfei.gao@austin.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

Rami Nammour
Total, U.S.
rami.nammour@total.com

MS23

Low-Dimensional Structure in Bayesian Inference Problems with Mixture Models

Efficient solutions to many Bayesian inference problems exploit the low-dimensional structure in the prior-to-posterior update. Such structure typically arises when the data are informative only on a subspace of the parameters. Identifying the subspace, often labeled as the likelihood-informed subspace (LIS) is particularly challenging when using non-gaussian priors, and/or non-linear maps between parameters and data. When the likelihood and/or the prior are prescribed using Gaussian mixture models, the LIS for each component of the posterior mixture is easily identified, and cumulatively they describe the data informed directions for the full problem. We rigorously analyze the ability of these component LIS in approximating the posterior distribution, and demonstrate the utility of the underlying ideas in a high-dimensional atmospheric retrieval problem.

Jayanth Jagalur Mohan, Ricardo Baptista
MIT
jagalur@mit.edu, rsb@mit.edu

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

MS23

Balanced Truncation for Bayesian Inference

We consider the Bayesian approach to the linear Gaussian inference problem of inferring the initial condition of a linear dynamical system from noisy output measurements taken after the initial time. In practical applications, the large dimension of the dynamical system state poses a computational obstacle to computing the exact posterior distribution. Balanced truncation is a system-theoretic method for model reduction which obtains an efficient reduced-dimension dynamical system by projecting the system operators onto state directions which simultaneously maximize energies defined by reachability and observability Gramians. We introduce Gramian definitions relevant to the inference setting and propose a balanced truncation approach based on these inference Gramians that yield a reduced dynamical system that can be used to cheaply approximate the posterior mean and covariance. Our Gramian definitions exploit natural connections between (i) the reachability and the prior covariance and (ii) the observability and the Fisher information. The resulting reduced model then inherits stability properties and error bounds from system theory, and in some settings yields an optimal posterior covariance approximation. Numerical demonstrations on two benchmark problems in model reduction show that our method can yield near-optimal posterior covariance approximations with order-of-magnitude state dimension reduction.

Elizabeth Qian
California Institute of Technology
eqian@caltech.edu

Jemima M. Tabeart
University of Edinburgh
jemima.tabeart@ed.ac.uk

Christopher A. Beattie
Virginia Polytechnic Institute and State University
beattie@vt.edu

Serkan Gugercin
Virginia Tech
Department of Mathematics
gugercin@vt.edu

Jiahua Jiang
Shanghai Tech University
j.jiang.3@bham.ac.uk

Peter R. Kramer
Rensselaer Polytechnic Institute
Department of Mathematical Sciences
kramep@rpi.edu

Akil Narayan
University of Utah
akil@sci.utah.edu

MS23

Geometric Ergodicity of Slice Sampling

For approximate sampling of a partially known distribution, e.g., posterior distributions in Bayesian inverse problems, the slice sampling methodology provides a machinery for the design and simulation of a suitable Markov chain without the necessity to tune any parameters as in many Metropolis-Hastings algorithms. In the machine learning community slice sampling is a frequently used approach, which appears not only there as standard sampling tool. In particular, the elliptical slice sampler attracted in the last decade considerable attention as a dimension-robust algorithm. However, from a theoretical point of view it is not well understood. In this talk, we show the geometric ergodicity of Markov chains generated by elliptical slice sampling as well as simple slice sampling with particular emphasis on their (in)dependence on the state space dimension.

Bjoern Sprungk
TU Freiberg
bjoern.sprungk@math.tu-freiberg.de

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

Viacheslav Natarovskii
University of Goettingen
Institute for Mathematical Stochastics

vnataro@uni-goettingen.de

MS24

Performance Bounds for PDE-Constrained Optimization under Uncertainty

Computational solutions of PDE-constrained optimization under uncertainty involve various approximations such as finite-dimensional approximation of control variables, numerical approximation of PDE solutions, sample average approximation of risk measures or reliability constraints, approximation of nonsmooth expressions and penalization of constraints. In this presentation, we develop a general framework for analyzing such approximations and establish performance bounds for the obtained approximating solutions. Specifically, under the convergence of each approximation, the performance of any cluster point of the computed controls are guaranteed to be sub-optimal with a certain tolerance. We demonstrate the framework by applying it to a concrete example of buffered probability constrained optimal control problem governed by an elliptic PDE with a random coefficient field and a distributed control function.

Johannes O. Royset
Operations Research Department
Naval Postgraduate School
joroyset@nps.edu

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

MS24

Asymptotic Properties of Monte Carlo Methods for PDE-Constrained Optimization under Uncertainty

The numerical solution of every PDE-constrained optimization problem subject to uncertainty requires at some point an approximation of the random parameters. This typically manifests itself in the use of Monte Carlo sampling approaches to replace the underlying probability measure $\mathbb{P}$ by an associated empirical probability measure $\mathbb{P}_N$. Similarly, if only data is available, then we may also take the perspective that the *true* measure $\mathbb{P}$ has been replaced by a discrete approximation. Either way, the solutions we compute, which may require an enormous amount of computing power, may only be understood as a single realization of a complex random process. If this is the case, then it is crucial to understand how the optimal values and optimal solutions behave as the sample size $N \to \infty$. One possibility to derive such results makes use of probability metrics. By taking this perspective, we first derive a general error bound for a class of risk-neutral PDE-constrained optimization problems. Afterwards, we demonstrate how additional assumptions on the structure of uncertainty in the differential operators can be exploited to obtain a rate of convergence using empirical process theory and a central limit theorem for the optimal value.

Thomas Surowiec
Fachbereich Mathematik und Informatik
Philipps-Universitat Marburg
surowiec@mathematik.uni-marburg.de

Werner Römisch
Humboldt-Universität Berlin
romisch@mathematik.hu-berlin.de

MS24

An Interior-Point Approach for Risk-Averse PDE-Constrained Optimization with Risk Measures

We propose a method for solving risk-averse PDE-constrained optimization problems where the used risk measure is a convex combination of the mean and the conditional value-at-risk (CVaR). Since these risk measures can be evaluated by solving a related inequality-constrained optimization problem, we propose a log-barrier technique to approximate the risk measure, which leads to a new continuously differentiable convex risk measure: the log-barrier risk measure. We prove consistency of the approximation via a variational convergence technique. Using the differentiability of the log-barrier risk measure, we derive first-order optimality conditions reminiscent of interior point approaches in nonlinear programming. We study the associated Newton systems in full and reduced form and provide a sufficient condition for local superlinear convergence in the continuous setting. For the discretization of the problem, we employ novel low-rank tensor methods. The presentation concludes with numerical examples.

Michael Ulbrich
Technical University of Munich
Chair of Mathematical Optimization
mulbrich@ma.tum.de

Sebastian Garreis
TU Muenchen, Germany
garreis@ma.tum.de

Thomas Surowiec
Fachbereich Mathematik und Informatik
Philipps-Universitat Marburg
surowiec@mathematik.uni-marburg.de

MS24

Optimal Control of PDEs under Uncertainty with Joint Chance State Constraints

We study optimal control of PDEs under uncertainty, where the state variable is subject to joint chance constraints. While we seek deterministic controls, the corresponding states are probabilistic due to uncertainty in the governing equation. Joint chance constraints require that realizations of this probabilistic state variable satisfy pointwise bounds with a given, typically high, probability. We consider linear and bilinear elliptic PDEs with infinite-dimensional uncertain parameters as governing equations. We argue that properties of the governing equations reduce the effective random space dimension and show how this can be used to approximate the chance probabilities. In particular, we use a spherical-radial decomposition of Gaussian random variables, which allows not only computation of the joint chance probabilities, but also their derivatives, enabling the use of efficient gradient based optimization algorithms.

Florian Wechsung
New York University
wechsung@cims.nyu.edu

Rene Henrion
Weierstrass Institute
Berlin, Germany
rene.henrion@wias-berlin.de

Georg Stadler
Courant Institute for Mathematical Sciences
New York University
stadler@cims.nyu.edu

MS25

Emulation and Uncertainty Quantification for a Coupled Model of Fluid Flow and Mechanical Deformation

Complex physical processes are often described by multiple types of physics coupled together with input and output parameters that are vector valued. Unfortunately, to generate statistics on these models we need to run numerous expensive simulations. Instead of running the full physical model for each design point, we use an emulator which acts as an interpolator to allow us to quickly evaluate the simulation at non-design points. We propose using a Gaussian Stochastic Process (or GaSP) emulator with the coupling accomplished by using a version of a linked emulator for the composite simulator. The vector output is handled by extending parallel partial emulators. In this work we describe the extension of GaSP emulators to coupled systems with vector output and apply the emulator to the Terzaghi consolidation problem. Terzaghi models the consolidation of a 100-inch column of mud which compacts after a load is dropped on the top of the column. The column is impermeable on the sides and bottom, but fluid is allowed to exit the top which results in changing porosity and permeabilty in the column. Mathematically we loosely couple fluid flow and mechanical deformation. In this loose coupling, the change in fluid pressure is passed from the flow equations to act as a load on the mechanical deformation. Solving then for displacement leads to a strain change and ultimately an update in porosity which is then fed back into the fluid equations for the subsequent set of time steps.

Tamara Dolski
University of Texas at Dallas
txk171630@utdallas.edu

Sue Minkoff
University of Texas - Dallas
sminkoff@utdallas.edu

Elaine Spiller
Marquette University
elaine.spiller@marquette.edu

MS25

Coupled Emulation of Poroelastic Deformation and Fluid Pressurization in Biomechanical Models of Articular Cartilage

Biphasic poroelastic models have been widely used to study coupling between deformation and flow in articular cartilage. Mathematical and computational solutions of boundary value problems that model in vitro or in vivo responses to time-varying loading can enhance our understanding of mechanotransduction from the tissue to the cellular scale. They also contribute to the development of optimal strategies for tissue engineering. Coupled emulators have a potential role in accelerating 3D computational biomechanical models (e.g. contact problems in diarthrodial joints). Such emulators can also contribute to coupling models of biomechanical deformation with models for biotransport and cell physiology in states of normal tissue homeostasis, or when alterations occur due to diseases such as osteoarthritis. This talk will present strategies for developing coupled Gaussian stochastic process (GaSP) emulators of biphasic poroelastic deformation, flow and pressurization in cylindrical cartilage explants. The associated boundary value problems can be viewed as a system that couples a forced elastostatics equation to a porous media flow equation. Exact analytical representations, based on orthogonal series solutions, are used to systematically evaluate the accuracy of the emulators. Several time-varying loading protocols, such as creep, stress relaxation and dynamic loading are considered and the effects of loading rates and frequencies on emulation accuracy will be discussed.

Mansoor A. Haider
North Carolina State University
Department of Mathematics
mahaider@ncsu.edu

MS25

Uncertainty Quantification for Trace Gas Sensor Modeling

Trace gas sensors which are compact and portable can be used to detect minute quantities of gases for applications ranging from disease diagnosis to industrial pollution control and homeland security. The sensor contains a laser heat source and a tiny quartz tuning fork that acts as a resonator. The laser source is tuned to the correct frequency range so that if the gas to be detected is present, it will absorb this energy. When this energy is absorbed by the gas molecules, both an acoustic and a thermal wave are generated and propagate out from this laser source. If the laser is positioned between the tines of the tuning fork, then these waves cause the tines to vibrate, and the velocity of the tines of the fork is proportional to the signal strength and hence to the amount of gas present. Mathematically we model this sensor system using a version of the Navier-Stokes equations for pressure and temperature in the fluid, namely, the Morse-Ingard Equations. We couple these equations to the displacement equation in the tuning fork. This two-way coupled system allows us to determine the damping in the system without a priori assumptions or experimental measurements. However, the system involves many input parameters which make it costly to determine statistics about the system. In this talk I will discuss the sensitivity analysis of the coupled system and reduced order modeling to maximize the strength of the signal generated by the gas.

Susan Minkoff
Department of Mathematical Sciences
University of Texas at Dallas
sminkoff@utdallas.edu

Ali Mozumder
Dept of Mathematical Sciences
University of Texas at Dallas
aliahammed.mozumder@utdallas.edu

John Zweck
University of Texas at Dallas
Department of Mathematical Sciences
zweck@utdallas.edu

MS25

Invariant Representations in Heterogeneous Catalysis

The main challenge in using machine learning models to accelerate computational catalysis is the development

of molecular representations and surface descriptors that can be used to predict adsorption energies. While several datasets are publicly available, their applicability to new systems is rather limited due to the intrinsic idiosyncrasies in their data. Namely, the heterogeneity in the density functional theory (DFT) makes the transfer of knowledge challenging between two different systems that make use of two different functionals. We propose to leverage the variability in these open-source datasets and learn invariant material and molecular representations by ignoring the spurious correlations specific to each dataset. To develop these generalizable molecular descriptors, we assume that different approximations in the exchange-correlation functional exhibit invariant effects on surface chemistry. Namely, significant differences in adsorption energies should not change with the DFT functional. This allows us to develop novel deep learning models that are attuned to heterogeneous catalysis by learning from pairwise energies within a functional while enforcing invariant correlation structures across different functionals. In our initial studies, we found statistically significant improvements from learning using data from a variety of DFT functionals as compared with only data calculated using the functional of interest.

Jawad Chowdhury
nu University of North Carolina at Charlotte
mchowdh5@uncc.edu

Charles Fricke, Andreas Heyden
University of South Carolina
cfricke@email.sc.edu, heyden@cec.sc.edu

Gabriel Terejanu
University of North Carolina at Charlotte
gabriel.terejanu@uncc.edu

MS26

Gradient-Based Bayesian Experimental Design for Implicit Models Using Mutual Information Lower Bounds

We introduce a framework for Bayesian experimental design (BED) with implicit models, where the data-generating distribution is intractable but sampling from it is still possible. In order to find optimal experimental designs for such models, our approach maximizes mutual information lower bounds that are parametrized by neural networks. By training a neural network on sampled data, we simultaneously update network parameters and designs using stochastic gradient-ascent. The framework enables experimental design with a variety of prominent lower bounds and can be applied to a wide range of scientific tasks, such as parameter estimation, model discrimination and improving future predictions. Using a set of intractable toy models, we provide a comprehensive empirical comparison of prominent lower bounds applied to the aforementioned tasks. We further validate our framework on a challenging system of stochastic differential equations from epidemiology.

Steven Kleinegesse, Michael Gutmann
University of Edinburgh
steven.kleinegesse@ed.ac.uk, michael.gutmann@ed.ac.uk

MS26

Physics-Informed Model-Based Reinforcement Learning with Quantified Uncertainties

Model-based reinforcement learning (MBRL) algorithms are believed to have higher sample efficiency compared to their model-free counterparts. However, the challenge of learning an accurate model largely limits the performance of MBRL used in a high-dimensional complex physics environment. To deal with the aleatoric uncertainty caused by the randomness of the system or/and the epistemic uncertainty due to limited training data, people proposed probabilistic MBRL algorithms based on model ensemble or Bayesian networks. Nonetheless, these models are usually built in a black-box manner, which poses great challenges to identifying and quantifying uncertainties from various sources. To tackle this issue, we introduce a physics-informed probabilistic MBRL method that leverages physics knowledge (e.g., conservation laws and governing equations) in model development and training processes. By incorporating the prior information of the environment, we carefully estimate the various uncertainties due to data noises, model learning, and model rollout length, and quantify their impact on the policy optimization. We will test our proposed method on dynamic control problems for several physical systems.

Xin-Yang Liu, Jian-Xun Wang
University of Notre Dame
xliu28@nd.edu, jwang33@nd.edu

MS26

Design of Experiments with Functional Variables

The aim of this work is to extend the usual optimal experimental design paradigm to experiments where the settings of one or more variables are functions. For these new experiments, a design consists of combinations of functions for each run of the experiment along with settings for non-functional variables. After briefly introducing the class of functional variables, basis function systems are described. Basis function expansion is applied to a functional linear model and expanded to a generalised functional linear model consisting of both functional and scalar factors, reducing the problem to an optimisation problem of a single design matrix.

Damianos Michaelides, Antony Overstall, Dave Woods
University of Southampton
dm3g15@soton.ac.uk, a.m.overstall@soton.ac.uk, d.woods@soton.ac.uk

MS26

Exploring Risk-Averse Design Criteria for Sequential Optimal Experimental Design in a Bayesian Setting

This work explores optimal experimental design in a Bayesian setting for nonlinear models. We explore a sequential approach to optimal design that updates the design as data is collected. Additionally, we investigate various ways of defining optimality through the choice of the design criterion. Standard Bayesian approaches often measure divergence in terms of the average behavior of the distributions quantifying uncertainty in parameter estimates or model outputs. However average behavior is often not conservative enough. Here, we compare these standard approaches to risk-averse measures of divergence which aim to avoid worst case scenarios by focusing on behavior at the tails of such distributions.

Rebekah White

Sandia National Laboratories
rebwhit@sandia.gov

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

Bart G. van Bloeman Waanders
Sandia National Laboratories
Optimization and Uncertainty Quantification Department
bartv@sandia.gov

Alen Alexanderian
NC State University
alexanderian@ncsu.edu

Drew P. Kouri
Optimization and Uncertainty Quantification
Sandia National Laboratories
dpkouri@sandia.gov

Jose Huerta
Sandia National Laboratories
jghuert@sandia.gov

MS27

Shadowing-Based Data Assimilation Method for Partially Observed Models

Data assimilation is broadly used in atmosphere and ocean science to correct error in the state estimation by incorporating information from measurements into the mathematical model. The widely-used variational data assimilation method has a drawback of a drastic increase of the number of local minima of the corresponding cost function as the number of measurements increases. The shadowing approach to data assimilation, which was pioneered by K. Judd and L. Smith in Physica D (2001), aims at estimating the whole trajectory at once. It has no drawback of several local minima. However, it is computationally expensive, requires measurements of the whole trajectory, and has an infinite subspace of solutions. We propose to decrease the computational cost by projecting the shadowing approach to the unstable subspace that typically has much lower dimension than the phase space. Furthermore, we propose a novel shadowing-based data assimilation method that lifts up the requirement of a fully-observed state. We prove convergence of the method and demonstrate in numerical experiments with Lorenz models that the developed data assimilation method substantially outperforms the variational data assimilation method.

Svetlana Dubinkina
VU Amsterdam
s.b.dubinkina@vu.nl

Bart M. de Leeuw
Centrum Wiskunde & Informatica
b.m.de.leeuw@cwi.nl

MS27

Trading Information for Statistical Properties with Possibility Theory

Possibility theory allows for modelling information directly rather than through random processes; as a consequence, changing the analogue of a p.d.f., referred to as a possibility function, is permitted as long as no artificial information is created. This is an important difference between possibility theory and probability theory as it means that some approximations can be replaced by modifications of the underlying possibility functions. As an example, the information about two unknown quantities can be made "independent" by discarding the joint information, and this can be used to derive a more principled version of the naive Bayes approach. Similarly, in data assimilation, localisation can be achieved without approximation, hence avoiding any unjustified optimism in the resulting quantification of uncertainty. This talk will introduce the necessary background in possibility theory and motivate its use in practical problems via a range of convenient properties and examples.

Jeremie Houssineau
University of Warwick
jeremie.houssineau@warwick.ac.uk

MS27

Analysis of An Interacting Particle System Using Diffusion Maps for Bayesian Inference

Ensemble Kalman type methods have seen an explosion in use in data assimilation applications and more recently for a range of learning tasks. Despite their desirable stability properties, they are not consistent with Bayes theorem for non-linear, non-Gaussian systems. Recently, a range of controlled particle filters have been proposed which aim to emulate the structure of Ensemble Kalman type methods whilst simultaneously providing consistent samples in the asymptotic limit. More specifically, such filters involve constructing a control law to steer particles such that the corresponding probability distribution satisfies a variational Bayes formula. We first provide an overview of this new class of filters and how they can be used for nonlinear ensemble data assimilation and Bayesian inverse problems. A framework which allows to derive these filters will be explored, which will also highlight the main differences among them. We then discuss recent analytic and numerical work on a diffusion map based approximation of one such filter, namely the Feedback Particle Filter.

Sahani Pathiraja
University of Potsdam
pathiraja@uni-potsdam.de

MS27

Stochastic Gradient Descent with a Noise Model Inspired by Overparametrized Learning Problems

In the literature on stochastic gradient descent, there are two types of convergence results: (1) SGD finds minimizers of convex objective functions and (2) SGD finds critical points of smooth objective functions. Classical results are obtained under the assumption that the stochastic noise is L^2-bounded and that the learning rate decays to zero at a suitable speed. We show that, if the objective landscape and noise possess certain properties which are reminiscent of deep learning problems, then we can obtain global convergence guarantees of first type under second type assumptions for a fixed (small, but positive) learning rate. The convergence is exponential, but with a large random coefficient. If the learning rate exceeds a certain threshold, we discuss minimum selection by studying the invariant distribution of a continuous time SGD model. We show that at a critical threshold, SGD prefers minimizers where

the objective function is 'flat' in a precise sense.

Stephan J. Wojtowytsch
Princeton University
swoj@tamu.edu

MS28

Semiparametric Machine Learning Methods for Overcoming Model Error

Parametric modeling is the most powerful paradigm in terms of ability to fit a large model from a reasonable amount of data. This power comes at the expense of inflexibility, and even a small mismatch between the parametric form and the truth can dramatically degrade the model's usefulness. In contrast, nonparametric (data-driven) models are extremely flexible, and can provably adapt to an enormous class of systems, but they have data requirements that grow exponentially with the dimension of the underlying system. In this talk we will explore the middle-ground of semiparametric modeling techniques that leverage the strengths of parametric and nonparametric methods. We introduce a framework that allows the flexible nonparametric models to fill in the gaps and correct low-dimensional model error in a parametric model. This framework uses an ensemble of states in the parametric model to represent the uncertainty in the current forecast or state estimate, and a full probability distribution is estimated for the nonparametric model's state. We overview how a combination of sampling, interpolation, and uncertainty quantification techniques are used to link the parametric and nonparametric models. Finally, some synthetic forecasting examples illustrate how the semiparametric framework can overcome certain types of model error.

Tyrus Berry
George Mason University
tyrus.berry@gmail.com

MS28

Computing Committors in Collective Variable Using Mahalanobis Diffusion Maps

The study of rare events in molecular and atomic systems such as conformal changes and cluster rearrangements has been one of the most important research themes in chemical physics. Key challenges are associated with long waiting times rendering molecular simulations inefficient, high dimensionality impeding the use of PDE-based approaches, and the complexity of transition processes limiting the predictive power of asymptotic methods. Diffusion maps are promising algorithms to mitigate these issues. We adapt the diffusion map with Mahalanobis kernel proposed by Singer and Coifman (2008) for the SDE describing molecular dynamics in collective variables in which the diffusion matrix is position-dependent and, unlike the case considered by Singer and Coifman, is not associated with a diffeomorphism. We offer an elementary proof showing that one can approximate the generator for this SDE discretized to a point cloud via the Mahalanobis diffusion map. We use it to calculate the committor functions in collective variables for two benchmark systems: alanine dipeptide, and Lennard-Jones-7 in 2D. For validating our committor results, we compare our committors to the finite-difference solution or by conducting a "committor analysis" as used by molecular dynamics practitioners. We contrast the outputs of the Mahalanobis diffusion map with those of the standard diffusion map with isotropic kernel and show that the former gives significantly more accurate estimates for committors.

Luke Evans
University of Maryland
evansal@umd.edu

MS28

Nonlinear Model Reduction for Slow-Fast Stochastic Systems near Unknown Invariant Manifolds

We introduce a nonlinear stochastic model reduction technique for high-dimensional stochastic dy- namical systems that have a low-dimensional invariant effective manifold with slow dynamics, and high-dimensional, large fast modes. Given only access to a black box simulator from which short bursts of simulation can be obtained, we design an algorithm that outputs an estimate of the invariant manifold, a process of the effective stochastic dynamics on it, which has averaged out the fast modes, and a simu- lator thereof. This simulator is efficient in that it exploits of the low dimension of the invariant manifold, and takes time steps of size dependent on the regularity of the effective process, and therefore typically much larger than that of the original simulator, which had to resolve the fast modes. The algorithm and the estimation can be performed on-the-fly, leading to efficient exploration of the effective state space, without losing consistency with the underlying dynamics. This construction enables fast and efficient simulation of paths of the effective dynamics, together with estimation of crucial features and observables of such dynamics, including the stationary distribution, identification of metastable states, and residence times and transition rates between them.

Mauro Maggioni
Johns Hopkins University
mauromaggionijhu@icloud.com

Felix Ye
University at Albany
xye2@albany.edu

Sichen Yang
Johns Hopkins University
syang114@jhu.edu

MS29

Model Discrepancy in CO2 Retrievals from the OCO-2 Satellite

The Orbiting Carbon Observatory 2 (OCO-2) collects space-based measurements of atmospheric CO2. The CO2 measurements are indirect since the instrument observes radiances (reflected sunlight) over a range of wavelengths and a physical model is inverted, via Bayes Theorem, to estimate CO2 concentration in the atmosphere. This inference is in fact an estimation of physical parameters (an inverse problem) which can be both biased and overconfident when model error is present but not accounted for. The OCO-2 mission addresses this problem in a few different ways, e.g. with a post-inference bias correction procedure based on ground measurements. This talk will discuss methods to account for structured and informative model error directly in the inversion procedure to lessen bias and provide more reliable uncertainty estimates.

Jenny Brynjarsdottir
Case Western Reserve University
Dept. of Mathematics, Applied Mathematics and

Statistics
jxb628@case.edu

MS29

Multivariate Fused Gaussian Process for Large Spatial Data

Large multivariate spatial data sets are common in environmental and climate sciences. This article proposes a flexible multivariate spatial statistical model for such data. Built upon Ma and Kang (2020), we model multivariate spatial processes in an additive form with two components to induce spatial dependence and a relationship between distinct variables: One component is of a low-rank format, and the other component is built in a conditional way with multivariate spatial conditional autoregressive (CAR) models. By combining these two components, the resulting model not only allows for efficient computation of parameter estimation and spatial prediction but is also flexible to describe spatial covariance and cross-covariance structures that can potentially be nonstationary or asymmetric. The performance of the proposed model that we call the multivariate fused Gaussian process (MFGP) is investigated through an extensive simulation study and a real-data example. The results show that MFGP, by borrowing information from complementary data, provides substantially improved spatial predictions compared to the univariate models. We also demonstrate that MFGP outperforms the multivariate model with the low-rank component solely or with a multivariate CAR model with a separable covariance matrix.

Emily L. Kang
Department of Mathematical Sciences
University of Cincinnati
kangel@ucmail.uc.edu

Miaoqi Li
Wells Fargo
li2mq@mail.uc.edu

Kerry Cawse-Nicholson, Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology
kerry-anne.cawse-nicholson@jpl.nasa.gov,
Amy.Braverman@jpl.nasa.gov

MS29

Objective Frequentist Uncertainty Quantification for Atmospheric CO2 Retrievals

The steadily increasing amount of atmospheric carbon dioxide (CO_2) is having an unprecedented impact on the global climate system. In order to better understand the sources and sinks of CO_2, NASA operates the Orbiting Carbon Observatory-2 & 3 satellites to monitor CO_2 from space. These instruments measure the radiance of the sunlight reflected off the Earth's surface, which is then inverted in an ill-posed inverse problem to obtain estimates of the atmospheric CO_2 concentration. In this work, we propose a new CO_2 retrieval method that uses known physical constraints on the state variables and direct inversion of the target functional of interest to construct well-calibrated frequentist confidence intervals based on convex programming. We compare the method with the current operational retrieval procedure, which uses prior knowledge in the form of probability distributions to regularize the problem, and demonstrate that the proposed intervals consistently achieve the desired frequentist coverage while the operational uncertainties might be poorly calibrated both at individual locations and over a spatial region. We also study the influence of individual nuisance state variables on the length of the proposed intervals and identify certain key variables that can greatly reduce the final uncertainty given additional deterministic or probabilistic constraints, and develop a principled framework to incorporate such information into our method.

Pratik Patil, Mikael Kuusela
Carnegie Mellon University
pratik@cmu.edu, mkuusela@andrew.cmu.edu

Jonathan Hobbs
Jet Propulsion Laboratory
jonathan.m.hobbs@jpl.nasa.gov

MS29

Assessing the Impact of Uncertainty in the Orbiting Carbon Observatory-2 Estimates of CO2 Concentration

Satellites that track atmospheric greenhouse gases, such as NASAs Orbiting Carbon Observatory-2 (OCO-2), measure spectral radiances at fine spatial and temporal resolutions. Retrieval algorithms are designed to estimate the atmospheric constituent of interest, such as carbon dioxide (CO2), by modelling the process of radiative transfer through the atmosphere. The retrieved atmospheric states inform key hypotheses in carbon cycle science, but these inferences require an assessment of the sources of uncertainty in the retrieval process, including algorithm parameter choices and the inherent variability of the atmospheric states and measured radiances. One of the fundamental uses of OCO-2s data product is carbon flux inversion modelling, which relies on the satellites global coverage to produce detailed maps of CO2 sources and sinks. We present the results of a series of simulation experiments and spatial statistical analyses designed to quantify the uncertainty for individual retrievals and aggregate summaries, assessing the impact of rigorous uncertainty quantification on global flux inversion modelling.

Joaquim Teixeira, Jonathan Hobbs
Jet Propulsion Laboratory
joaquim.p.teixeira@jpl.nasa.gov,
jonathan.m.hobbs@jpl.nasa.gov

Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology
Amy.Braverman@jpl.nasa.gov

MS30

Verification of Hydrodynamic Simulation Codes for Modeling the Effects of Curvature on Detonation Propagation

The propagation speed of a curved detonation front is an important property to characterize the performance of a High Explosive (HE). Data on the detonation speed in finite dimensional charges is often used to calibrate models of HE combustion. Simulating these experiments can be computationally expensive and is often the computational bottleneck in such calibrations. Two different approaches to modeling this phenomenon are considered: the first uses an assumption of large curvature to simplify the problem into a system of Ordinary Differential Equations (ODE) in a single dimension, the second solves a system of partial

differential equations in two dimensions. The former approach is orders of magnitude faster than the latter but is only applicable for large curvatures, while the latter has no such limitations. There are three parts to this verification study. First, the implementation of the two codes are verified against exact analytic solutions for single- and two-step Arrhenius kinetics. Second, a solution verification exercise is performed on both codes to understand the effect of spatial discretization on the error in the detonation speed. Finally, the two codes are compared to identify the regime where the faster ODE solver an be used with acceptable accuracy.

Stephen A. Andrews
Los Alamos National Laboratory
saandrews@lanl.gov

Andrew K. Henrick
University of Notre Dame
Los Alamos National Laboratory
ahenrick@nd.edu

Tariq D. Aslam
Los Alamos National Laboratory
Group WX-9, Shock and Detonation Physics
aslam@lanl.gov

MS30

Code-Verification Techniques for the Method-of-Moments Implementation of the Electric-Field Integral Equation

Though the method-of-moments implementation of the electric-field integral equation plays an important role in computational electromagnetics, it provides many code-verification challenges due to the different sources of numerical error. In this work, we provide an approach through which we can apply the method of manufactured solutions to isolate and verify the solution-discretization error. We accomplish this by manufacturing both the surface current and the Green's function. Because the arising equations are poorly conditioned, we reformulate them as a set of constraints for an optimization problem that selects the solution closest to the manufactured solution. We demonstrate the effectiveness of this approach for cases with and without coding errors.

Brian A. Freno, Neil Matula, William Johnson
Sandia National Laboratories
bafreno@sandia.gov, nmatula@sandia.gov, wajohns@sandia.gov

MS30

Discretization Error Transport Equations for Unsteady Flows

Discretization errors are one of the sources of uncertainty in scientific computing. They occur when a numerical method is used to approximate the solution to a set of partial differential or integral equations, and are complicated functions of the chosen grid, time step, numerical method, and the solution. Our group at Virginia Tech has recently focused on the use of error transport equations (ETE) for providing higher-order estimates of the discretization error in the field solution as well as in derived solution functionals (usually surface integrals such as the drag force) for computational fluid dynamics applications. This talk will focus on the extension of these methods from steady-state problems to time-dependent ones. ETE offers significant advantages over more traditional approaches such as Richardson extrapolation (which requires multiple, systematically-coarsened grids) and adjoint methods (which can only provide discretization error estimates in the derived solution functionals). Furthermore, while adjoint methods applied to unsteady problems require the original problem to be solved forward in time and then the adjoint problem to be solved backward in time, we have shown that ETE can be evolved along with the original problem in time, thus requiring significantly less storage. Time-dependent results will include solutions and ETE for the linear advection-diffusion equation, viscous Burgers equation, the Euler equations, and the Navier-Stokes equations.

Christopher J. Roy
Virginia Tech, Department of Aerospace and Ocean Engineering
cjroy@vt.edu

Hongyu Wang, William Jordan, Michael Ganotaki
Crofton Department of Aerospace and Ocean Engineering
Virginia Tech
hongyuwa@vt.edu, wajordan@vt.edu, mganotak@vt.edu

MS30

Numerical Verification Procedure for Large Deformation Analyses of Hyperelasticity

In recent years, the importance of verification and validation of numerical simulations has increased in engineering fields. The objective of this work is to propose numerical verification procedures for large deformation problems of hyperelasticity, which is one of the typical problems of nonlinear solid mechanics. In conventional verification procedures to compare the numerical solutions and exact ones of second-order partial differential equations, source terms are calculated from the given artificial solutions, such as manufactured solutions. However, second-order spatial derivatives of solutions are hardly calculated, especially in nonlinear problems of solid. To circumvent such difficulty, the author developed an alternative procedure to calculate equivalent nodal vectors of body forces in the finite element method without calculation of the spatial derivative of stresses corresponding to the second-order derivative of the solution. This procedure is realized by the assumption of the weak exact solution and the calculation of equivalent nodal force vectors associated with stresses of the exact solution corresponding to the second-order spatial derivatives of the solution. We apply the proposed procedure to the method of manufactured solution and the method of nearby problems for large deformation problems of hyperelasticity and several representative numerical results are presented to discuss the effectiveness and properties of our approach.

Takahiro Yamada
Yokohama National University
tyamada@ynu.ac.jp

MS31

Bayesian Optimization of Functional Output in Inverse Problems

Motivated by the parameter identification problem of a reaction-diffusion transport model in a vapor phase infiltration processes, we propose a Bayesian optimization procedure for solving the inverse problem that aims to find an input setting that achieves a desired functional output. The proposed algorithm improves over the standard

single-objective Bayesian optimization by (i) utilizing the generalized chi-square distribution as a more appropriate predictive distribution for the squared distance objective function in the inverse problems, and (ii) applying functional principal component analysis to reduce the dimensionality of the functional response data, which allows for efficient approximation of the predictive distribution and the subsequent computation of the expected improvement acquisition function.

Chaofan Huang, Yi Ren, Emily McGuinness, Mark Losego, Ryan Lively, V. Roshan Joseph
Georgia Institute of Technology
chaofan.huang@gatech.edu, yren48@gatech.edu, emcguinness6@gatech.edu, losego@gatech.edu, ryan.lively@chbe.gatech.edu, roshan@gatech.edu

MS31

Predictive Model for the Sublimation Enthalpy of Organic Molecules Powered by Density Functional Theory and Machine Learning

A fast and accurate method to predict the sublimation enthalpy of organic and organic-inorganic hybrid molecular precursors is critical to several applications, including membrane separations, microelectronics, energy storage, and smart coatings. Density functional theory (DFT) has been widely applied in materials research due to its unparalleled success in fundamental property prediction. Nevertheless, DFT calculations tend to be time-consuming, especially for large systems. Data-driven machine learning (ML) methods are may be trained on a parent dataset (which could come from expensive DFT computations); once trained, the ML models can make property predictions almost instantaneously for new cases. In this work, we employed DFT to establish a sublimation enthalpy dataset for a benchmark set of organic molecules, and have built an ML model, trained on the DFT dataset, for the prediction of the sublimation enthalpy for new organic molecules. Our ML model was based on Gaussian process regression model. The initial ML model was coupled with different active-learning iterative cycles to progressively perform sublimation enthalpy DFT calculations, thus systematically augmenting the training dataset. Our results demonstrate that an active learning framework based on the predicted uncertainty of the ML models can systematically improve the predictive capability for the sublimation enthalpy of organic molecules in a smart way.

Yifan Liu
Georgia Institute of Technology
yifaninreallife@gatech.edu

MS31

Black-Box Optimization with a Novel Nonlocal Gradient and Its Applications

The problem of minimizing multi-modal loss functions with a large number of local optima frequently arises in model calibration, architecture design and machine learning problems. Since the local gradient points to the direction of the steepest slope in an infinitesimal neighborhood, an optimizer guided by the local gradient is often trapped in a local minimum. To address this issue, we develop a novel nonlocal gradient to skip small local minima by capturing major structures of the loss's landscape in black-box optimization. The nonlocal gradient is defined by a directional Gaussian smoothing (DGS) approach. The key idea of DGS is to conducts 1D long-range exploration with a large smoothing radius along d orthogonal directions in $\mathbb{R}^d$, each of which defines a nonlocal directional derivative as a 1D integral. Such long-range exploration enables the nonlocal gradient to skip small local minima. The d directional derivatives are then assembled to form the nonlocal gradient. We use the Gauss-Hermite quadrature rule to approximate the d 1D integrals to obtain an accurate estimator. The superior performance of our method is demonstrated in several benchmark tests as well as physics-informed machine learning problems.

Hoang A. Tran
Oak Ridge National Laboratory
Computer Science and Mathematics Division
tranha@ornl.gov

Jiaxin Zhang, Dan Lu, Guannan Zhang
Oak Ridge National Laboratory
zhangj@ornl.gov, lud1@ornl.gov, zhangg@ornl.gov

MS31

Deep-Green Inversion (DGI) to Extract Traction Separation Relationship at Material Interfaces

The traction-separation relationship of an interface is a critical component to understand and model the delamination behavior of multi-layer composites in situations where large scale bridging is active. Limited by current experimental techniques, the extraction of traction-separation relationships often relies on inverse approaches, where far field measurements are used as input data. In this work, a data-driven model based on Greens function embedded neural networks is proposed (namely Deep Green Inversion, or DGI), where the input consists of far field displacement fields while the output is the desired but initially unknown tractions and separations on the interface. Specifically, Greens functions are embedded as a loss function term along with other terms based on mean squared error and the field equations associated with loaded elastic bodies. The developed approach is then successfully validated for mode-I and mixed-mode cohesive zone extraction problems using only far-field displacement synthetic data that are generated from numerical solutions to the problems. As part of the validation process and consideration of any limitations of the approach, the amount of displacement data required to produce robust traction-separation relations is deliberated. The traction-separation relationships extracted via the DGI neural network developed here agree very well with the results obtained via a direct extraction approach.

Congjie Wei, Chenglin Wu
Missouri University of Science and Technology
cw6ck@mst.edu, wuch@mst.edu

Jiaxin Zhang
Oak Ridge National Laboratory
zhangj@ornl.gov

MS32

Entropy-Diminishing Reduced Basis Method for Cross-Diffusion Systems

In this work, we construct a reduced model for the resolution of a cross-diffusion system modelling the Physical Vapor Deposition (PVD) process within the fabrication of solar cells. Our cross diffusion system is a system of parametric nonlinear degenerated parabolic partial differential equations whose numerical resolution is nontrivial.

We employ the cell-centered finite volume method for the space discretization and the backward Euler scheme for the time discretization. To reduce the computational cost of the numerical resolution, we construct an efficient reduced model that approximates accurately the solution of our cross-diffusion problem. Furthermore, we prove that the generated reduced model respects the mathematical properties of the continuous solution, namely, the mass conservation, the positivity of the solution, the preservation of the volume filling constraint, and the decay of entropy. In the numerical experiments we test our strategy and confirm the strength of our approach.

Jad Dabaghi
CERMICS, Ecole des Ponts ParisTech
jad.dabaghi@enpc.fr

Virginie Ehrlacher
CERMICS ENPC, Paris
virginie.ehrlacher@enpc.fr

MS32

Nonlinear Model Order Reduction Using Geodesic Shooting in the Diffeomorphism Group

Parametrized hyperbolic conservation laws constitute a difficult task for classical model order reduction techniques. The solution manifold of advection dominated equations is typically highly nonlinear, and therefore, linear model order reduction methods are incapable of creating appropriate reduced order models. We propose a new approach for nonlinear model order reduction for this class of problems. The method is based on a diffeomorphic deformation of the considered space-time domain and deploys ideas from image analysis. By considering space-time solutions, shock formation or shock interaction are already incorporated in the solution snapshots. To obtain an efficient online algorithm, we represent elements from the Lie group of diffeomorphisms by elements of the Lie algebra of vector fields via the exponential map. In the Hilbert space of vector fields we can apply well-known model order reduction techniques to obtain a reduced subspace. Numerical experiments suggest that suitable linear reduced models in the space of vector fields can be created. The diffeomorphisms are finally deployed to transform reference solution snapshots.

Hendrik Kleikamp
Applied Mathematics, University of Münster
hendrik.kleikamp@uni-muenster.de

Mario Ohlberger
University of Muenster
Applied Mathematics Muenster
mario.ohlberger@uni-muenster.de

Stephan Rave
University of Muenster
stephan.rave@uni-muenster.de

MS32

Nonlinear Model Reduction for Wildland Fire Simulation

In this talk, we present a new hyperreduction scheme for a recently proposed nonlinear model order reduction (MOR) technique which is based on dynamically transformed ansatz functions. In contrast to standard model reduction methods based on linear subspaces, the considered nonlinear MOR scheme allows for an effective reduction even when applied to transport-dominated systems involving the propagation of sharp fronts. The transformations applied to the modes are parameterized by time-dependent path variables which are themselves unknowns of the reduced-order model (ROM). In the case that the full-order model (FOM) is linear, the ROM contains state-dependent coefficient matrices whose efficient offline/online decomposition is addressed. For the case of a nonlinear FOM, we extend the idea of the (discrete) empirical interpolation method by using dynamically transformed ansatz functions for the nonlinearity and we demonstrate how to achieve an efficient offline/online decomposition. Finally, we discuss the application of the new method to a nonlinear parameter-dependent model for the propagation of combustion waves in the context of wildland fire simulation.

Felix Black, Philipp Schulze
Technische Universität Berlin, Germany
black@math.tu-berlin.de, pschulze@math.tu-berlin.de

Benjamin Unger
University of Stuttgart, Germany
benjamin.unger@simtech.uni-stuttgart.de

MS33

Bi-Fidelity Training of Neural Networks Using ℓ_1-Regularization

Neural networks have recently been at the forefront of scientific machine learning (SciML) research as they are capable of accurately representing a functional relationship between inputs and a quantity of interest. However, these networks often require a large dataset during training to prevent overfitting. For engineering systems, we often have multiple models to describe the behavior. Some of these models, known as high-fidelity, are capable of describing the behavior with higher levels of accuracy, but in general, are expensive to simulate. On the other hand, inexpensive models, known as low-fidelity, often produce inaccurate predictions. In this study, sparsity promoting ℓ_1-regularization is utilized to train neural networks in the presence of a small training dataset from a high-fidelity model. Two variants of ℓ_1-regularization are used, which inform the neural network with parameters from an identical network trained using data from low-fidelity models of the problem leading to a generalization of transfer learning of neural networks. Numerical examples demonstrate that the bi-fidelity ℓ_1-regularization methods are capable of producing errors one order of magnitude smaller than errors from the networks trained only using the high-fidelity dataset.

Subhayan De, Alireza Doostan
Department of Aerospace Engineering Sciences
University of Colorado, Boulder
subhayan.de@colorado.edu,
Alireza.Doostan@Colorado.EDU

MS33

Multifidelity Active Learning for the Failure Analysis of Triso Nuclear Fuel

The TRISO nuclear fuel is a robust fuel which has been proposed to be used in several advanced reactor concepts such as microreactors. Given the importance of preventing fission products release from the nuclear fuel, accurate high-fidelity simulations of TRISO models are required for

understanding its failure behavior. However, not only is the computational model of TRISO expensive, but also, its fuel parameters can be uncertain, requiring a statistical characterization of failure. To accelerate the statistical failure analysis of TRISO, we use a multifidelity modeling strategy with active learning. Specifically, we replace the expensive TRISO model with a low-fidelity model and train a Gaussian Process to learn when a call to the expensive model is required during the statistical failure analysis. We explore different options for the low-fidelity model which include data-driven and physics-based models. Overall, across several TRISO models, the multifidelity active learning algorithm accurately predicts their failure probabilities irrespective of the choice of the low-fidelity model. However, we noticed that there are differences in the number of times the high-fidelity model is called depending upon the quality of the low-fidelity model. The algorithm also accurately characterizes the probability distributions of TRISO fuel parameters that are most likely to cause its failure. This information is important for the optimal fuel design and fabrication.

Som L. Dhulipala
Idaho National Laboratory
Som.Dhulipala@inl.gov

Michael D. Shields
Johns Hopkins University
michael.shields@jhu.edu

Wen Jiang, Benjamin Spencer
Idaho National Laboratory
wen.jiang@inl.gov, benjamin.spencer@inl.gov

Promit Chakroborty
Johns Hopkins University
pchakro1@jhu.edu

MS33

Exploration of Multifidelity UQ Methods for Monte Carlo Radiation Applications in Stochastic Media

Stochastically mixed media are present in several radiation transport (RT) problems, including inertial confinement fusion, boiling water reactors and atmospheric transport. Stochastic mixing (SM) directly impacts the possibility to perform accurate RT computations, at a reduced computational cost, by requiring ensembles realizations for the media. Moreover, uncertainty quantification (UQ) is also required for evaluating statistics of the quantities of interest in the presence of sources of uncertainty such as the relative abundance of materials or the distributions of material chunk sizes. Consequently, reliable and accurate UQ for RT problems, in the presence of SM, can easily become computationally intractable. In this contribution, we plan to investigate the use of multifidelity UQ (MF) strategies to reduce the computational burden associated to these analyses. The idea of MF UQ is to efficiently fuse the information from a large number of low-fidelity computations with an handful of high-fidelity runs. In this context, however, classical MF UQ cannot be efficiently deployed without also considering how to treat the ensemble computations, associated to the SM, to maximize the amount of correlation between models. We will focus on MF UQ strategies based on RT models defined on both the quality of the numerics and the complexity of the SM approximations. Several numerical test cases will be presented and discussed to compare MF UQ against relevant single-fidelity approaches

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

Aaron Olson
Sandia National Laboratories
aolson@sandia.gov

MS33

Multifidelity Data Fusion in Convolutional Encoder/Decoder Networks

Although an accurate high-fidelity prediction is desired, there is often insufficient high-fidelity data for data-driven methods. Capturing and leveraging the relationship between low-fidelity and high-fidelity models can enable an efficient training of these methods in a restricted data regime. In this work, we focus on convolutional neural networks to predict and quantify uncertainty from a multifidelity data ensemble. In addition to its spatial awareness, these networks have the advantage of applying the same parameters over a large space, limiting the number of parameters required to learn patterns in high-dimensional data as compared to a fully-connected network. Specifically, we use networks assembled from encoders, decoders and skip connections, allowing us the capacity to operate in data regimes where the dimensionality of the input and output varies, implying staggered encoder/decoders, and where it remains the same, implying symmetric encoder and decoder components. We also embed the network with dropblock layers to quantify the uncertainty as a Monte Carlo estimator and investigate the factors responsible for the amount of variability generated in the network output. We demonstrate our networks predictive capacity for several verification functions and problems derived from solutions to partial differential equations.

Lauren Partin
University of Notre Dame
lhensley@nd.edu

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

Ahmad A. Rushdi
Stanford University
rushdi@stanford.edu

Michael S. Eldred
Sandia National Laboratories
Optimization and Uncertainty Quantification Dept.
mseldre@sandia.gov

Daniele E. Schiavazzi
University of Notre Dame
dschiavazzi@nd.edu

MS34

Efficient Marginalization-Based Mcmc Methods for Hierarchical Bayesian Inverse Problems

Hierarchical models in Bayesian inverse problems are characterized by an assumed prior probability distribution for the unknown state and measurement error precision, and hyper-priors for the prior parameters. Combining these

probability models using Bayes' law often yields a posterior distribution that cannot be sampled from directly, even for a linear model with Gaussian measurement error and Gaussian prior, both of which we assume in this paper. In such cases, Gibbs sampling can be used to sample from the posterior, but problems arise when the dimension of the state is large. This is because the Gaussian sample required for each iteration can be prohibitively expensive to compute, and because the statistical efficiency of the Markov chain degrades as the dimension of the state increases. The latter problem can be mitigated using marginalization-based techniques, but these can be computationally prohibitive as well. In this paper, we combine the low-rank techniques of Brown, Saibaba, and Vallelian (2018) with the marginalization approach of Rue and Held (2005). We consider two variants of this approach: delayed acceptance and pseudomarginalization. We provide a detailed analysis of the acceptance rates and computational costs associated with our proposed algorithms and compare their performances on two numerical test cases - image deblurring and inverse heat equation.

Arvind Saibaba
North Carolina State University
saibab@ncsu.edu

Johnathan M. Bardsley
University of Montana
bardsleyj@mso.umt.edu

Andrew Brown
Clemson
ab7@g.clemson.edu

Alen Alexanderian
NC State University
alexanderian@ncsu.edu

MS34

Sparsity Through Hyperpriors: Solutions and Uncertainty Quantification

The search for a sparse solution of an inverse problems is naturally suited for the Bayesian framework, where sparsity, regarded as an a priori belief about the solution, can be characterized in the definition of the prior. This talk will address how different levels of sparsity can be promoted through computationally feasible hierarchical priors, and it will discuss how to account for both the sensitivity of the solution entries to the measurements and for the signal to noise ratio. Computed examples will illustrate the computational efficiency of the resulting inverse solver.

Daniela Calvetti
Case Western Reserve Univ
Department of Mathematics, Applied Mathematics and Statistic
dxc57@case.edu

Erkki Somersalo
Case Western Reserve University
ejs49@case.edu

MS34

Bayesian Inverse Problems with Horseshoe Priors for Edge-Preserving Inferences

In many large-scale inverse problems, such as computed tomography and image deblurring, characterization of sharp edges in the unknown solution is a desired attribute. Within the Bayesian approach to inverse problems, edge-preservation is oftentimes achieved using heavy-tailed Markov random field priors. Another strategy, popular in the statistical community, is the application of so-called shrinkage or sparsity-promoting priors. An advantage of this formulation lies in expressing the prior as a Gaussian distribution, depending of global and local variance hyperparameters which are endowed with heavy-tailed hyperpriors. In this talk, we revisit the shrinkage horseshoe prior and we discuss a simulation framework to solve the resulting hierarchical Bayesian inverse problem. Our formulation uses a Gibbs sampler, where the Gaussian conditional is sampled efficiently with an optimization algorithm, while the heavy-tailed conditionals are simulated with the NUTS variant of the Hamiltonian Monte Carlo method. Our numerical procedure is able to estimate the solution in an efficient manner while achieving edge-preservation. A simple deconvolution model is used to illustrate the approach.

Felipe Uribe, Yiqiu Dong
Technical University of Denmark
Department of Applied Mathematics and Computer Science
furca@dtu.dk, yido@dtu.dk

Per Christian Hansen
Technical University of Denmark
DTU Compute
pcha@dtu.dk

MS35

Multilevel Representations of Isotropic Gaussian Random Fields on the Sphere

Representation of random fields on hypersurfaces is important in many applications such as (cell) biology, climatology and astrophysics. We will construct series expansion of isotropic Gaussian random field on a sphere with independent Gaussian coefficients and localized basis functions. The basis functions are obtained by applying the square root of the covariance operator to spherical needlets. The localization property is especially useful in adaptive algorithms. In addition, we will present numerical illustrations and an application to random elliptic PDEs on the sphere.

Ana Djurdjevac
Freie Universität Berlin
anadjurdjevac@gmail.com

Markus Bachmayr
Université Pierre et Marie Curie
bachmayr@ljll.math.upmc.fr

MS35

A Higher Order Perturbation Approach for Electromagnetic Scattering Problems on Random Domains

We consider time-harmonic electromagnetic scattering problems on perfectly conducting scatterers with uncertain shape. Thus, the scattered field will also be uncertain. Based on the knowledge of the two-point correlation of the domain boundary variations around a reference domain, we derive a perturbation analysis for the mean of the scattered field. Therefore, we compute the second shape derivative of the scattering problem for a single perturbation. Tak-

ing the mean, this leads to an at least third order accurate approximation with respect to the perturbation amplitude of the domain variations. To compute the required second order correction term, a tensor product equation on the domain boundary has to be solved. We discuss its discretization and efficient solution using boundary integral equations. Numerical experiments in three dimensions are presented.

Jürgen Dölz
University of Bonn
doelz@ins.uni-bonn.de

MS35

Population-based Image Segmentation and Shape Regression

In medical image analysis, population-based shape analysis allows us to distinguish disease groups from normal controls. To achieve this, extracting regions of interest, i.e., shapes, from medical images, and estimating the relationship between shapes and their associated variables, e.g., age, are the two fundamental tasks to address. However, we are suffering from limited data and labels, as well as lacking efficient algorithms for performing shape regression and uncertainty quantification. In this talk, I will present our recent works on these topics, and demonstrate their effectiveness in studying brain disease.

Yi Hong
Shanghai Jiaotong University
yi.hong@sjtu.edu.cn

MS35

Lesa: Longitudinal Elastic Shape Analysis of Brain Subcortical Structures

Over the past 30 years, magnetic resonance imaging has become a ubiquitous tool for accurately visualizing the change and development of the brain subcortical structures (e.g., hippocampus) across time and group. Although subcortical structures act as information hubs of the nervous system, their quantification is still in its infancy due to challenges in shape extraction, representation, and modeling. Here, we develop a simple and efficient framework for longitudinal elastic shape analysis (LESA) of subcortical structures. In LESA, subcortical structures are segmented, extracted, and represented as parameterized 3-dimensional (3D) surfaces. Integrating ideas from elastic shape analysis of static surfaces, principal component analysis (PCA) of shapes, and statistical modeling of sparse longitudinal data, LESA provides a unique toolbox for systematically quantifying the development and changes of longitudinal subcortical surface shapes. The key novelties of LESA include: (i) it can efficiently capture complex subcortical structures using a small number of basis functions and (ii) it can accurately predict the spatiotemporal shape changes of the human subcortical structures. We applied LESA to analyze three longitudinal neuroimaging data sets and showcase its wide applications in estimating continuous shape trajectories, building life-span growth patterns, and comparing shape differences among different groups.

Zhengwu Zhang
UNC Chapel Hill
zz10c@email.unc.edu

MS36

Bayesian Excursions in the Treatment of Model Error

Advances in observational and computational assets have led to revolutions in the range and quality of results in many science and engineering settings. However, needs for treating model errors and assessing their impacts have kept pace with those advances. In this manuscript, we discuss a general Bayesian strategy to incorporate and assess model error in complex, physical-statistical models. We distinguish between two classes of models: (i) sufficiently manageable models which allow for a model error process to be incorporated into the dynamics of the field of interest and (ii) very high-dimensional settings where computer model output is treated as biased data. In each case, we illustrate the benefits of accounting for model error through an oceanographic example.

Radu Herbei
The Ohio State University
herbei@stat.osu.edu

MS36

Bayesian Spanning Treed Gaussian Process for High-Dimensional Output Simulators and Its Extensions

Motivated by a real-life application of a Storm Surge simulator with high-dimensional outputs, we propose a new Bayesian emulator called Bayesian Spanning treed Gaussian process suitable to analyze computer models with non-stationary massive outputs in the input domain. Central to the design of the Bayesian model is the idea of coupling Bayesian spanning tree algorithms with parallel partial Gaussian process regression model. The proposed approach is further extended to the co-kriging setting enabling the analysis of multi-fidelity computer models with non-stationary outputs. Given certain assumptions on the Bayesian model, we introduce a suitable stochastic mechanism that facilitates predictions in a principal manner. The good performance of our method is demonstrated in a benchmark example, while our method is used for the analysis of a surge simulator in the multi-fidelity setting.

Kieran Richards, Georgios Karagiannis
University of Durham
kieran.96@outlook.com, georgios.stats@gmail.com

Alex Konomi
University of Cincinnati
alex.konomi@uc.edu

MS36

Computationally Efficient Statistical Emulators for Complex Forward Models in Remote Sensing: An Application to the Oco-2 Mission

Observing system uncertainty experiments (OSUEs) have been recently proposed as a cost-effective way to perform probabilistic assessment of retrievals for NASAs Orbiting Carbon Observatory-2 (OCO-2) mission. One important component in the OCO-2 retrieval algorithm is a full-physics forward model that describes the mathematical relationship between atmospheric variables such as carbon dioxide and radiances measured by the remote sensing in-

strument. This complex forward model is computationally expensive but large-scale OSUEs require evaluation of this model numerous times, which makes it infeasible for comprehensive experiments. To tackle this issue, we develop a statistical emulator to facilitate large-scale OSUEs in the OCO-2 mission with independent emulation. The proposed emulator reduces dimensionality in the input space and explores a full scale approximation for the output space. The full scale approximation for the output space accounts both the large and small scale variation in a lower space input dimension. Within this approximation model, we develop a computationally efficient Bayesian inference framework. We also explore different strategies for the input dimension reduction.

Alex Konomi
University of Cincinnati
alex.konomi@uc.edu

Emily L. Kang
Department of Mathematical Sciences
University of Cincinnati
kangel@ucmail.uc.edu

Jonathan Hobbs
Jet Propulsion Laboratory
jonathan.m.hobbs@jpl.nasa.gov

MS36

Auto-Differentiable Ensemble Kalman Filters

Data assimilation is concerned with sequentially estimating a temporally-evolving state. This task, which arises in a wide range of scientific and engineering applications, is particularly challenging when the state is high-dimensional and the state-space dynamics are unknown. In this talk we will introduce a machine learning framework for learning dynamical systems in data assimilation. Our auto-differentiable ensemble Kalman filters (AD-EnKFs) blend ensemble Kalman filters for state recovery with machine learning tools for learning the dynamics. In doing so, AD-EnKFs leverage the ability of ensemble Kalman filters to scale to high-dimensional states and the power of automatic differentiation to train high-dimensional surrogate models for the dynamics. Numerical results using the Lorenz-96 model show that AD-EnKFs outperform existing methods that use expectation-maximization or particle filters to merge data assimilation and machine learning. In addition, AD-EnKFs are easy to implement and require minimal tuning.

Yuming Chen, Daniel Sanz-Alonso, Rebecca Willett
University of Chicago
ymchen@uchicago.edu, sanzalonso@uchicago.edu, willett@uchicago.edu

MS37

Uncertainty Quantification for Stochastic Neural Networks

Abstract: Uncertainty quantification of deep neural networks (DNN) is a critical issue in deep learning. In our UQ for DNN framework, the DNN architecture is the neural ordinary differential equations (Neural-ODE), which formulates the evolution of potentially huge hidden layers in the DNN as a discretized ordinary differential equation (ODE) system. To characterize the randomness caused by the uncertainty of models and noises of data, we add a multiplicative Brownian motion noise to the ODE as a stochastic diffusion term, which changes the ODE to a stochastic differential equation (SDE). The deterministic DNN becomes a stochastic neural network (SNN). In the SNN, the drift parameters serve as the prediction of the network, and the stochastic diffusion governs the randomness of network output, which helps quantify the epistemic uncertainty of deep learning. I will present results on convergence as well as numerical experiments.

Yanzhao Cao
Department of Mathematics & Statistics
Auburn University
yzc0009@auburn.edu

Richard Archibald
Computational Mathematics Group
Oak Ridge National Labratory
archibaldrk@ornl.gov

Feng Bao
Florida State University
bao@math.fsu.edu

MS37

Ultrafast data reduction at the extreme edge for particle physics and beyond

We discuss applications and opportunities for the real-time, low-latency data reduction using machine learning in embedded systems. This talk will focus on particle physics examples spanning application spaces from front-end data compression to real-time feature extraction for data filtering to experimental control in systems with FPGAs and ASICs with algorithm latencies as low as 25 nanoseconds. We will review essential methods for designing, optimizing, and validating algorithms in hardware and emerging tool flows to accelerate algorithm development. We will also discuss first explorations into the robustness and generalizability of such techniques.

Tran Nhan
Fermi National Accelerator Laboratory
ntran@fnal.gov

MS37

Quantifying Uncertainty in E3SM via Functional Tensor Network Approximations

The predictive performance of the Energy Exascale Earth System Model (E3SM) is challenged by the modeling choices for a large ensemble of physical processes. This results in a large number of uncertain parameters and computationally expensive numerical simulations which makes both forward and inverse uncertainty quantification studies difficult. To overcome these challenges, we will focus on constructing surrogate models that exploit the model structure via low-rank functional tensor networks approximations. We will approximate the components of the large scale model with functional forms tailored to the model behavior. We will then cast the training of the functional tensor network model in a Bayesian framework and use a Stein variational inference approach to construct a probabilistic model that approximates the discrepancy between the surrogate and the original model predictions. We will focus on the land model component of E3SM and present results pertaining to global sensitivity analysis and model calibration at a regional scale.

Cosmin Safta

Sandia National Laboratories
csafta@sandia.gov

MS37

Empirical Bayesian Inference Using a Support Informed Sparse Prior

We develop a new empirical Bayesian inference algorithm for solving a linear inverse problem given multiple measurement vectors (MMV) of under-sampled and noisy observable data. Specifically, by exploiting the joint sparsity across the multiple measurements in the sparse domain of the underlying signal or image, we construct a new support informed sparsity promoting prior. While a variety of applications can be modeled using this framework, in this talk we discuss classification and target recognition from synthetic aperture radar (SAR) data which are acquired from neighboring aperture windows. Our numerical experiments demonstrate that using this new prior not only improves accuracy of the recovery, but also reduces the uncertainty in the posterior when compared to standard sparsity producing priors. We also discuss how our method can be used to combine and register different types of data acquisition. This is joint work with Theresa Scarnati formerly of the Air Force Research Lab Wright Patterson and now working at Qualis Corporation in Huntsville, AL, and Jack Zhang, recent bachelor degree recipient at Dartmouth College and now enrolled at University of Minnesotas PhD program in mathematics.

Theresa A. Scarnati
Air Force Research Laboratory
tscarnati24@gmail.com

Anne Gelb
Dartmouth College
Department of Mathematics
annegelb@math.dartmouth.edu

MS38

Uncertainty Quantification and Correctability for Directed Graphical Models and Applications in Materials Design

In this talk, we focus on directed graphical models as they allow us to integrate in a natural way expert knowledge, physical modeling, heterogeneous and correlated data and quantities of interest (QoI). For exactly this reason, multiple sources of model uncertainty are inherent within the modular structure of the graphical model. We present information-theoretic, robust uncertainty quantification methods and non-parametric stress tests for directed graphical models to assess the effect and the propagation through the graph of multi-sourced model uncertainties to QoIs. We use these methods to rank the different sources of uncertainty and provide a mathematically rigorous approach to correctability that guarantees a systematic selection for improvement of the most under-performing components of a graphical model while controlling potential new errors created in the process in other parts of the model. We demonstrate these methods in Bayesian networks built for trustworthy prediction of materials screening to increase the efficiency of chemical reactions in catalysis. Based on the Sabatier's principle, the optimal oxygen binding energy has to be our QoI while a Bayesian network is built from expert knowledge (volcano curves), as well as various available experimental and computational data and their correlations or conditional independence. We quantify its model uncertainties, we rank their impact from least to most influential, and we correct the most under-performing ones.

Panagiota Birmpa
University of Massachusetts Amherst
birmpa@math.umass.edu

Jinchao Feng
John Hopkins University
jfeng34@jhu.edu

Markos A. Katsoulakis
University of Massachusetts, Amherst
Dept of Mathematics and Statistics
markos@math.umass.edu

Luc Rey-Bellet
Department of Mathematics & Statistics
University of Massachusetts
luc@math.umass.edu

MS38

Rons: Reduced-Order Nonlinear Solutions for Pdes with Conserved Quantities

Reduced-order models where the solution is assumed as a linear combination of prescribed modes are rooted in a well-developed theory. However, more general models where the reduced solutions depend nonlinearly on time-dependent variables have thus far been derived in an ad hoc manner. Here, we introduce Reduced-order Nonlinear Solutions (RONS): a unified framework for deriving reduced-order models that depend nonlinearly on a set of time-dependent variables. The set of all possible reduced-order solutions are viewed as a manifold immersed in the function space of the PDE. The parameters are evolved such that the instantaneous discrepancy between reduced dynamics and the full PDE dynamics is minimized. In the special case of linear parameter dependence, our reduced equations coincide with the standard Galerkin projection. Furthermore, any number of conserved quantities of the PDE can readily be enforced in our framework. Since RONS does not assume an underlying variational formulation for the PDE, it is applicable to a broad class of problems, and in particular fluid dynamics. We demonstrate the efficacy of RONS on three fluid flows: an advection-diffusion equation, the propagation of surface waves, and vortex dynamics in ideal fluids.

Mohammad Farazmand
Department of Mathematics
NC State University
farazmand@ncsu.edu

William Anderson
Department of Mathematics
North Carolina State University,
wmander3@ncsu.edu

MS38

Dynamical Low Rank Approximation for High-Dimensional Data Assimilation

In this talk we present our recent advances in the topic of efficient data assimilation involving large-dimensional chaotic random dynamical systems, where non-Guassian features are captured via the Dynamical low rank approximation (DLRA). DLRA, also called Dynamically orthogo-

nal (DO) approximation can be viewed as a reduced basis method, thus solvable at a relatively low computational cost, in which the solution is expanded as a linear combination of few well chosen deterministic functions with random coefficients. The peculiarity of the DLR method is that the spatial basis is computed on the fly and is free to evolve in time, thus adjusting at each time to the current structure of the random solution. First, we consider the signal to be approximated by a simple DLR method in the forecast step and examine various techniques of performing the analysis step in the DLR subspace. Then, we present two possibilities of enriching the DLR solution with a Gaussian approximation evolving in the remaining part of the phase space, and discuss how to perform the analysis step which eventually results in an updated DLR subspace.

Fabio Nobile
EPFL, Switzerland
fabio.nobile@epfl.ch

Yoshihito Kazashi
École polytechnique fédérale de Lausanne
y.kazashi@uni-heidelberg.de

Kody Law
University of Manchester
kodylaw@gmail.com

Vidlickova Eva
EPFL
eva.vidlickova@epfl.ch

MS38

Statistical Reduced-Order Models and Machine Learning-Based Closure Strategies for Turbulent Dynamical Systems

The capability of using imperfect statistical reduced-order models to capture crucial statistics in complex turbulent systems is investigated. Much simpler and more tractable block-diagonal models are proposed to approximate the complex and high-dimensional turbulent dynamical equations using both parameterization and machine learning strategies. A systematic framework of correcting model errors with empirical information theory is introduced, and optimal model parameters under this unbiased information measure can be achieved in a training phase before the prediction. It is demonstrated that crucial principal statistical quantities in the most important large scales can be captured efficiently with accuracy using the reduced-order model in various dynamical regimes of the flow field with distinct statistical structures. In addition, new machine learning strategies are proposed to learn the expensive unresolved processes directly from data.

Di Qi
Purdue University
qidi@purdue.edu

Andrew Majda
Courant Institute NYU
jonjon@cims.nyu.edu

John Harlim
Pennsylvania State University
jharlim@psu.edu

MS39

Multi-Objective Bayesian Optimization over High-Dimensional Search Spaces

The ability to optimize multiple competing objective functions with high sample efficiency is imperative in many applied problems across science and industry. Multi-objective Bayesian optimization (BO) achieves strong empirical performance on such problems, but even with recent methodological advances, it has been restricted to simple, low-dimensional domains. Most existing BO methods exhibit poor performance on search spaces with more than a few dozen parameters. In this work we propose MORBO, a method for multi-objective Bayesian optimization over high-dimensional search spaces. MORBO performs local Bayesian optimization within multiple trust regions simultaneously, allowing it to explore and identify diverse solutions even when the objective functions are difficult to model globally. We show that MORBO significantly advances the state-of-the-art in sample-efficiency for several high-dimensional synthetic and real-world multi-objective problems, including a vehicle design problem with 222 parameters, demonstrating that MORBO is a practical approach for challenging and important problems that were previously out of reach for BO methods.

Sam Daulton
Facebook
sdaulton@fb.com

David Eriksson
UBER
deriksson@fb.com

Maximilian Balandat, Eytan Bakshy
Facebook
balandat@fb.com, ebakshy@fb.com

MS39

Leveraging Replication in Sequential Design Tasks

An increasing number of time-consuming simulators exhibit a complex noise structure that depends on the inputs. Advances in Gaussian process modeling with input-dependent noise, especially via replication (iid repetitions of the same experiment), allow efficient modeling with better uncertainty quantification on the predictions. We focus here on strategies for balancing replication and exploration for various sequential design goals, possibly in parallel batches. These goals include global model accuracy, optimization, contour finding and dimension reduction, thus motivating further approaches detailed along this symposium. Illustration on synthetic examples are provided as well as a large scale massively parallel real world epidemiology problem.

Mickael Binois
INRIA
mickael.binois@inria.fr

MS39

Combining Calibrated Simulator with Observations to Infer Photometric Redshift

We are interested in obtaining a full probabilistic description of the redshift for cosmological objects based on pho-

tometric surveys. However, such surveys are limited and do not span the entire color spectrum which creates gaps in the input space for any statistical model to effectively reproduce the entire response surface. To overcome this challenge, the problem of redshift estimation is posed as an inverse problem where the observed photometric surveys are augmented with a simulation that, when calibrated, fills in the gaps in the prior training data and aids in quantifying the uncertainty in redshift estimation. We will discuss the emulation and calibration of this function input–function output simulator and how such simulation augmentation technique improves uncertainty quantification in photometric redshift estimation.

Arindam Fadikar
Argonne National Laboratory
afadikar@anl.gov

MS39

Scalable High-Dimensional Bayesian Optimization

Bayesian optimization has become a powerful method for the sample-efficient optimization of expensive black-box functions. These functions do not have a closed-form and are evaluated for example by running a complex economic simulation, by an experiment in the lab or in a market, or by a CFD simulation. Usecases arise in machine learning, e.g., when tuning the configuration of an ML model or when optimizing a reinforcement learning policy. Of particular interest are high-dimensional constrained settings, where we are looking for a solution that satisfies inequality constraints of the form c(x) ¡= 0 and is globally optimal for the objective function among all feasible solutions. These problems are difficult for current approaches due to the curses of dimensionality, the heterogeneity of the underlying functions, and the often small and non-convex sets of feasible points. Due to the lack of sample efficient methods, practitioners usually fall back to evolutionary strategies or heuristics. In this talk I will start with a brief introduction to Bayesian optimization and then present scalable Bayesian optimization algorithms that address the above challenges via local surrogates and sampling strategies. Moreover, I will show comprehensive experimental results that demonstrate that the proposed methods achieve excellent results and outperform the state-of-the-art methods.

Matthias Poloczek
Amazon EconTech
matthias.poloczek@gmx.de

MS40

Informative Path Planning for Anomaly Detection in Environment Exploration and Monitoring

An unmanned autonomous vehicle (UAV) is sent on a mission to explore and reconstruct an unknown environment from a series of measurements collected by Bayesian optimization. The success of the mission is judged by the UAVs ability to faithfully reconstruct any anomalous features present in the environment, with emphasis on the extremes (e.g., extreme topographic depressions or abnormal chemical concentrations). We show that the criteria commonly used for determining which locations the UAV should visit are ill-suited for this task. We introduce a number of novel criteria that guide the UAV towards regions of strong anomalies by leveraging previously collected information in a mathematically elegant and computationally tractable manner. We demonstrate superiority of the proposed approach in several applications, including reconstruction of seafloor topography from real-world bathymetry data, as well as tracking of dynamic anomalies. A particularly attractive property of our approach is its ability to overcome adversarial conditions, that is, situations in which prior beliefs about the locations of the extremes are imprecise or erroneous.

Antoine Blanchard
Massachusetts Institute of Technology
ablancha@mit.edu

Themistoklis Sapsis
Massachusetts Institute of Techonology
sapsis@mit.edu

MS40

Surrogate-Based Sequential Bayesian Experimental Design Using Non-Stationary Gaussian Processes

Inferring arbitrary quantities of interest (QoI) using limited computational or, in realistic scenarios, financial budgets, is a challenging problem that requires sophisticated strategies for the optimal allocation of the available resources. Bayesian optimal experimental design identifies the optimal set of design locations for the purpose of solving a parameter inference problem and the optimality criterion is typically associated with maximizing the worth of information in the experimental measurements. Sequential design strategies further identify the optimal design in a sequential manner, starting from a initial budget and iteratively selecting new optimal points until either an accuracy threshold is reached, or a cost limit is exceeded. In this paper, we present a generic sequential Bayesian experimental design framework that relies on maximizing an information theoretic design criterion, namely the Expected Information Gain, in order to infer QoIs formed as nonlinear operators acting on black-box functions. Our framework relies on modeling the underlying response function using non-stationary Gaussian Processes, thus enabling efficient sampling from the QoI in order to provide Monte Carlo estimators for the design criterion. We demonstrate the performance of our method on an engineering problem of steel wire manufacturing and compare it with two classic approaches: uncertainty sampling and expected improvement.

Piyush Pandita
School of Mechanical Engineering, Purdue University, West Lafayette, Indiana, 47906
piyush.pandita@ge.com

Panagiotis Tsilifis
Ecole Polytechnique Fédérale de Lausanne
pantsili@gmail.com

Nimish Awalgaonkar, Ilias Bilionis
Purdue University
nawalgao@purdue.edu, ibilion@purdue.edu

MS40

Data-Efficient, Adaptive Learning in Optimization under Uncertainty: Applications in Materials' Design

A fully integrated approach to the design and optimization of novel materials requires modeling the entire process-structure-property chain in a manner that accounts for the

multitude of uncertainties that are present. This gives rise to a stochastic inversion problem, as optimality can only be defined in terms of an appropriate expectation. While the complexity and cost of the structure-property map implies that one has to resort to data-driven surrogates, this in itself poses a considerable challenge; changes in process parameters can cause radical changes in the resulting microstructures and invalidate surrogates dependent on representative training data. To overcome this bottleneck we propose a data-efficient, adaptive learning strategy coupled with an Expectation-Maximization algorithm in order to drive the stochastic inversion problem. An acquisition function based on the design objective scores candidate microstructures according to their ability to inform and further guide the material design towards optimality, i.e. additional, informative microstructures and their effective physical properties are identified and constructed on the fly. We demonstrate for two-phase binary random media the adaptive data-refinement of a surrogate based on a convolutional neural network, and illustrate not only the ability to successfully perform the stochastic inversion of the entire process-structure-property chain, but also to considerably mitigate the dependence on expensive data.

Maximilian Rixner
Technical University of Munich
Professorship of Continuum Mechanics
maximilian.rixner@tum.de

Phaedon-Stelios Koutsourelakis
Technical University of Munich
p.s.koutsourelakis@tum.de

MS41

Adaptive Tikhonov Strategies for Stochastic Ensemble Kalman Inversion

Ensemble Kalman inversion (EKI) is a derivative-free optimizer aimed at solving inverse problems, taking motivation from the celebrated ensemble Kalman filter. The purpose of this article is to consider the introduction of adaptive Tikhonov strategies for EKI. This work builds upon Tikhonov EKI (TEKI) which was proposed for a fixed regularization constant. By adaptively learning the regularization parameter, this procedure is known to improve the recovery of the underlying unknown. For the analysis, we consider a continuous-time setting where we extend known results such as well-posdeness and convergence of various loss functions, but with the addition of noisy observations. Furthermore, we allow a time-varying noise and regularization covariance in our presented convergence result. In turn we present three adaptive regularization schemes, which are highlighted from both the deterministic and Bayesian approaches for inverse problems, which include bilevel optimization, the MAP formulation and covariance learning. We numerically test these schemes and the theory on linear and nonlinear partial differential equations, where they outperform the non-adaptive TEKI and EKI.

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

Claudia Schillings
University of Mannheim
Institute of Mathematics
c.schillings@uni-mannheim.de

Xin T. Tong
National University of Singapore
Department of Mathematics
tongthomson@gmail.com

Simon Weissmann
University of Mannheim
simon.weissmann@uni-heidelberg.de

MS41

Mean-Field Approximations to Filtering and Inverse Problems

Recently, renewed interest in classical methods for filtering has been observed using tools from kinetic theory. In particular, we will present results on the ensemble Kalman filter (EnKF) in application towards inverse problems and Bayesian inversion. Using meanfield approximation we study convergence and stability of the filtering method and develop a novel modified and unconditionally stable version of the EnKF. Numerical results and comparison with existing methods will be presented.

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

MS41

On Unbiased Discretization for Discretized Models

In this work, we consider computing expectations w.r.t. probability measures which are subject to discretization error. Examples include partially observed diffusion processes or inverse problems, where one may have to discretize time and/or space, in order to practically work with the probability of interest. Given access only to these discretization, we consider the construction of unbiased Monte Carlo estimators of expectations w.r.t. such target probability distributions. It is shown how to obtain such estimators using a novel adaptation of randomization schemes and Markov simulation methods. Under appropriate assumptions, these estimators possess finite variance and finite expected cost. There are two important consequences of this approach: (i) unbiased inference is achieved at the canonical complexity rate, and (ii) the resulting estimators can be generated independently, thereby allowing strong scaling to arbitrarily many parallel processors. Several algorithms are presented, and applied to Bayesian inverse problems

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

MS41

The Effects of Subsampling in Observation Space on the Accuracy and Stability of Continuous Square Root Filters

The family of Ensemble Square Root filters (ESRFs) is very popular in many application areas as they are known to be computationally feasible/robust and capable of signal tracking even in nonlinear and high dimensional settings. A large class of these discrete filters can, by taking a time limit, be linked to a deterministic version of the Ensemble Kalman Bucy filter (EnKBF). The long-time stability and accuracy of this EnKBF has recently been investigated for an idealised setting, specifically the underlying system is assumed to be fully observed. Here we study

the properties of the continuous filter for subsampled data (i.e., partial observations), orthogonal transformations and misspecified parameters. Subsampling as well as orthogonal transformations have in practice been identified to increase robustness. Analyzing these techniques in combination with filtering with respect to accuracy and stability allows to gain key insights into properties of the EnKBF and by extension into properties of the large family of ES-RFs

Jana de Wiljes
University of Potsdam
wiljes@uni-potsdam.de

MS42

Multi-fidelity modeling of multi-scale porosity defects in cast alloys

Cast alloys often contain heterogeneous pores that significantly affect alloy behavior in high-performance applications. To understand the cross-scale impact of microscale porosity characteristics on the cast components macro-mechanical properties, expensive multi-scale simulations are typically required. In this talk, we will introduce a multi-fidelity and multiscale framework to simulate the behavior of metallic components containing process-induced pores. Major components of our approach include: (1) a porosity-oriented 3D microstructure reconstruction algorithm which mimics the materials local heterogeneity, (2) a novel mechanistic reduced-order model which significantly reduces computational costs by projecting solution variables into a lower dimensional space where the materials elasto-plastic behaviors are approximated, and (3) a novel multi-fidelity method based on latent-map Gaussian processes that streamline assimilating any number of data sources with variable fidelity level.

Ramin Bostanabad
Northwestern University
raminb@uci.edu

Shiguang Deng, Tammer Labolle
University of California Irvine
s.deng@uci.edu, jeweisla@uci.edu

MS42

Efficient Bayesian Inference Approaches for Matrix and Tensor Learning: Applications to Surrogate Modeling of Complex Systems

Matrix/tensor completion is the problem of recovering a matrix or tensor from a fraction of its entries. It is a classical problem with many practical applications, such as designing recommender systems and predicting drug-protein interactions. Additionally, in the context of data-driven modeling, we can also pose the exploration of discretized parameter spaces as matrix/tensor completion problems. The development of Bayesian inference models based on low-rank factorizations is an active research area. These probabilistic approaches provide robust uncertainty estimates on the recovered matrix entries, which is a major advantage over classical optimization-based approaches. A standard strategy in the literature is to assign zero-mean Gaussian priors on the columns or rows of factor matrices to create a conjugate system. This choice of prior leads to simple implementations; however, it also causes symmetries in the posterior distribution that can severely reduce the efficiency of Markov-chain Monte-Carlo (MCMC) sampling approaches. In this talk, we propose simple modifications to the prior choice; these provably break the posterior symmetries and maintain/improve the accuracy of the predictions. Specifically, we provide conditions on the prior parameters that break the posterior symmetries. For example, we show that using non-zero linearly independent prior means significantly reduces the autocorrelation of MCMC samples, and can also lead to smaller reconstruction errors.

Saibal De
Sandia National Laboratories, U.S.
sde@sandia.gov

Alex Gorodetsky
University of Michigan
goroda@umich.edu

MS42

Moopsnet: Learning Multiple-Output Operators Defined on Product Spaces Simultaneously

Very recently, several approaches have been developed to learn PDE solution operators by using neural networks such as DeepONet and Fourier neural operator. DeepONet has been used successfully in many applications, including various explicit operators, such as integrals and fractional Laplacians, as well as implicit operators that represent deterministic and stochastic differential equations. More generally, DeepONet can learn multiphysics and multiscale operators spanning across many scales. However, these approaches only consider a single input function. Here, we propose the first neural network that can learn multiple-output operators defined on product spaces simultaneously.

Lu Lu, Shuai Meng
University of Pennsylvania
lulu1@seas.upenn.edu, shuaim@seas.upenn.edu

Pengzhan Jin
Peking University
jpz@math.pku.edu.cn

MS42

Bayesian Inverse Uncertainty Quantification of Nuclear Reactor Simulators

The Best Estimate plus Uncertainty (BEPU) approach for nuclear systems modeling and simulation requires that the prediction uncertainty must be quantified in order to prove that the investigated design stays within acceptance criteria. A rigorous Uncertainty Quantification (UQ) process should simultaneously consider multiple sources of quantifiable uncertainties: (1) parameter uncertainty due to randomness or lack of knowledge; (2) experimental uncertainty due to measurement noise; (3) model uncertainty caused by missing/incomplete physics and numerical approximation errors, and (4) code uncertainty when surrogate models are used. We performed Bayesian inverse UQ on a nuclear reactor thermal-hydraulics code, TRACE, based on BFBT data. We also integrated results from inverse UQ and quantitative validation to provide robust predictions so that all these sources of uncertainties can be taken into consideration. The quantitative validation metric is based on Bayesian hypothesis testing. The resulting metric, called the Bayes factor, is used to form weighting factors to combine the prior and posterior knowledge of the parameter uncertainties in a Bayesian model averaging process. In this way, model predictions will be able to integrate the

results from inverse UQ and validation to account for all available sources of uncertainties.

Xu Wu, Ziyu Xie, Farah Alsafadi
North Carolina State University
xwu27@ncsu.edu, zxie22@ncsu.edu, fralsafa@ncsu.edu

MS43

Simple, Low-Cost and Accurate Data-Driven Geophysical Forecasting with Learned Kernels

Modelling geophysical processes as low-dimensional dynamical systems and regressing their vector field from data is a promising approach for learning emulators of such systems. We show that when the kernel of these emulators is also learned from data (using kernel flows, a variant of cross-validation), then the resulting data-driven models are not only faster than equation-based models but are easier to train than neural networks such as the long short-term memory neural network. In addition, they are also more accurate and predictive than the latter. When trained on geophysical observational data, for example the weekly averaged global sea-surface temperature, considerable gains are also observed by the proposed technique in comparison with classical partial differential equation-based models in terms of forecast computational cost and accuracy. When trained on publicly available re-analysis data for the daily temperature of the North American continent, we see significant improvements over classical baselines such as climatology and persistence-based forecast techniques. Although our experiments concern specific examples, the proposed approach is general, and our results support the viability of kernel methods (with learned kernels) for interpretable and computationally efficient geophysical forecasting for a large diversity of processes.

Romit Maulik
Argonne National Laboratory
rmaulik@anl.gov

Boumediene Hamzi
Imperial College London
b.hamzi@imperial.ac.uk

Houman Owhadi
Applied Mathematics
Caltech
owhadi@caltech.edu

MS43

Learning Multiple-Output Operators Defined on Product Spaces Simultaneously

Very recently, several approaches have been developed to learn PDE solution operators by using neural networks such as deep operator networks (DeepONet) and Fourier neural operator. DeepONet has been used successfully in many applications, including various explicit operators, such as integrals and fractional Laplacians, as well as implicit operators that represent deterministic and stochastic differential equations. More generally, DeepONet can learn multiphysics and multiscale operators spanning across many scales. However, these approaches only consider a single input function. Here, we propose the first neural network that can learn multiple-output operators defined on product spaces simultaneously.

Shuai Meng, Lu Lu
University of Pennsylvania
shuaim@seas.upenn.edu, lulu1@seas.upenn.edu

Pengzhan Jin
Chinese Academy of Sciences
jpz@lsec.cc.ac.cn

MS43

Operator Learning for Inverse Problems

In the present work, we develop a model to learn the map from operators to functions. The model is employed in the contexts of inverse problem for differential equations with unknown function coefficients. In particular, we aim at learning the underlying map from the Dirichlet-to-Neumann operator, defined for a wide collection of boundary data, to the equation coefficient. The use of this model is supported by rigorously proven statements demonstrating the uniqueness and well-posedness of such map. Moreover, we also apply the same idea in the context of physics informed neural networks and present extensive numerical experiments.

Roberto Molinaro
Seminar for Applied Mathematics
ETH Zurich
roberto.molinaro@sam.math.ethz.ch

Michael Prasthofer
ETH Zurich, Switzerland
michael.prasthofer@sam.math.ethz.ch

MS43

PDE-constrained optimization via self-supervised operator learning

Design and optimal control problems are among the fundamental, ubiquitous tasks we face in science and engineering. In both cases, we aim to represent and optimize an unknown (black-box) function that associates a performance/outcome to a set of controllable variables through an experiment. In cases where the experimental dynamics can be described by partial differential equations (PDEs), such problems can be mathematically translated into PDE-constrained optimization tasks, which quickly become intractable as the number of control variables and the cost of experiments increases. In this work we leverage physics-informed deep operator networks (DeepONets) – a self-supervised framework for learning the solution operator of parametric PDEs – to build fast and differentiable surrogates for rapidly solving PDE-constrained optimization problems, even in the absence of any paired input-output training data. The effectiveness of the proposed framework will be demonstrated across different applications involving continuous functions as control or design variables, including time-dependent optimal control of heat transfer, and drag minimization of obstacles in Stokes flow. In all cases, we observe that DeepONets can minimize infinite dimensional cost functionals in a matter of seconds, yielding a significant speed up compared to traditional PDE-constrained optimization approaches that employ adjoint solvers to optimize over finite-dimensional state-spaces.

Sifan Wang, Paris Perdikaris
University of Pennsylvania
sifanw@sas.upenn.edu, pgp@seas.upenn.edu

MS44

Ice Sheet Uncertainty Quantification Via Product-

Convolution Hessian Approximation

Quantifying the uncertainty in the basal sliding friction field is critical for understanding antarctic ice sheet dynamics. Existing methods for quantifying this uncertainty rely on low rank approximation of the Hessian in the inverse problem, but these methods are sub-optimal because the rank of the Hessian is large. We present a new product-convolution method for approximating the Hessian, and show that our method can approximate the ice sheet Hessian cheaply, despite the high rank.

Nick Alger
The University of Texas at Austin
Center for Computational Geosciences and Optimization
nalger225@gmail.com

Tucker Hartland, Noemi Petra
University of California, Merced
thartland@ucmerced.edu, npetra@ucmerced.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS44

Estimating Parameters of the Nonlinear Cloud and Rain Equation from a Large-Eddy Simulation

Predatorprey dynamics have been suggested as simplified models of stratocumulus clouds, with rain acting as a predator of the clouds. We describe a mathematical and computational framework for estimating the parameters of a simplified model from a large eddy simulation (LES). In our method, we extract cycles of cloud growth and decay from the LES and then search for parameters of the simplified model that lead to similar cycles. We implement our method via Markov chain Monte Carlo. Required error models are constructed based on variations of the LES cloud cycles. This computational framework allows us to test the robustness of our overall approach and various assumptions, which is essential for the simplified model to be useful. Our main conclusion is that it is indeed possible to calibrate a predatorprey model so that it becomes a reliable, robust, but simplified representation of selected aspects of a LES. In the future, such models may then be used as a quantitative tool for investigating important questions in cloud microphysics.

Matthias Morzfeld
University of California, San Diego
Scripps Institution of Oceanography
mmorzfeld@ucsd.edu

Spencer C. Lunderman
University of Arizona
Department of Mathematics
Lunderman@math.arizona.edu

Franziska Glassmeier
TU Delft
f.glassmeier@tudelft.nl

Graham Feingold
National Oceanic and Atmospheric Administration
graham.feingold@noaa.gov

MS44

Fast Methods for Bayesian Inverse Problems with Uncertain Pde Forward Models with Application to Ice Sheet Flow Inverse Problems

We consider the Bayesian inference of the uncertain basal sliding coefficient field in the presence of uncertainty in the thermal properties of a nonlinear Stokes ice sheet model. To account for the associated model uncertainty (error), we employ the Bayesian Approximation Error (BAE) approach to approximately premarginalize simultaneously over both the noise in measurements and uncertainty in the forward model. To reduce the computational cost, we linearize the model discrepancy, which leads to an explicit formula for the likelihood and use this approximation as a control variate in combination with BAE. We carry out approximative posterior uncertainty quantification based on a linearization of the parameter-to-observable map centered at the maximum a posteriori (MAP) basal sliding coefficient estimate, i.e., by taking the Laplace approximation. We study the performance of this hybrid BAE approach in the context of an ice sheet inverse problem, where the basal sliding coefficient field is the parameter of primary interest, which we seek to infer, and the thermal properties of the ice (that enter into the rheology) represent so-called nuisance (secondary uncertain) parameters. The results demonstrate both the need to account for model uncertainty and the computational tractability of doing so.

Noemi Petra
University of California, Merced
npetra@ucmerced.edu

Ruanui Nicholson
Department of Engineering Science
University of Auckland
ruanui.nicholson@auckland.ac.nz

Umberto Villa
Electrical and Systems Engineering
Washington University in St Louis
uvilla@wustl.edu

Radoslav G. Vuchkov
University of California, Merced
rvuchkov@ucmerced.edu

MS44

Using Uncertainty to Improve Predictability

One can improve predictions by exchanging interpretability for predictability. Many machine learning techniques are designed this way. We will show that uncertainties can, in some cases, be used to improve predictions. The domain of application is time dependent processes.

Juan M. Restrepo
Oak Ridge National Lab
restrepojm@ornl.gov

Dallas Foster
Massachusetts Institute of Technology

fostdall@mit.edu

MS45

Efficient Hyper-Reduced Data-Driven Nonlinear Manifold Reduced Order Model

Traditional linear subspace reduced order models (LS-ROMs) are able to accelerate physical simulations, in which the intrinsic solution space falls into a subspace with a small dimension, i.e., the solution space has a small Kolmogorov n-width. However, for physical phenomena not of this type, such as advection-dominated flow phenomena, a low-dimensional linear subspace poorly approximates the solution. To address cases such as these, we have developed an efficient nonlinear manifold ROM (NM-ROM), which can better approximate high-fidelity model solutions with a smaller latent space dimension than the LS-ROMs. Our method takes advantage of the existing numerical methods that are used to solve the corresponding full order models (FOMs). The efficiency is achieved by developing a hyper-reduction technique in the context of the NM-ROM. Numerical results show that neural networks can learn a more efficient latent space representation on advection-dominated data from 2D Burgers equations with a high Reynolds number. A speed-up of up to 11.7 for 2D Burgers equations is achieved with an appropriate treatment of the nonlinear terms through a hyper-reduction technique.

Youngsoo Choi
Lawrence Livermore National Laboratory
choi15@llnl.gov

MS45

Stochastic approximation for nonlocal operators

In this talk, we will consider a new stochastic approximation framework for nonlocal operators via physics-informed deep neural networks. We will demonstrate the efficiency of the proposed method for solving forward and inverse integro equations such as nonlocal Possion models.

Ling Guo
Shanghai Normal University
lguo@shnu.edu.cn

MS45

Manifold Learning for Forward and Inverse UQ in High Dimensions

High-fidelity physics-based models often generate very high-dimensional quantities of interest (QoIs) (e.g., spatio-temporally varying material deformation or fluid flow). For such models, the construction of a surrogate model becomes computationally intractable. To address this issue, we introduce a manifold learning-based method for the construction of surrogate models for complex stochastic systems. Our first objective is to identify the embedding of a set of high-dimensional data representing quantities of interest of the stochastic model. We employ Grassmannian diffusion maps (GDMaps), a two-step non-linear dimension reduction technique which allows the reduction of data dimensionality and the identification of meaningful independent directions of the manifold on which the data live. Non-intrusive polynomial chaos expansion is employed to map stochastic input parameters and diffusion coordinates of the reduced space by adaptively constructing the polynomial base onto the non-linear manifold. Adaptive clustering in conjunction with geometric harmonics, is employed to map newly generated realizations of latent features to the ambient space which allows for the generation of highly accurate out-of-sample predictions. We illustrate the proposed framework and perform forward UQ on a system of advection-diffusion-reaction equations modeling a first-order chemical reaction between two species. An inverse UQ problem is also investigated on a 2D steady-state heat equation.

Katiana Kontolati
Johns Hopkins University
kontolati@jhu.edu

Dimitrios Loukrezis
Institut Theorie Elektromagnetischer Felder, TU Darmstadt
Graduate School Computational Engineering, TU Darmstadt
loukrezis@temf.tu-darmstadt.de

Ketson dos Santos
Ecole Polytechnique Fédérale de Lausanne
kmd2191@columbia.edu

Dimitrios Giovanis, Michael D. Shields
Johns Hopkins University
dgiovan1@jhu.edu, michael.shields@jhu.edu

MS45

Learning the G-Limits in Homogenization Problems Via Physics Informed Neural Network

Multiscale equations with a scale separation can be approximated by the corresponding homogenized equations with slowly varying homogenized coefficients (the G-limit). The traditional homogenization techniques typically rely on the periodicity of the multiscale coefficients, thus finding the G-limits requires some other approaches in more general settings. We consider the inverse problem of recovering the G-limits from the (noisy) measurement of multiscale solutions. In this work, we develop an efficient physics-informed neural networks (PINNs) algorithm for the inverse problem when the multiscale coefficients are not necessarily periodic or with a known form. We demonstrate that our approach could produce desirable approximations to the G-limits and, consequently, homogenized solutions.

Jun Sur Richard Park
University of Iowa
junsur-park@uiowa.edu

Xueyu Zhu
University of Iowa
Department of Mathematics
xueyu-zhu@uiowa.edu

MS46

Entropy-Based Adaptive Design for Contour Finding and Estimating Reliability

In reliability analysis, methods used to estimate failure probability are often limited by the costs associated with model evaluations. Many of these methods, such as multi-fidelity importance sampling (MFIS), rely upon a computationally efficient, surrogate model like a Gaussian process (GP) to quickly generate predictions. The quality of the

GP fit, particularly in the vicinity of the failure region(s), is instrumental in supplying accurately predicted failures for such strategies. We introduce an entropy-based GP adaptive design that, when paired with MFIS, provides more accurate failure probability estimates and with higher confidence. We show that our greedy data acquisition strategy better identifies multiple failure regions compared to existing contour-finding schemes. We then extend the method to batch selection, without sacrificing accuracy. Illustrative examples are provided on benchmark data as well as an application to an impact damage simulator for National Aeronautics and Space Administration (NASA) spacesuits.

Robert Gramacy
Virginia Tech
rbg@vt.edu

MS46

Scalable Bayesian Transport Maps for Computer-Model Emulation and Data Assimilation

Output ensembles from computer models can be modeled using multivariate probabilistic distributions. A multivariate distribution can be described by a triangular transport map from the target distribution to a simple reference distribution. We propose Bayesian nonparametric inference on the transport map by modeling its components using Gaussian processes. This enables regularization and accounting for uncertainty in the map estimation, while still resulting in a closed-form invertible posterior map. We then focus on inferring the distribution of a spatial field from a small output ensemble. We develop specific transport-map priors that are highly flexible but shrink toward a Gaussian field with Matern-type covariance. The approach is scalable to high-dimensional fields due to data-dependent sparsity and parallel computations. We discuss extensions, including Dirichlet process mixtures for marginal non-Gaussianity. We present numerical results to demonstrate the accuracy, scalability, and usefulness of our methods, including statistical emulation of non-Gaussian climate-model output. This idea can also be used to infer forecast distributions in ensemble-based data assimilation.

Matthias Katzfuss
Texas A&M University
katzfuss@gmail.com

Florian Schäfer
Georgia Tech
florian.schaefer@cc.gatech.edu

MS46

Scaling Up Bayesian Uncertainty Quantification for Inverse Problems Using Deep Neural Networks

Due to the importance of uncertainty quantification (UQ), Bayesian approach to inverse problems has recently gained popularity in applied mathematics, physics, and engineering. However, traditional Bayesian inference methods based on Markov Chain Monte Carlo (MCMC) tend to be computationally intensive and inefficient for such high dimensional problems. To address this issue, a surrogate based method, calibration-emulation-sampling (CES), has recently been proposed for large dimensional UQ problems. In this work, we propose a novel CES approach for Bayesian inference based on deep neural network models for the emulation phase. The resulting algorithm is computationally more efficient and more robust against variations in the training set. Further, by using an autoencoder (AE) for dimension reduction, we have been able to speed up our Bayesian inference method up to three orders of magnitude. Overall, our method, henceforth called Dimension-Reduced Emulative Autoencoder Monte Carlo (DREAMC) algorithm, is able to scale Bayesian UQ up to thousands of dimensions for inverse problems. Using two low-dimensional (linear and nonlinear) inverse problems we illustrate the validity of this approach. Next, we apply our method to two high-dimensional numerical examples (elliptic and advection-diffusion) to demonstrate its computational advantages over existing algorithms.

Shiwei Lan
California Institute of Technology
slan@caltech.edu

MS46

Large Scale Kriging by Substituting Optimization for Inversion

Gaussian processes are popular tools for prediction that have been shown to be orders of magnitude more accurate that modern competitors on a host of prediction tasks. However, the computational cost of fitting them can be daunting. Inspired by the recent deployments of ultra-large scale optimization in deep learning, this talk illustrates how carefully written optimization problems can be used to replace the usual matrix decomposition used to fit Gaussian process predictors.

Matthew Plumlee
Northwestern University
mplumlee@northwestern.edu

MS47

PyApprox: Approximation and Probabilistic Analysis of Data

PyApprox is a Python software package that provides flexible and efficient tools for credible data-informed decision making. PyApprox implements methods addressing various issues surrounding high-dimensional parameter spaces and limited evaluations of expensive simulation models with the goal of facilitating simulation-aided knowledge discovery, prediction and design. This talk will provide an overview of the various methods available in PyApprox for surrogate modeling, sensitivity analysis, uncertainty quantification, optimal experimental design and risk assessment. Focus will be given to automated tools that do not require extensive hand-tuning of hyper-parameters. The methods overview will be complemented by various numerical examples that demonstrate the plug an play nature of PyApprox which is designed to allow rapid prototyping and comparison of existing methods.

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

MS47

Lagun: An Open Source Platform for Data Exploration and Optimization

Guided by needs of real applications, the open source R/Shiny platform, Lagun, has been developed to provide a user-friendly interface to methods dedicated to the exploration and the analysis of expensive experiments (typically

complex numerical simulations or costly experimentations, which lead to datasets with a small sample size). Some guided workflows are provided to help non-expert users to apply safely the proposed methodologies. Laguns main current functionalities are (i) Optimized design of experiments to generate a dataset with a good spatial repartition of the simulations/experiments of the system in the inputs/parameters domain; (ii) Visual exploration tools to explore the complete dataset composed of inputs and system outputs/responses for insightful visual analyses, identification of the main trends and of the most influential inputs; (iii) Build surrogate models that infer a predictive relationship between inputs and outputs and use them extensively in uncertainty quantification, sensitivity analysis, deterministic optimization, optimization under uncertainty or in more intensive graphical studies; (iv) Direct connection with simulation scripts to perform automatic and sequential simulations, in particular for surrogate and simulation-based optimizations and calibration of numerical simulation. Lagun aims to integrate state-of-the-art methods via the Safran-IFPEN collaborative project and via contributions of research developers in the UQ community.

Sébastien Da Veiga
Safran Tech, FRANCE
sebastien.da-veiga@safrangroup.com

Clément Bénard
Safran Tech
clement.benard@safrangroup.com

David Chazalviel
Tech'Advantage
chazalviel@tech-advantage.com

Ayoub El Bachir
Safran Engineering Services
ayoub.el-bachiri@safrangroup.com

Delphine Sinoquet
IFP Energies nouvelles
delphine.sinoquet@ifpen.fr

MS47

Uqlab 2.0 and Uqcloud: Open-Source Vs. Cloud-Based Uncertainty Quantification Tools

Developing scientific software that can reach an heterogeneous, rapidly evolving international audience is a challenging process, albeit rewarding. By the end of 2022, UQLab will have reached 10 years of development, during which it went from being the "new kid on the block", to one of the most recognizable household names for uncertainty quantification software worldwide. This milestone will bring a significant shift in the UQLab development and distribution model, with two main changes aimed at reaching an ever larger audience. At first, the next major release 2.0 of UQLab (Q1 2022) will see the full release of its source code, including that of its "core". The use of the very permissive "BST 3-clause" open source license is expected to result in higher accessibility (and likely adoption rate) for both academic and industrial/commercial users. In parallel Q1 2022 will also see the official launch of UQCloud, the cloud-based sibling of UQLab developed from - and continuously integrated with - the same codebase. Through its secure REST API, UQCloud provides the complete array of uncertainty quantification tools included in UQLab, without any programming language restrictions. In this contribution we present an overview of the timeline and implications of this important transition, as well as our view on how modern, industry-driven agile and continuous integration practices can positively affect scientific software development.

Bruno Sudret
ETH Zurich
sudret@ibk.baug.ethz.ch

Stefano Marelli
Chair of Risk, Safety and Uncertainty Quantification
Institute of Structural Engineering, ETH Zurich
marelli@ibk.baug.ethz.ch

Christos Lataniotis
ETH Zurich
lataniotis@ibk.baug.ethz.ch

MS47

UQpy Version 4: Refactored, Continuously Integrated and Docker Ready for An Enhanced User and Developer Experience

UQpy software is an open-source python library that enables a wide range of uncertainty quantification studies on physical and mathematical systems. The software itself is split into nine different modules, with each one addressing a separate aspect of uncertainty quantification, ranging from sampling methods to surrogate modeling of complex physical systems and state-of-the-art dimension reduction techniques. The use of the Python programming language, with its simplified syntax and readability, allows scientists to focus on an expedited research process rather than complicated language constructs. The role of UQpy software is two-fold. On the one hand, it serves as a user-ready toolbox for uncertainty quantification, that includes niche uncertainty quantification methods. On the other hand, its modularity allows for quick extensibility, rendering it an ideal platform to test and implement new uncertainty quantification techniques. Version 4 of UQpy introduces an enhanced object-oriented approach to the code architecture, thus allowing for a simplified user and developer experience. Continuous Integration of the software ensures that the releases follow strict code quality rules and automated testing verifies the validity of the scientific output. Finally, packaging and distribution of UQpy in system agnostic format and multiple packaging archives such as PyPI, conda-forge and Docker images allows the dissemination and use of UQpy by lifting systems specific limitations.

Dimitrios Tsapetis, Dimitrios Giovanis, Michael D. Shields
Johns Hopkins University
dtsapet1@jhu.edu, dgiovan1@jhu.edu, michael.shields@jhu.edu

MS48

Multilevel Monte Carlo Methods for Estimating the Expected Value of Information

Motivated by applications to medical decision making under uncertainty, we study Monte Carlo estimation of the expected value of partial perfect information (EVPPI) and the expected value of sample information (EVSI). Both EVPPI and EVSI are defined as nested expectations, for which the standard (nested) Monte Carlo methods require $O(\varepsilon^{-3})$ or $O(\varepsilon^{-4})$ computational costs to achieve the root-mean-square accuracy ε. To reduce these costs

to $O(\varepsilon^{-2})$, we introduce antithetic multilevel Monte Carlo (MLMC) estimators for these quantities in this talk. Under some assumptions on decision models, the antithetic property of the MLMC estimator enables to prove such a computational complexity for estimating EVPPI [M. Giles and T. Goda, Decision-making under uncertainty: using MLMC for efficient estimation of EVPPI, Stat. Comput., 29 (2019), 739–751]. The result can be extended to EVSI, by directly using the Bayes' formula and showing auxiliary results on the MLMC estimation of nested ratio expectations [T. Hironaka, M. Giles, T. Goda and H. Thom, Multilevel Monte Carlo estimation of the expected value of sample information, SIAM/ASA J. Uncertainty Quantif., 8 (2020), 1236–1259]. Numerical experiments support our theoretical analysis.

Takashi Goda
University of Tokyo
goda@frcer.t.u-tokyo.ac.jp

Michael B. Giles
Mathematical Institute
Oxford University
Mike.Giles@maths.ox.ac.uk

Tomohiko Hironaka
University of Tokyo
hironaka-tomohiko@g.ecc.u-tokyo.ac.jp

Howard Thom
Bristol Medical School
University of Bristol
howard.thom@bristol.ac.uk

MS48

Multilevel Markov Chain Monte Carlo Methods for Large Scale Bayesian Inverse Problems

We present a novel class of Multi-Level Markov Chain Monte Carlo (ML-MCMC) algorithms and apply them in the context of large-scale Bayesian inverse problems (BIPs), where the likelihood function involves a complex differential model, which is then approximated on a sequence of increasingly accurate discretizations. The key point of this algorithm is to construct highly coupled Markov chains together with the standard Multi-level Monte Carlo argument to obtain a better cost-tolerance complexity than a single-level MCMC algorithm. We present several approaches to generate these highly coupled chains by sampling, e.g., from a maximal coupling of the proposals for each marginal Markov chain. By doing this, we are allowed to create novel ML-MCMC methods for which, contrary to previously used models, the proposals at each iteration can depend on the current state of this chain, while at the same time, creating chains that are highly correlated. The presented methods are tested on an array of both academic and large-scale BIPs, where a clear computational advantage can be observed with respect to their single-level counterpart.

Juan Pablo Madrigal Cianci
École polytechnique fédérale de Lausanne
EPFL
juan.madrigalcianci@epfl.ch

Fabio Nobile
EPFL, Switzerland
fabio.nobile@epfl.ch

Anamika Pandey
RWTH Aachen
pandey@uq.rwth-aachen.de

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

MS48

Multilevel Quasi-Monte Carlo for Random Elliptic Eigenvalue Problems

Random eigenvalue problems are useful models for quantifying the uncertainty in several applications from the physical sciences and engineering, e.g., structural vibration analysis, the criticality of a nuclear reactor or photonic crystal structures. In this paper we present a simple multilevel quasi-Monte Carlo (MLQMC) method for approximating the expectation of the minimal eigenvalue of an elliptic eigenvalue problem with coefficients that are given as a series expansion of countably-many stochastic parameters. The MLQMC algorithm is based on a hierarchy of discretisations of the spatial domain and truncations of the dimension of the stochastic parameter domain. To approximate the expectations, randomly shifted lattice rules are employed. This paper is primarily dedicated to giving a rigorous analysis of the error of this algorithm. A key step in the error analysis requires bounds on the mixed derivatives of the eigenfunction with respect to both the stochastic and spatial variables simultaneously. An accompanying paper [Gilbert and Scheichl, 2021], focusses on practical extensions of the MLQMC algorithm to improve efficiency, and presents numerical results.

Robert Scheichl
University of Heidelberg
r.scheichl@uni-heidelberg.de

MS48

Adaptive Multilevel Monte Carlo for Probabilities

We consider numerical approximations of $\mathbb{P}(G \in \Omega)$ where the d-dimensional random variable G cannot be sampled directly, but there is a hierarchy of increasingly accurate approximations which can be sampled. The cost of standard Monte Carlo estimation scales poorly with accuracy in this setup since it compounds the approximation cost for a large number of samples. Multilevel Monte Carlo techniques improve the cost slightly, but are known to perform worse when sampling discontinuous observables such as the event $G \in \Omega$. We propose a general adaptive framework which is able to return the Multilevel Monte Carlo complexities seen for smooth or Lipschitz observables. Our assumptions and numerical analysis are kept general allowing the methods to be used for a wide class of problems ranging from financial risk estimation and option pricing to computing failure probabilities.

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

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

Aretha L. Teckentrup
University of Edinburgh

a.teckentrup@ed.ac.uk

MS49

Identifying Underlying Systems by Adapted Optimal Transport Maps

Different observations of a relation between inputs and outputs are often reported in terms of histograms (discretizations of the source and the target densities). Transporting these densities to each other provides insight regarding the underlying relation. In (forward) uncertainty quantification, one typically studies how the distribution of inputs to a system affects the distribution of the system responses. In this talk we focus on the identification of the system (the transport map) itself, once the input and output distributions are determined, and suggest a modification of current practice by including data from what we call an observation process. We hypothesize that there exists a smooth manifold underlying the relation; the sources and the targets are then partial observations (possibly projections) of this manifold. Knowledge of such a manifold implies knowledge of the relation, and thus of the right transport between source and target observations.

Caroline Moosmueller
University of California, San Diego
cmoosmueller@ucsd.edu

Yannis Kevrekidis
Johns Hopkins University
yannisk@jhu.edu

Felix Dietrich
Technical University of Munich
felix.dietrich@tum.de

MS49

Partitioning Data Classes Using Optimal Transport Embeddings

Problems related to detection, estimation, clustering, and classification using data emanating from physical sensors (e.g. signals and images) often present difficult mathematical challenges due to nonlinearities present when modeling complex phenomena. When the data emanates from processes related to transport phenomena, solutions based on optimal transport and other Lagrangian embeddings capable of yielding high accuracy solutions for low computational cost have emerged. These embeddings are able to fully represent signals and image and can be viewed as mathematical transforms. We describe some of their mathematical properties related to partitioning data classes, thus supporting high classification accuracies in certain signal and image processing problems. Results with simulated and real data are also shown.

Gustavo Rohde
University of Virginia
gustavo.rohde@gmail.com

MS49

Neural-Network Approaches for Likelihood-Free Inversion

In this talk, we discuss neural network approaches for solving ill-posed inverse problems in a data driven-way. Here, we will bypass the construction of a likelihood (and prior) term and aim to solve inverse problems in a likelihood-free manner. To this end, we extend our recent work on using neural networks for solving high-dimensional optimal transport problems to inverse problems. We will also draw from advances in generative modeling. Specifically, our goal is to learn a transport map between a simple latent distribution and the posterior distribution defined by input-output pairs of the forward problem.

Lars Ruthotto
Department of Mathematics and Computer Science
Emory University
lruthotto@emory.edu

MS49

Optimal Mass Transport Methods in Medicine

We will describe the uses of optimal mass transport (OMT) methods for the analysis of medical data. This will include the new matrix-valued and vector-valued extensions. Both genomic and image data will be described. Because of their weak continuity, metrics based on OMT are ideal for handling the noise typical in medical data.

Allen Tannenbaum
Stony Brook University
Allen.Tannenbaum@stonybrook.edu

MS50

Parametric Model Embedding: A Novel Approach to Design-Space Dimensionality Reduction in Shape Optimization

The need for increasingly performing functional-surface designs is constantly growing in many engineering fields, requiring increasingly accurate analyses and innovative solutions. The latter can be achieved via the simulation-driven design optimization paradigm, which integrates shape parameterization models, numerical solvers, and optimization algorithms. The demand for highly innovative designs often requires global optimization on ever-larger design spaces, with an ever-increasing number of design variables, leading unavoidably to the so-called curse of dimensionality (CoD). Methodologies for reducing the design-space dimensionality have been recently developed based on unsupervised learning methods (e.g., principal component analysis). These methods provide reduced dimensionality representations capable of maintaining a certain degree of design variability. Although such methodologies have demonstrated their capability to alleviate the CoD, they usually do not allow to use the original design-space parameterization. This aspect represents a limit to their widespread use in the industrial field, where the design parameters often pertain to well-established parametric models (e.g., CAD models). This work discusses how to embed the parametric-model original parameters in a reduced-dimensionality representation. The proposed method uses a generalized feature space and aims at resolving a prescribed design variability by properly selecting the latent dimensionality.

Andrea Serani
Consiglio Nazionale delle Ricerche
Istituto di Ingegneria del Mare (CNR-INM)
andrea.serani@cnr.it

Matteo Diez
CNR-INM
National Research Council-Institute of Marine
Engineering

matteo.diez@cnr.it

MS50

Spline Chaos: An Efficient Representation of Stochastic Processes

In this talk, the 2018 presented spline chaos is demonstrated, where the space of orthogonal polynomials in the Wiener-Askey chaos expansion is replaced by a spline space. B-spline basis functions are used in the chaos expansion, which are i.a. prevalent in isogeometeric analysis. Strong and weak convergence of the B-spline chaos is proven and substantiated by numerical results. Further, several stochastic differential equations are numerically solved by a stochastic Galerkin type approach.

Christoph Eckert, Michael Beer
Leibniz University Hannover
eckert@irz.uni-hannover.de, beer@irz.uni-hannover.de

Pol D. Spanos
Rice University
spanos@rice.edu

MS50

Stochastic Isogeometric Analysis on Arbitrary Multipatch Domains by Spline Dimensional Decomposition

This paper presents a new stochastic method by integrating spline dimensional decomposition (SDD) of a high-dimensional random function and isogeometric analysis (IGA) on arbitrary multipatch geometries to solve stochastic boundary-value problems from linear elasticity. The method, referred to as SDD-mIGA, involves (1) analysis-suitable T-splines with significant approximating power for geometrical modeling, random field discretization, and stress analysis; (2) Bezier extraction operator for isogeometric mesh refinement; and (3) a novel Fourier-like expansion of a high-dimensional output function in terms of measure-consistent orthonormalized splines. The proposed method can handle arbitrary multipatch domains in IGA and uses standard least-squares regression to efficiently estimate the SDD expansion coefficients for uncertainty quantification applications. Analytical formulae have been derived to calculate the second-moment properties of an SDD-mIGA approximation for a general output random variable of interest. Numerical results, including those obtained for a 54-dimensional, industrial-scale problem, demonstrate that a low-order SDD-mIGA is capable of efficiently delivering accurate probabilistic solutions when compared with the benchmark results from crude Monte Carlo simulation.

Ramin Jahanbin
Front End Analytics
raminj@feasol.com

Sharif Rahman
The University of Iowa
sharif-rahman@uiowa.edu

MS50

Recent Advances on Iga-Based Multi-Index Stochastic Collocation

Multi-Index Stochastic Collocation (MISC) is a method of the multi-level family, aimed at reducing computational costs when repeatedly solving a parametric PDE for Uncertainty Quantification purposes. This cost reduction is achieved by exploiting multiple hierarchies of discretizations; in particular, anisotropic grids are considered. Moreover, the random variables are sampled in a deterministic way, by using tensor grids instead of Monte Carlo samples. In this talk, we employ Isogeometric Analysis (IGA) instead of the more traditional finite elements solvers. IGA solvers employ splines as basis functions, which enables a simpler meshing process, exact geometry representation and high-continuity basis functions. Finally, IGA solvers fit particularly well in the MISC framework due to their tensor-based construction. The effectiveness of the methodology is showcased by a few numerical examples. This talk will also serve as introduction to the Minisymposium, discussing other uses of spline-based technologies for UQ.

Lorenzo Tamellini
Istituto di Matematica Applicata e Tecnologie Informatiche
tamellini@imati.cnr.it

Chiara Piazzola
Consiglio Nazionale delle Ricerche - Istituto di Matematica Applicata e Tecnologie Informatiche (CNR - IMATI)
chiara.piazzola@imati.cnr.it

Joakim Beck
King Abdullah University of Science and Technology (KAUST)
joakim.beck@kaust.edu.sa

Raúl Tempone
Chair of Mathematics for Uncertainty Quantification
RWTH-Aachen University, Germany
tempone@uq.rwth-aachen.de

MS52

Squishing the Banana: Transport Map-Accelerated Adaptive Importance Sampling

Sampling from probability distributions with complex structure can be very challenging, as the ubiquitous Metropolis-Hastings method can exhibit poor mixing. In many applications this complexity can manifest itself in the target distribution being concentrated on a lower dimensional manifold. Without using a method that exploits this structure, proposals will often be made off the manifold and rejected. Transport maps have recently been used as a tool to effectively simplify this complex structure and accelerate Metropolis-Hastings algorithms in such a scenario. In this talk we will see how the same approach can accelerate and stabilize ensemble importance sampling schemes, a family of methods which have favourable properties, including the potential to be very efficient on parallel architectures.

Simon Cotter
University of Manchester
School of Mathematics
Simon.cotter@manchester.ac.uk

Ioannis Kevrekidis
Dept. of Chemical Engineering
Princeton University
yannis@princeton.edu

Paul Russell

University of Manchester
Department of Mathematics
paul.russell@postgrad.manchester.ac.uk

MS52

Deep Tensor Train Approximation for Rare Event Simulation

Estimation of event probability is a famous statistical task. Headline applications include risk assessment of a waste repository or an infection surge. In most problems neither the random variable defining the event nor its distribution are available explicitly. Instead, those are parametrized by a high-dimensional random vector, and the sought probability needs to be computed as a multivariate integral of a complicated and numerically expensive function (e.g. a PDE solution) over the failure region. The standard unbiased Monte Carlo approach requires the number of samples inversely proportional to both the squared error and the event probability, which can exceed billions if the event is rare. We develop a method to approximate a biasing distribution by building upon the deep composition of Rosenblatt transport maps, induced by a set of probability densities bridging from a tractable reference measure to the optimal biasing distribution. Each map is computed using a Tensor Train approximation of the pullback of the bridging density by the previously computed map. The bridging is done by tightening the width of the sigmoid approximation of the indicator function of the event. Moreover, we can estimate a non-tractable (e.g. posterior) density using the same approach but bridging with tempered densities. Numerical experiments with Bayesian inverse ODE and PDE problems demonstrate little to no scaling of the computing complexity with the failure probability.

Sergey Dolgov
University of Bath
Department of Mathematical Sciences
s.dolgov@bath.ac.uk

Tiangang Cui
Monash University
tiangang.cui@monash.edu

Robert Scheichl
University of Heidelberg
r.scheichl@uni-heidelberg.de

MS52

A Scalable and Efficient Bayesian Inference Framework for a Nonlinear Stokes Ice Sheet Flow Model

Solving Bayesian inverse problems suffers from the twin difficulties of the high dimensionality of the uncertain parameters and (possibly) computationally expensive forward models. In this talk, we consider the problem of reducing the state and parameter dimension for Bayesian inverse problems. To reduce the parameter dimension, we exploit the underlying problem structure (e.g., local sensitivity of the data to parameters, the smoothing properties of the forward model, and the covariance structure of the prior) and identify a likelihood-informed parameter subspace that shows where the change from prior to posterior is most significant. For the state dimension reduction (i.e., to establish a low-dimensional manifold for a nonlinear forward problem), we employ a proper orthogonal decomposition (POD) combined with the discrete empirical interpolation method (DEIM). The resulting joint parameter and state dimension reduction leads to a scalable and efficient scheme that can be used to explore the posterior distribution of the parameter or subsequent predictions. We demonstrate the accuracy and efficiency of this approach for an inverse problem governed by the nonlinear Stokes ice sheet flow model, where the parameter of interest is the basal sliding coefficient field.

Ki-Tae Kim
University of California Merced
kkim107@ucmerced.edu

Noemi Petra
University of California, Merced
npetra@ucmerced.edu

Benjamin Peherstorfer
Courant Institute of Mathematical Sciences
New York University
pehersto@cims.nyu.edu

Tiangang Cui
Monash University
tiangang.cui@monash.edu

MS52

Solving High-Dimensional Nonlinear Filtering Problems Using a Tensor Train Decomposition Method

In this talk, we propose an efficient numerical method to solve high-dimensional nonlinear filtering (NLF) problems. Specifically, we use the tensor train decomposition method to solve the forward Kolmogorov equation (FKE) arising from the NLF problem. Our method consists of offline and online stages. In the offline stage, we use the finite difference method to discretize the partial differential operators involved in the FKE and extract low-dimensional structures in the solution tensor using the tensor train decomposition method. In the online stage using the precomputed low-rank approximation tensors, we can quickly solve the FKE given new observation data. Therefore, we can solve the NLF problem in a real-time manner. Under some mild assumptions, we provide convergence analysis for the proposed method. Finally, we present numerical results to show the efficiency and accuracy of the proposed method in solving up to six-dimensional NLF problems.

Zhongjian Wang
Department of Statistics and the College
University of Chicago
zhongjian@uchicago.edu

sijing li
Department of Mathematics
The University of Hong Kong
lsj17@hku.hk

Stephen S. -T. Yau
Department of Mathematical Sciences
Tsinghua University
yau@uic.edu

Zhiwen Zhang
Department of Mathematics,
The University of Hong Kong,

zhangzw@hku.hk

MS53

Optimality Conditions and Regularization for OUU with Almost Sure State Constraints

In this talk, we present necessary and sufficient optimality conditions for convex stochastic optimization problems subject to almost sure equality and conical constraints. We refine classical results by Rockafellar and Wets from two-stage stochastic optimization to include states belonging to the Bochner space of essentially bounded random variables with images in a reflexive and separable Banach space. Under certain conditions, the optimality conditions given are necessary and sufficient. Lagrange multipliers exhibit different regularity depending on whether or not the assumption of relatively complete recourse is satisfied. We propose a Moreau–Yosida regularization for such problems and show consistency of the optimality conditions for the regularization problem as the regularization parameter is taken to infinity. Algorithmic approaches using stochastic approximation are discussed and an application to PDE-constrained optimization under uncertainty is presented.

Caroline Geiersbach
Weierstrass Institute for Applied Analysis and Stochastics Berlin
geiersbach@wias-berlin.de

MS53

An Augmented Lagrangian Approach for Risk-Averse PDE-Constrained Optimization with State Constraints

Many science and engineering applications necessitate the optimization of systems described by PDEs with uncertain inputs including noisy problem data and unknown boundary or initial conditions. One can formulate these problems as risk-averse optimization problems in Banach space, which upon discretization, become enormous nonsmooth nonlinear programs. The analysis and numerical solution of these problems is further complicated when pointwise constraints on the PDE solution are present. The Lagrange multipliers for these constraints are often measures, leading to mesh dependence if naively discretized and solved with off-the-shelf optimization methods. To address this challenge, we present a general method for risk-averse, state-constrained PDE optimization motivated by the method of multipliers. At each iteration, our algorithm solves smooth equality-constrained subproblems using a composite-step SQP method that exploits inexact linear system solves. We prove convergence of our algorithm in infinite dimensions and demonstrate its efficiency on numerical examples.

Drew P. Kouri
Optimization and Uncertainty Quantification
Sandia National Laboratories
dpkouri@sandia.gov

MS53

Optimal Design Under Uncertainty of Chemoepitaxial Guideposts for the Directed Self-Assembly of Block Copolymer Systems

Directed self-assembly (DSA) of block-copolymers (BCPs) is a promising strategy for the manufacturing of nanoscale devices. In the particular form of this process known as chemoepitaxy, BCP mixtures which naturally form nanoscale structures upon phase separation are guided by chemically patterned substrates to produce morpologies of interest. We aim to study the problem of designing optimal substrate chemical patterns given desired equilibrium morphologies. The chemoepitaxial process can be modeled using the nonlocal Cahn-Hilliard equations arising from the minimization of the Ohta-Kawasaki free-energy, with additional terms accounting for the substrate interactions. However, the specific equilibrium state produced given a fixed substrate design depends on the initial condition used for the free-energy minimization. We thus pose the optimal design problem as one under uncertainty and look to quantify the reliability of substrate designs subject to uncertainty in the initial condition. In this talk, we present a solution strategy for the optimal design problem and numerically demonstrate its effectiveness in producing robust substrate designs.

Dingcheng Luo
Oden Institute for Computational Engineering and Sciences,
The University of Texas at Austin, Austin, TX, USA
dc.luo@utexas.edu

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS53

A Framework for Machine Learning Based Optimization Under Uncertainty

We propose in this talk a general framework for machine learning based optimization under uncertainty and inverse problems. Our approach replaces the complex forward model by a surrogate, e.g. a neural network, which is learned simultaneously in a one-shot sense when estimating the unknown parameters from data or solving the optimal control problem. By establishing a link to the Bayesian approach, an algorithmic framework is developed which ensures the feasibility of the parameter estimate / control w.r. to the forward model.

Claudia Schillings
University of Mannheim
Institute of Mathematics
c.schillings@uni-mannheim.de

Philipp A. Guth, Simon Weissmann
University of Mannheim
pguth@mail.uni-mannheim.de, simon.weissmann@uni-heidelberg.de

MS55

Local Approximation of Expected Utility Surface for Nonlinear Bayesian Optimal Experimental Design

Data acquisition is time-consuming, expensive, and even dangerous. Furthermore, the data may not constrain the parameters-of-interest. We present a rigorous optimal experimental design (OED) workflow that identifies efficient experiments by maximizing the information gain while

minimizing costs and risks. Under a Bayesian framework, the collected data are used to update a prior distribution. This defines a posterior distribution that illustrates the uncertainty reduction. The amount of information gain is quantified by the Kullback-Leibler (KL) divergence from the posterior distribution to the prior. However, numerical evaluation of the KL divergence requires nested Monte Carlo, which is intractable because of the expensive physics-based forward models that must be evaluated to compute the posterior density function. Therefore, we replace the posterior density with a computationally cheaper local polynomial surrogate model. However, the surrogate model introduces bias into the estimate of the KL divergence. Repeatedly refining the surrogate model ensures asymptotic decay of the surrogate bias, and using a bias-variance trade-off to trigger refinements yields a rate-optimal strategy. Local polynomials are appealing because our refinement strategy can minimize error in regions of non-trivial posterior support. We demonstrate our method by optimizing the location of sensors that observe the concentration of a contaminate, inferring the conductivity field in a physical domain.

Andrew D. Davis
MIT
add8536@cims.nyu.edu

Xun Huan
University of Michigan
xhuan@umich.edu

MS55

Multilevel Double Loop Monte Carlo Method with Importance Sampling for Bayesian Optimal Experimental Design

An optimal experimental set-up maximizes the value of data for statistical inferences. The efficiency of strategies for finding optimal experimental set-ups is particularly important for experiments that are time-consuming or expensive to perform. When the experiments are modeled by Partial Differential Equations (PDEs), multilevel methods have been proven to reduce the computational complexity of their single-level counterparts when estimating expected values. For a setting where PDEs can model experiments, we propose a multilevel method for estimating the widespread criterion known as the Expected Information Gain (EIG) in Bayesian optimal experimental design. We propose a Multilevel Double Loop Monte Carlo (MLDLMC), where the Laplace approximation is used for importance sampling in the inner expectation. The method's efficiency is demonstrated by estimating EIG for inference of the fiber orientation in composite laminate materials from an electrical impedance tomography experiment.

Luis Espath
RWTH Aachen University
espath@uq.rwth-aachen.de

Joakim Beck
King Abdullah University of Science and Technology (KAUST)
joakim.beck@kaust.edu.sa

Ben Mansour Dia
CPG, King Fahd University of Petroleum and Minerals
diabenmansour@gmail.com

Raul F. Tempone
Mathematics, Computational Sciences & Engineering
King Abdullah University of Science and Technology
raul.tempone@kaust.edu.sa

MS55

Deep Neural Network Surrogates for PDE-Constrained Bayesian OED

We consider optimal experimental design for Bayesian inverse problems governed by PDE forward models. Specially, we seek to find a sensor network geometry that maximizes the expected information gain (EIG). The problem is intractable using brute force Monte Carlo sampling. To make the EIG computation tractable, we invoke a projected neural network surrogate approximation of the parameter-to-observable (p2o) map. The architecture of the projected network exploits the geometry and intrinsic low-dimensionality of the p2o map, resulting in a parsimonious representation that requires few training data. Initial results are very promising.

Keyi Wu
University of Texas Austin
keyiwu@math.utexas.edu

Thomas O'Leary-Roseberry
The University of Texas at Austin
Oden Institute for Computational and Engineering Sciences
tom@oden.utexas.edu

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS55

Bayesian Design for Parameter Estimation Using Supervised Regression Methods

We propose a computationally efficient simulation-based approach to find Bayesian optimal designs when the goal of inference is parameter estimation. We apply supervised regression methods to simulated data in order to facilitate the computation of the expected loss criteria used in Bayesian optimal design. In particular, we consider squared error loss, for which our approach is straightforward to implement. This approach requires considerably fewer simulations from the candidate models than previous approaches using approximate Bayesian computation. The approach is particularly useful in the presence of models with intractable likelihoods but can also provide computational advantages when the likelihoods are manageable. The methods are applied to find optimal designs for a diverse range of models such as Markov process models or nonlinear mixed effects models.

Markus Hainy
Johannes Kepler University Linz
markus.hainy@jku.at

Chris Drovandi
Queensland University of Technology

Australia
c.drovandi@qut.edu.au

David Price
University of Melbourne
david.price1@unimelb.edu.au

MS56

Bayesian Improved Cross Entropy Method for Categorical Distribution Model

We propose a modification of the improved cross entropy iCE method to enhance its performance for network reliability assessment. The iCE method employs a smooth transition from the nominal density to the optimal importance sampling IS density and updates the parametric distribution model through cross entropy minimization $Papaioannouet.al.$, 2019. The efficiency and the accuracy of the iCE method are largely influenced by the choice of the parametric model. In the context of reliability of systems with multi-state component states, the obvious choice of the parametric model is the Categorical distribution. When updating such distribution model through standard iCE, the probability assigned to certain category often converges to 0 due to lack of occurrence of the samples, resulting in a reduced sample space for the final importance sampling distribution. However, such category is not necessarily negligible for the failure event. To circumvent this problem, we propose an accurate yet efficient algorithm termed Bayesian improved cross entropy method $BiCE$, in which the posterior predictive distribution is employed to update the parametric family instead of the weighted maximum likelihood estimation. A set of examples are used to illustrate the efficiency and the accuracy of the proposed method.

Jianpeng Chan
Technical University of Munich
Engineering Risk Analysis Group
jianpeng.chan@tum.de

Iason Papaioannou
Engineering Risk Analysis Group
TU M{ü}nchen
iason.papaioannou@tum.de

Daniel Straub
Engineering Risk Analysis Group
TU München
straub@tum.de

MS56

Physically Driven EV-GDEE for the Estimation of Time-Variant Failure Probabilities of High-Dimensional Stochastic Dynamical Systems

Time-variant reliability assessment of engineering systems subjected to stochastic excitations, especially for rare events, is of paramount importance for the performance-based decision-making of design, but is still of great challenge due to the nonlinear and random coupling in high-dimensional systems. For this purpose, an ensemble-evolving-based generalized density evolution equation $EV-GDEE$ is established, as a one- or two-dimensional partial differential equation, with respect to the response of interest in a high-dimensional system. The equivalent drift coefficient in the EV-GDEE represents the physically driving force of evolution of the probability density function PDF in the ensemble sense, and is identified as the conditional expectation of the original drift function. In this sense, the proposed method can be called as the physically driven EV-GDEE. Some representative dynamical analyses of the underlying physical system are performed for the identification of the equivalent drift coefficient. Then, the EV-GDEE can be solved to capture the transient PDF of the response of interest. For the purpose of reliability, an absorbing boundary process is constructed. Its EV-GDEE can be established and solved to obtain time-variant first-passage reliability. The proposed method shows the high accuracy of the failure probability even in the order of magnitude 10^{-4} - 10^{-6} for rare events, which are achieved with only hundreds of dynamical analyses.

Meng-Ze Lyu, Jianbing Chen
College of Civil Engineering, Tongji University
chenjb@tongji.edu.cn, chenjb@tongji.edu.cn

MS56

Sensitivity Analysis Methodology for Extreme Financial Risks Using Splitting Methods based on Reversible Transformations

We develop a new methodology to compute sensitivities of financial risk statistics based on rare-event using the idea of splitting. Most popular financial risk quantities are computed as expectation of a functional of the relevant risk-governing stochastic model when the underlying stochastic object takes value in a very small probability set. We represent the derivative of these quantities also as an expectation by using an integration- by-parts formula in the setting of Malliavin calculus. We derive explicit representation of the Malliavin weight which allows us to compute the derivative as an expectation of another functional of the stochastic model when the underlying stochastic object takes value in a very small probability set. Using the idea of splitting, we show that the derivative can be computed efficiently using a recently proposed simulation method based on reversible shaking transformations We demonstrate the application of our methodology using an example of option Greeks for an out-of-the-money European option.

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

Ankush Aggarwal
University of Glasgow
James Watt School of Engineering
ankush.aggarwal@glasgow.ac.uk

Stefano De Marco
Ecole Polytechnique
stefano.de-marco@polytechnique.edu

Gang Liu
Squarepoint Capital
gang.liu@squarepoint-capital.com

MS57

Stationary Density Estimation of Ito Diffusion with Deep Learning

In this talk, I will discuss a deep learning approach for estimating the stationary density of ergodic It diffusions. Given the time series of the stochastic process, we employ neural network models to solve supervised learning tasks to

uncover the drift and diffusion coefficients. Subsequently, we estimate the density by finding a neural network-based regression solution to the stationary Fokker-Planck equations with the estimated drift and diffusion coefficients. Convergence study, accounting generalization errors induced by estimation of the drift and diffusion coefficients, and the PDE solvers, will be presented. I will demonstrate numerical results on a 20-dimensional Langevin dynamical system.

John Harlim
Pennsylvania State University
jharlim@psu.edu

MS57

Transition Path Theory with Neural and Tensor Networks

Deep neural-network/ tensor method can be used for compressing high-dimensional functions arising from partial differential equations (PDE). In this talk, we focus on using these methods for solving for the committor function, which enables us to study the transition processes between metastable states in chemistry applications.

Yuehaw Khoo
Department of Statistics
University of Chicago
ykhoo@galton.uchicago.edu

Yian Chen, Jeremy Hoskins
University of Chicago
yianc@uchicago.edu, jeremyhoskins@uchicago.edu

Michael Lindsey
New York University
michael.lindsey@cims.nyu.edu

Jianfeng Lu
Mathematics Department
Duke University
jianfeng@math.duke.edu

Lexing Ying
Stanford University
Department of Mathematics
lexing@stanford.edu

MS57

Computing Committor Functions for the Study of Rare Events Using Deep Learning

The committor function is a central object in understanding transition events between metastable states in complex systems. It has a simple mathematical description it satisfies the backward Kolmogorov equation. However, computing the committor function for realistic systems at low temperature is a challenging task, due to the curse of dimensionality and the scarcity of transition data. In this talk, I will present a computational approach that overcomes these issues and achieves good performance on complex benchmark problems with rough energy landscapes. The new approach combines deep learning, importance sampling and feature engineering techniques. This establishes an alternative practical method for studying rare transition events among metastable states of complex, high dimensional systems.

Qianxiao Li
National University of Singapore
Singapore
qianxiao@nus.edu.sg

Bo Lin, Weiqing Ren
National University of Singapore
matboln@nus.edu.sg, matrw@nus.edu.sg

MS57

Learning Control Forces to Characterize Dynamical Phase Transitions

Sampling the collective, dynamical fluctuations that lead to nonequilibrium pattern formation requires probing rare regions of trajectory space. Recent approaches to this problem based on importance sampling, cloning, and spectral approximations, have yielded significant insight into nonequilibrium systems, but tend to scale poorly with the size of the system, especially near dynamical phase transitions. Here we propose a machine learning algorithm that samples rare trajectories and estimates the associated large deviation functions using a many-body control force by leveraging the flexible function representation provided by deep neural networks, importance sampling in trajectory space, and stochastic optimal control theory. We show that this approach scales to hundreds of interacting particles and remains robust at dynamical phase transitions.

Grant Rotskoff
Stanford University
rotskoff@stanford.edu

MS58

Uncertainties in Data-Drive Earth System Predictions

Deep learning is a tool that is increasingly being leveraged by the earth system modeling community due its capacity to ingest large quantities of data to discover patterns and relationships, incorporate limited observed data, and make enhanced, faster data-driven earth-system predictions. When using earth-system models directly to make predictions, the various sources of uncertainties can be quantified, e.g. the distinction between internal variability and model uncertainty. However, when we instead make enhanced predictions with data-driven models, a further source of uncertainty is introduced by the data-driven system, and we also must ensure that the other sources of uncertainty are properly accounted for and quantified when making a prediction with such models. This becomes especially important in the case of impacts assessments where decisions are routinely required to be made under uncertainty. In this talk, I will present our recent efforts to quantify the various sources of uncertainty in our predictions of climate quantities from data-driven models. This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344 and is supported by LLNL Laboratory Directed Research and Development project 22-SI-008.

Gemma J. Anderson, Baoxiang Pan, Andre Goncalves, Don Lucas, Celine Bonfils, Jiwoo Lee
Lawrence Livermore National Laboratory
anderson276@llnl.gov, pan11@llnl.gov, andreric@llnl.gov,

ddlucas@llnl.gov, bonfils2@llnl.gov, lee1043@llnl.gov

MS58

ML4UQ: Machine Learning to Enable Efficient Uncertainty Quantification in Climate Models

We will talk how to leverage machine learning (ML) to advance uncertainty quantification (UQ) in climate models. First, we will discuss different ML-based surrogate modeling techniques to accelerate UQ. These techniques include building a surrogate model of the climate model to learn the complex input-output relationship, and construction of a surrogate model of the likelihood function in performing UQ. Additionally, we will introduce an assumption-free, scalable, and robust prediction interval method to quantify the ML model uncertainty so as to produce accurate and credible surrogate model predictions. This prediction interval method can not only quantify the uncertainty of the climate model surrogates but also the uncertainty of the data-driven ML models in simulating the climate system. We will demonstrate the methods using land surface models.

Dan Lu, Dan Ricciuto
Oak Ridge National Laboratory
lud1@ornl.gov, ricciutodm@ornl.gov

Siyan Liu
Oak Ridge National Laboratory, US
lius1@ornl.gov

MS58

Global Sensitivity Analysis Using the Ultra-Low Resolution Energy Exascale Earth System Model

For decades, the Arctic has been warming at least twice as fast as the rest of the globe. As a first step towards quantifying parametric uncertainty in Arctic feedbacks, we perform a variance-based global sensitivity analysis (GSA) using a fully-coupled, ultra-low resolution (ULR) configuration of version 1 of the Department of Energys Energy Exascale Earth System Model (E3SMv1). The study randomly draws 139 realizations of ten model parameters spanning three E3SMv1 components, which are used to generate 75-year long projections of future climate using a fixed pre-industrial forcing. We quantify the sensitivity of six Arctic-focused quantities of interest (QOIs) to these parameters using main effect, total effect and Sobol sensitivity indices computed with a Gaussian process (GP) emulator. A sensitivity index-based ranking of model parameters shows that the atmospheric parameters in the CLUBB (Cloud Layers Unified by Binormals) scheme have significant impact on sea ice status and the larger Arctic climate. We also use our GP emulator to predict the response of varying each variable when the impact of other parameters are averaged out. These results allow one to assess the non-linearity of a parameters impact on a QOI and investigate the presence of local minima encountered during the spin-up tuning process. Our study confirms the necessity of performing global analyses involving fully-coupled climate models, and motivates follow-on investigations involving the ULR model.

Irina K. Tezaur
Sandia National Laboratories
ikalash@sandia.gov

Kara Peterson
Sandia Natl. Labs
kjpeter@sandia.gov

Amy Powell
Sandia National Laboratories
ajpowel@sandia.gov

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

Erika Roesler
Sandia National Laboratories
elroesl@sandia.gov

MS58

E3sm Atmosphere Surrogate Construction and Calibration Using Machine Learning and Reduced Order Modeling Methods

We develop surrogate construction methods for spatially varying scalar fields from the E3SM Atmosphere Model (EAM). The goal is to integrate surrogates into climate model development. This differs from approaches where surrogate construction lags model development and tuning, and where surrogate targets are typically single scalars. Surrogate models for spatial fields could improve expert climate model tuning and provide field data to automated tuning algorithms. There is a tradeoff between improving the surrogate (i.e. accuracy, dimensionality, and number of uncertain climate model parameters) and the expense of generating the sample E3SM simulations. We study the effects of sample size, simulation length, simulation resolution, and surrogate dimensionality to determine the most efficient surrogate methods for a desired surrogate accuracy. Surrogates are constructed using reduced order modeling (ROM) where principal component analysis (PCA) is applied to the ensemble target fields. The method is a two-stage process in which we perform PCA, and then use machine learning methods to fit the reduced or latent space, including multivariate polynomial regression, random forests, support vector machines, kernel ridge regression, Gaussian processes, k-nearest neighbor, and feed forward neural networks. Finally, surrogate models are used to drive a Bayesian calibration of the uncertain climate model parameters. SAND 2021-12647A.

Benjamin Wagman, Kamaljit Chowdhary
Sandia National Laboratories
bmwagma@sandia.gov, kchowdh@sandia.gov

MS59

Accelerating Physical Simulations with Reduced Order Models

Although many model reduction schemes have been developed to reduce the computational cost of simulations while minimizing the error introduced in the reduction process, there are challenges especially in nonlinear advection-dominated problems such as sharp gradients, moving shock fronts, and turbulence, which hinder those model reduction schemes from being practical. In this talk, we will present recent developments in reduced order models for such simulations, with which we will demonstrate both good accuracy and considerable speed-up, enabling faster simulation and optimization solvers.

Dylan M. Copeland
TBA2
copeland11@llnl.gov

Siu Wun Cheung, Youngsoo Choi, Kevin Huynh
Lawrence Livermore National Laboratory
cheung26@llnl.gov, choi15@llnl.gov, huynh24@llnl.gov

MS59

Three Approaches to Handle Uncertainty in Sequential Decision-Making

In many practical problems from online advertisement to healthcare and computational finance, it is extremely important to have guarantees on the performance of the policy generated by our algorithms. This reduces the risk of deploying our strategy and helps us to convince the product (hospital, investment) manager that it is safe and not going to harm the business. In this talk, we discuss three different approaches to this fundamental problem, we call them model-based, model-free, and online. In the model-based approach, we first use the batch of data and build a simulator that mimics the behavior of the dynamical system under studies (online advertisement, hospitals ER, financial market), and then use this simulator to generate data and learn a policy. The main challenge here is to have guarantees on the performance of the learned policy, given the error in the simulator. This line of research is closely related to the area of robust learning and control. In the model-free approach, we learn a policy directly from the batch of data (without building a simulator), and the main question is whether the learned policy is guaranteed to perform at least as well as a baseline strategy. This line of research is related to off-policy evaluation and control. In the online approach, the goal is to control the exploration of the algorithm in a way that never during its execution the loss of using it instead of the baseline strategy is more than a given margin.

Mohammad Ghavamzadeh
Google Research, US
ghavamza@google.com

MS59

Fast Approximate Bayesian Uncertainty Quantification Using Conditional Generative Adversarial Networks

Uncertainty quantification (UQ) is crucial in engineering and science, but also computationally intensive when coupled with an expensive physics-based model. Conventional inverse UQ methods (e.g. Markov Chain Monte Carlo (MCMC) algorithms and variational inference) require restarting the inference procedure when encountering a new measurement or a new experiment condition. Therefore, they become prohibitive for tasks where repeated Bayesian inference is needed, such as optimal experimental design. We thus take a different approach that avoids online physics-based model simulation or optimization and aim to directly build a single generative model that integrates conditional distributions of the parameter conditioned under different possible measurements and experiment designs. In particular, we employ the conditional generative adversarial network (cGAN), which learns the mapping from the joint space of latent variable, observable and design variable to the space of posterior distributions, by solving a minimax game between two deep neural networks. Only one cGAN needs to be trained offline, and different observation and design values can be passed into it during online usage to achieve fast approximate posterior sampling. We illustrate the convergence of our method and demonstrate it on a problem of optimal experimental design of combustion reaction application.

Chengyang Huang, Xun Huan
University of Michigan
chengyah@umich.edu, xhuan@umich.edu

MS59

Bayesian Approach for Inverse Problems in Tomographic Imaging

Nonlinear and infinite-dimensional inverse problems appear in many applications including tomographic imaging. Solving such inverse problems using Bayesian approach needs appropriate function space setting and a measure theoretic framework, as well as sampling methods to approximate the posterior density, which is computationally challenging due to the expensive PDE-based forward simulations. Therefore, it is essential to reduce the uncertainty of the model parameters to be estimated as well as to acquire the most informative data to update the prior knowledge about the uncertain quantities. We will present information-based optimal experimental design of a tomography inverse problem with application in medical imaging with the goal of optimizing the solution of the associated Bayesian inverse problem.

Leila Taghizadeh
Vienna University of Technology
leila.taghizadeh@ma.tum.de

Ahmad Karimi, Clemens Heitzinger
TU Wien
ahmad.karimi@tuwien.ac.at,
clemens.heitzinger@tuwien.ac.at

MS60

Physics-Informed Machine Learning for Uncertainty Sensitivity Analysis

This work considers global sensitivity analysis (GSA) for situations where both a physics-based model and experimental observations are available, and investigates physics-informed machine learning (PIML) strategies to effectively combine the two sources of information in order to maximize the accuracy of sensitivity estimates. When computational models (either physics-based or data-driven) are used for the sensitivity analysis of engineering systems, the sensitivity estimate is affected by the accuracy and uncertainty of the model. Two representative machine learning (ML) techniques are considered, namely, deep neural networks (DNN) and Gaussian process (GP) modeling, and two strategies for incorporating physics knowledge within these techniques are investigated, namely: (i) incorporating loss functions in the ML models to enforce physics constraints, and (ii) pre-training and updating the ML model using simulation and experimental data respectively. Four different models are built for each type (DNN and GP), and the uncertainties in these models are included in the Sobol indices computation. The DNN-based models, with many degrees of freedom in terms of model parameters and training options, are found to result in smaller bounds on the sensitivity estimates when compared to the GP-based models. The proposed methods are illustrated for additive manufacturing and lake temperature modeling examples.

Berkcan Kapusuzoglu, Sankaran Mahadevan
Vanderbilt University
berkcan.kapusuzoglu@vanderbilt.edu,

sankaran.mahadevan@vanderbilt.edu

MS60

Quantifying Uncertainty in Predictions of Physics Constrained Neural Networks

Data sparsity is a common issue to train machine learning models such as neural networks for engineering and scientific applications, where experiments and simulations are expensive. Recently physics-constrained neural networks (PCNNs) were developed to improve the training efficiency and interpretability of data-driven models by incorporating physical knowledge as constraints to guide the training process. However, prediction error is still a challenge for the PCNN framework. The model-form uncertainty of the PCNN model comes from the subjective choice of neural network structures, including the architecture, number of hidden units and layers, activation functions, etc. The parameter uncertainty of the PCNN model arises from the imperfect training process as well as imprecise physical knowledge used in the PCNN. In this work, the quantification of the parameter uncertainty in the PCNN model is studied. Based on the training errors of different physical constraints, the overall error of PCNN predictions is estimated. A heat transfer example is used to demonstrate the potential of the proposed approach for uncertainty quantification in PCNNs.

Luka Malashkhia
Georgia Institute of Technology
lukamalashkhia@gatech.edu

Dehao Liu
Texas A&M University
dehao.liu@tamu.edu

Yan Wang
Georgia Institute of Technology
yan-wang@gatech.edu

MS60

Quantifying Uncertainty in Machine Learning Models for Time Series Classification

Time series classification (TSC) is a fundamental challenge in many Scientific Machine Learning (SciML) problems essential to national security, including stockpile surveillance analysis. With the increase of time series data availability, many new classes of TSC algorithms have been proposed. In this talk, we present an empirical study of recent Machine and Deep Learning (ML/DL) methods and architectures for TSC with uncertainty quantification considerations and evaluate them on a number of univariate benchmark datasets. This study aims at enhancing the overall credibility of ML/DL predictions and opening new doors for SciML models to be credibly deployed to costly and high stakes problems.

Ahmad A. Rushdi
Stanford University
rushdi@stanford.edu

Erin Acquesta, Jose Huerta, Kyle Neal, William J. Rider
Sandia National Laboratories
eacques@sandia.gov, jghuert@sandia.gov, kneal@sandia.gov, wjrider@sandia.gov

MS60

Machine Learning for Single Particle Motion in Plasmas

The application of machine learning to particle advancement in kinematic plasma simulations is the motivation behind this work. Fully kinetic particle-in-cell simulations are necessary for resolution of high-frequency effects in magnetic confinement fusion plasmas near the wall, but the application of this simulation methodology to larger domains of the plasma is limited by computational cost and speed limitations. Both gyrokinetic and fully kinetic particle-in-cell methods often employ the Boris-Bunemann particle advancement algorithm due to its excellent long-simulation-time accuracy and efficiency. This work discusses the implementation of machine learning in plasma physics for the purpose of regression and classification tasks relating to experimental and computational data. The implementation of a neural network to the task of particle advancement is addressed, where training data is provided by the Boris-Bunemann algorithm with and without frequency correction. The aim of this work is to implement a neural network to accurately predict single charged particle motion in an electromagnetic field, capturing relevant physical effects and potentially providing a computationally efficient alternative to the Boris-Bunemann particle advancement algorithm for use in kinematic plasma simulations.

Sonata Valaitis
University of Illinois Urbana-Champaign
sonatav2@illinois.edu

Kathryn Maupin
Sandia National Laboratories
kmaupin@sandia.gov

Davide Curreli
University of Illinois at Urbana-Champaign
dcurreli@illinois.edu

MS61

DIRT: Tensorised Rosenblatt Transport for High-Dimensional Stochastic Computation

Characterising intractable high-dimensional random variables is one of the fundamental challenges in stochastic computation. It has broad applications in statistical physics, machine learning, uncertainty quantification, econometrics, and beyond. The recent surge of transport maps offers a mathematical foundation and new insights for tackling this challenge. In this talk, we will present a functional tensor-train (TT) based order-preserving construction of inverse Rosenblatt transport in high dimensions. It characterises intractable random variables via couplings with tractable reference random variables. By integrating our TT-based approach into a nested approximation framework inspired by deep neural networks, we are able to significantly expand its capability to random variables with complicated nonlinear interactions and concentrated density functions. We demonstrate the efficacy of the resulting deep inverse Rosenblatt transport (DIRT) on a range of applications in statistical learning and uncertainty quantification, including parameter estimation for dynamical systems, PDE-constrained inverse problems, and Bayesian

filtering.

Tiangang Cui
Monash University
tiangang.cui@monash.edu

MS61

Transport Information Flows for Bayesian Sampling Problems

In AI and inverse problems, the Markov chain Monte Carlo (MCMC) method is a classical model-free method for sampling target distributions. A fact is that the optimal transport first-order method (gradient flow) forms the MCMC scheme, known as Langevin dynamics. A natural question arises: Can we propose accelerated or high order optimization techniques for MCMC methods? We positively answer this question by applying optimization methods from optimal transport and information geometry, known as transport information geometry. E.g., we introduce a theoretical framework for Newton's flows in probability space w.r.t. the optimal transport metric. Several numerical examples are given to demonstrate the effectiveness of the proposed optimization-based sampling methods.

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

MS61

The Stein Geometry in Machine Learning: Gradient Flows, Optimal Transport and Large Deviations

Sampling or approximating high-dimensional probability distributions is a key challenge in computational statistics and machine learning. This talk will present connections to gradient flow PDEs, optimal transport and interacting particle systems, focusing on the recently introduced Stein variational gradient descent methodology and some variations. The construction induces a novel geometrical structure on the set of probability distributions related to a positive definite kernel function. We discuss the corresponding geodesic equations, infinitesimal optimal transport maps, as well as large deviation functionals. This is joint work with A. Duncan (Imperial College London), L. Szpruch (University of Edinburgh) and M. Renger (Weierstrass Institute Berlin).

Nikolas Nüsken
Universität Potsdam
nuesken@uni-potsdam.de

MS62

Gaussian Process Surrogate Modeling for Expensive Experiments with Multiple Mesh Fidelities

With breakthroughs in scientific computing, computer simulations are quickly replacing physical experiments in modern scientific and engineering problems. Such simulations, however, are very complex and time-consuming, and predictive models are used to emulate the expensive computer code. In many problems, these simulations are often performed in multiple stages, with the accuracy (or fidelities) at each stage controlled by a tuning parameter. This provides a flexible multi-stage, multi-fidelity framework for efficiently simulating lower-fidelity training data. We propose a new Multi-stage Multi-Fidelity Gaussian process (M^2GP) model, which leverages this multi-stage, multi-fidelity simulation data to efficiently train a probabilistic emulator for the high-fidelity, expensive computer code. We provide recommendations of two non-stationary kernel specifications which work well in different scenarios. The improvement of M^2GP over existing methods is then demonstrated in numerical experiments, an application to emulation of cantilever beam deflection under stress, and an application to emulation of heavy-ion collisions, which shed light on the origins of the Universe shortly after the Big Bang.

Yi Ji
Duke University
yi.ji@duke.edu

Henry Shaowu Yuchi
Georgia Institute of Technology
shaowu.yuchi@gatech.edu

Simon Mak, Derek Soeder, Jean-Francois Paquet, Steffen A Bass
Duke University
sm769@duke.edu, derek.soeder@duke.edu, jeanfrancois.paquet@duke.edu, bass@duke.edu

V. Roshan Joseph, C. F. Jeff Wu
Georgia Institute of Technology
roshan@gatech.edu, jeff.wu@isye.gatech.edu

MS62

Hybrid Multilevel Monte Carlo Polynomial Chaos Method for Global Sensitivity Analysis

Sensitivity Analysis supports the development of mathematical models by studying how the variability of a models output can be assigned to the different sources of uncertainty (and their interactions). Unfortunately, in the context of high-fidelity computer simulations, Global Sensitivity Analysis (GSA) often requires thousands of model evaluations, which practically make its computational cost intractable. In this contribution, we propose to alleviate the GSA cost by employing a multifidelity (MF) approach. MF approaches fuse information from several sources, e.g. models with a varying discretization, and are, in general, more efficient than their single fidelity counterparts. In particular, we show how Multilevel Monte Carlo (MLMC) can be used to accelerate the GSA workflow, by serving as a tool for the non-intrusive spectral projection step, which requires a quadrature in a potentially large space, in the construction of a polynomial chaos expansion (PCE). Subsequently, main and total effect Sobol indices may be computed analytically from the PCE coefficients, enabling efficient GSA. The sample allocation of our MLMC-PCE method is derived as an optimal solution targeted at the GSA statistics of interest, e.g. first order or total Sobol indices. We will present several numerical results based on verification problems and we will discuss the comparison of our GSA method with other MF GSA approaches and their single-fidelity counterpart.

Michael B. Merritt
North Carolina State University
mbmerrit@ncsu.edu

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

Teresa Portone
Sandia National Laboratories
tporton@sandia.gov

Michael S. Eldred
Sandia National Laboratories
Optimization and Uncertainty Quantification Dept.
mseldre@sandia.gov

Pierre Gremaud
Department of Mathematics
North Carolina State University
gremaud@ncsu.edu

MS62

Optimal Experimental Design for MFNets: Applications to Bayesian Learning of Multi-Fidelity Data-Driven Surrogates

This work examines optimal experimental design procedures for a Bayesian multi-fidelity uncertainty quantification framework, called MFNets. The framework is used to construct a surrogate from different information sources of varying cost and accuracy. The main benefits of MFNets include its ability to utilize non-hierarchical training data and the flexibility of the underlying directed acyclic graph which MFNets uses to encode connections between information sources. To select the most informative designs for the surrogate construction from a set of candidates, we consider several design criteria, such as maximizing a posterior variance or minimizing a posterior integrated variance. We demonstrate the experimental design for MFNets in synthetic and physics-based examples, and also compare the MFNets-based results with other state-of-the-art approaches, e.g., the classical co-Kriging model.

Trung Pham, Alex Gorodetsky
University of Michigan
trungp@umich.edu, goroda@umich.edu

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

MS62

Building Multiscale Representations and Surrogate Using a Greedy Approach

Multiscale strategies have been effectively used for physics and data driven modeling tasks in several communities such as PDE solvers, signal processing and statistical learning. The current talk introduces a novel weighted multiscale kernel based statistical surrogate, that greedily learns the structure of the underlying physics using a sparsity constraint. This presentation will provide a general analysis of the approximation properties for the proposed method, guidelines for setting up algorithmic hyperparameters, and experimental results demonstrating stability and robustness of predictions and approximations with respect to widely used surrogate models such as Gaussian Processes. The theoretical results presented for the proposed method will be further justified by surrogate construction for rain induced debris flow simulations (shallow water type system of equations). Additionally, we will cover details on how such a model can be used for uncertainty quantification and propagation in the modeling pipeline.

Prashant Shekhar
Embry Riddle Aeronautical University
shekharp@erau.edu

Abani Patra
Tufts University
abani.patra@tufts.edu

MS63

Higher-Order Quasi-Monte Carlo Rules for Domain Uncertainty Quantification Using Periodic Random Variables

Domain uncertainty quantification is the study of how uncertainties in the shape of a domain affect the output of a computational model. In this talk, we discuss a parameterization for an uncertain domain using a random perturbation field in which a countable number of independent random variables enter the random field as periodic functions. The random field can be constructed to have a prescribed mean and covariance function. As an application, we design tailored quasi-Monte Carlo cubature rules that achieve dimensionally independent, higher-order convergence rates for high-dimensional numerical integration problems associated with the output of an elliptic PDE model problem subject to uncertainty in the domain shape. In practice, the series expansion for the random perturbation field needs to be truncated to a finite number of terms. To this end, we also analyze the resulting dimension truncation error as well as the finite element discretization error of the PDE model problem. The theoretical rates are assessed in an ensemble of numerical experiments.

Harri H. Hakula
Aalto University
Harri.Hakula@aalto.fi

Helmut Harbrecht
Universitaet Basel
Department of Mathematics and Computer Science
helmut.harbrecht@unibas.ch

Vesa Kaarnioja
Aalto University
Department of Mathematics and Systems Analysis
vesa.kaarnioja@iki.fi

Frances Y. Kuo
School of Mathematics and Statistics
University of New South Wales
f.kuo@unsw.edu.au

Ian H. Sloan
University of New South Wales
School of Mathematics and Stat
i.sloan@unsw.edu.au

MS63

Various Search Algorithms for Good Lattice Rules

Lattice rules are among the most prominently studied quasi-Monte Carlo methods to approximate multivariate integrals. A rank-1 lattice rule to approximate a d-dimensional integral is fully specified by its generating vector and its number of points. While there are many results on the existence of "good' rank-1 lattice rules, there are no explicit constructions for good generating vectors for dimensions $d \geq 3$. This is why one usually resorts to computer search algorithms. Motivated by earlier work of Korobov from 1963 and 1982, we present two variants of

search algorithms for good lattice rules and show that the resulting rules exhibit a convergence rate in weighted function spaces that can be arbitrarily close to the optimal rate. Moreover, contrary to most other algorithms, we do not need to know the smoothness of our integrands in advance, the generating vector will still recover the convergence rate associated with the smoothness of the particular integrand, and, under appropriate conditions on the weights, the error bounds can be stated without dependence on d.

Peter Kritzer
Institut fuer Finanzmathematik
Universitaet Linz, Austria
peter.kritzer@oeaw.ac.at

Adrian Ebert
RICAM
Austrian Academy of Sciences
adrian.ebert@ricam.oeaw.ac.at

Dirk Nuyens
Department of Computer Science
KU Leuven
dirk.nuyens@cs.kuleuven.be

Onyekachi Osisiogu
RICAM
Austrian Academy of Sciences
onyekachi.osisiogu@ricam.oeaw.ac.at

MS63

Scaled Lattice Rules for Integration on $\mathbb{R}^d$ Achieving Higher-Order Convergence

In this talk, we show that by simply scaling lattice rules from the unit cube $[0,1]^d$ to properly sized boxes on $\mathbb{R}^d$, taking into account all errors, we can achieve higher-order convergence in approximating an integral on $\mathbb{R}^d$ where the order of convergence matches the smoothness of the integrand function in a certain Sobolev space of dominating mixed smoothness. Our method only assumes that we can evaluate the integrand function f and does not assume a particular density nor the ability to sample from it. In particular for the analysis we show that the method of adding Bernoulli polynomials to a function to make it "periodic' on a box without changing its integral value over the box, is equivalent to an orthogonal projection from a well chosen Sobolev space of dominating mixed smoothness to an associated periodic Sobolev space of the same dominating mixed smoothness, which we call a Korobov space. We conduct numerical experiments comparing with (i) direct product of Gauss–Hermite quadrature, (ii) Sparse grid based on Gauss–Hermite, and (iii) scaled interlaced Sobol' sequence. This talk is based on Dirk Nuyens and Yuya Suzuki, "Scaled lattice rules for integration on $\mathbb{R}^s$ achieving higher-order convergence with error analysis in terms of orthogonal projections onto periodic spaces," arXiv preprint arXiv:2108.12639, 2021.

Yuya Suzuki, Dirk Nuyens
Department of Computer Science
KU Leuven
yuya.suzuki@ntnu.no, dirk.nuyens@cs.kuleuven.be

MS64

Approximation Properties of Two-Layer Neural Networks with Values in a Banach Space

Approximation properties of infinitely wide neural networks have been studied by several authors in the last few years. New function spaces have been introduced that consist of functions that can be efficiently (i.e., with dimension-independent rates) approximated by neural networks of finite width. Typically, these functions are supposed to act between Euclidean spaces, typically with a high-dimensional input space and a lower-dimensional output space. As neural networks gain popularity in inherently infinite-dimensional settings such as inverse problems and imaging, it becomes necessary to analyse the properties of neural networks as nonlinear operators acting between infinite-dimensional spaces. In this talk, I will present dimension-independent Monte-Carlo rates for neural networks acting between Banach spaces with a partial order (vector lattices), where the ReLU no linearity will be interpreted as the lattice operation of taking the positive part.

Yury Korolev
University of Cambridge
DAMTP
y.korolev@maths.cam.ac.uk

MS64

On the Approximation Theory of Operator Learning Architectures

In this talk, I will present recent work on the approximation theory of operator learning architectures, including DeepONets and Fourier Neural Operators (FNOs). By a general decomposition of the error into encoding, approximation and reconstruction errors, both upper and lower bounds on the total error can be derived. The lower bounds represent fundamental limitations of current architectures. Potential ways to overcome such limitations by suitably modifying these architectures will be discussed. The abstract error and complexity estimates are then illustrated for prototypical examples of advection-dominated operators associated with hyperbolic PDEs. For these concrete examples, it is shown that the fundamental limitations faced by operator learning frameworks such as DeepOnets and FNOs can be overcome by our proposed neural operator architecture.

Samuel Lanthaler
ETH Zurich
samuel.lanthaler@sam.math.ethz.ch

Siddhartha Mishra
ETH Zurich
Zurich, Switzerland
siddhartha.mishra@sam.math.ethz.ch

MS64

Deep Learning and Accelerated Algorithms for High-Dimensional Hilbert-Valued Functions from Limited Data

Reconstructing high-dimensional functions taking values in abstract spaces is a key challenge arising in many mathematical modeling and computational science applications, e.g. computational uncertainty quantification (UQ). Such problems are often posed in terms of parameterized partial differential equations (PDEs) taking values in a Hilbert or Banach space. Tackling this problem is difficult due to the large expense of obtaining samples and high problem dimensionality. The last decade has seen huge advances in algorithms for such problems based on compressed sensing (CS), enabling stable and accurate resolution of Hilbert-valued functions from limited amounts of sampling data.

Simultaneously, deep neural networks (DNNs) have begun to emerge as promising tools for scientific computing. Our novel approach to this problem is fully algorithmic, combining CS techniques, orthogonal polynomial and finite element discretization, and efficient first-order optimization schemes with acceleration based on the weighted Square-Root LASSO decoder for Hilbert-valued functions. We also present DNN approaches for the Banach-valued case, including an algorithm and full theoretical analysis with explicit guarantees on the error and sample complexity. Our theoretical results for DNNs present a clear accounting of all sources of error. We present a general result for both classes of functions and conclude with the main application for a specific class of parametric PDEs.

Sebastian Moraga, Ben Adcock, Nick Dexter
Simon Fraser University
sebastian_moraga_scheuermann@sfu.ca,
ben_adcock@sfu.ca, ndexter@sfu.ca

Simone Brugiapaglia
Department of Mathematics and Statistics
Concordia University
simone.brugiapaglia@concordia.ca

MS64

Exploiting Jacobian Information in Parametric Surrogate Modeling

Outer-loop problems arising in scientific applications (such as optimization, uncertainty quantification and inverse problems) require repeated evaluation of computationally intensive numerical models, such as those arising from discretization and solution of ordinary and partial differential equations. The cost of these evaluations makes solution using the model prohibitive, and efficient accurate surrogates are a key to solving these problems in practice. In this talk we will discuss how compressed derivative information can be exploited to aid the design of parsimonious neural network architectures to be deployed in high dimensional inference. These reduced-basis architectures outperform conventional data-driven approaches when limited training data are available due to computational costs of evaluating high dimensional nonlinear PDEs.

Thomas O'Leary-Roseberry
The University of Texas at Austin
Oden Institute for Computational and Engineering Sciences
tom@oden.utexas.edu

Umberto Villa
Electrical and Systems Engineering
Washington University in St Louis
uvilla@wustl.edu

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS65

Delta-UQ: Accurate Uncertainty Quantification via Anchor Marginalization

TBD...

Jayaraman Thiagarajan
Lawrence Livermore National Labs
jayaramanthi1@llnl.gov

Rushil Anirudh
Lawrence Livermore National Laboratory
anirudh1@llnl.gov

MS65

Trustworthy Machine Learning via Logic Inference

Advances in machine learning have led to rapid and widespread deployment of learning based inference and decision making for safety-critical applications, such as autonomous driving and security diagnostics. Current machine learning systems, however, assume that training and test data follow the same, or similar, distributions, and do not consider active adversaries manipulating either distribution. Recent work has demonstrated that motivated adversaries can circumvent anomaly detection or other machine learning models at test time through evasion attacks, or can inject well-crafted malicious instances into training data to induce errors in inference time through poisoning attacks. In this talk, I will describe my recent research about security and privacy problems in machine learning systems. In particular, I will introduce several adversarial attacks in different domains, and discuss potential defensive approaches and principles, including game theoretic based and knowledge enabled robust learning paradigms, towards developing practical robust learning systems with robustness guarantees.

Bo Li
University of Illinois Urbana-Champagne
lbo@illinois.edu

MS65

Information-Preserving Bayesian Models for Efficient and Robust Learning

TBD..

Sandeep Madireddy
Argonne National Laboratory
smadireddy@anl.gov

MS65

Locally Valid and Discriminative Prediction Intervals for Deep Learning Models

Crucial for building trust in deep learning models for critical real-world applications is efficient and theoretically sound uncertainty quantification, a task that continues to be challenging. Useful uncertainty information is expected to have two key properties: It should be valid and discriminative . Moreover, when combined with deep learning (DL) methods, it should be scalable and affect the DL model performance minimally. Most existing Bayesian methods lack frequentist coverage guarantees and usually affect model performance. The few available frequentist methods are rarely discriminative and/or violate coverage guarantees due to unrealistic assumptions. Moreover, many methods are expensive or require substantial modifications to the base neural network. Building upon recent advances in conformal prediction and leveraging the classical idea of

kernel regression, we propose Locally Valid and Discriminative (LVD) prediction intervals, a simple, efficient and lightweight method to construct discriminative prediction intervals (PIs) for almost any DL model. With no assumptions on the data distribution, such PIs also offer finite-sample local coverage guarantees. We empirically verify, using diverse datasets, that besides being the only locally valid method for DL, LVD also exceeds or matches the performance of existing uncertainty quantification methods, while offering additional benefits in scalability and flexibility.

Shubhebdu Trivedi
Massachusetts Institute of Technology
shubhendu@csail.mit.edu

MS66

Necessary and Sufficient Conditions for Asymptotically Optimal Linear Prediction of Gaussian Processes on Compact Metric Spaces

Optimal linear prediction (aka. kriging) of a random field $\{Z(x)\}_{x\in\mathcal{X}}$ indexed by a compact metric space $(\mathcal{X}, d_{\mathcal{X}})$ can be obtained if the mean value function $m\colon \mathcal{X}\to\mathbb{R}$ and the covariance function $\varrho\colon \mathcal{X}\times\mathcal{X}\to\mathbb{R}$ of Z are known. We consider the problem of predicting the value of $Z(x^*)$ at some location $x^*\in\mathcal{X}$ based on observations at locations $\{x_j\}_{j=1}^n$ which accumulate at x^* as $n\to\infty$ (or, more generally, predicting $\varphi(Z)$ based on $\{\varphi_j(Z)\}_{j=1}^n$ for linear functionals $\varphi, \varphi_1, \ldots, \varphi_n$). Our main result characterizes the asymptotic performance of linear predictors (as n increases) based on an incorrect second order structure $(\widetilde{m}, \widetilde{\varrho})$, without any restrictive assumptions on $\varrho, \widetilde{\varrho}$ such as stationarity. We, for the first time, provide necessary and sufficient conditions on $(\widetilde{m}, \widetilde{\varrho})$ for asymptotic optimality of the corresponding linear predictor holding uniformly with respect to φ. These general results are illustrated by weakly stationary Gaussian fields on $\mathcal{X}\subset\mathbb{R}^d$, and for non-stationary SPDE-based generalized Whittle-Matrn fields.

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

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

MS66

Solving PDEs via Gaussian Processes

Solving PDEs is a fundamental task in scientific computing. Recently there has been a growing interest in automating the solution process by casting the task as a machine learning problem. This talk is concerned with a simple Gaussian process framework that addresses this problem with some theoretical guarantee. We will present several successful numerical examples in nonlinear elliptic PDEs, time-dependent PDEs, high dimensional PDEs, and inverse problems. We will also discuss the efficient implementation of this algorithm by generalizing ideas in fast algorithms for Gaussian processes and hierarchical parameter learning.

Yifan Chen
Caltech
yifanc@caltech.edu

MS66

Statistical Finite Element Methods for Nonlinear Partial Differential Equations

I will present a statistical finite element method for nonlinear, time-dependent phenomena, illustrated in the context of nonlinear internal waves (solitons). The statistical finite element method (statFEM) is a statistical augmentation of the finite element method that enables model-data synthesis through the admission of model misspecification inside of the governing equations, as represented by a Gaussian process. The method is Bayesian, coherently updates model mismatch upon receipt of observed data, and is applicable to a wide range of problems across science and engineering for which finite element methods are appropriate. I'll first introduce the statFEM, before detailing the method for nonlinear problems. I'll conclude the talk by discussing a case study of the Korteweg-de Vries equation for solitons, applying the method to experimental data.

Connor Duffin
University of Cambridge
cpd32@cam.ac.uk

Edward Cripps, Thomas Stemler
University of Western Australia
edward.cripps@uwa.edu.au, thomas.stemler@uwa.edu.au

Mark Girolami
Imperial College London
mag92@eng.cam.ac.uk

MS66

Balancing Numerical and Statistical Accuracy in the SPDE Approach to Gaussian Processes

The stochastic partial differential equation approach to Gaussian processes (GPs) represents Matérn GP priors in terms of n finite element basis functions and Gaussian coefficients with sparse precision matrix. Such representations enhance the scalability of GP regression and classification to datasets of large size N by setting $n\approx N$ and exploiting sparsity. In this talk we reconsider the standard choice $n\approx N$ through an analysis of the estimation performance. We demonstrate that under mild conditions one can set $n\ll N$ without hindering the estimation accuracy, leading to a second layer of computational gain.

Daniel Sanz-Alonso, Ruiyi Yang
University of Chicago
sanzalonso@uchicago.edu, yry@uchicago.edu

MS67

Reliability-Based Inversion: Stepwise Uncertainty Reduction Strategies?

In the design and analysis of computers experiments, "inversion" refers, broadly speaking, to the problem of finding the values of the inputs of a model that lead to outputs with given properties. In such problems, the object of interest is a subset of the input space, or its volume under a given measure. Robust formulations of the inversion problem, where the objective is to solve an inversion problem with respect to a given subset of the input variables in the presence of uncertainty on the others, are particularly important in many applications. We focus in this talk on a partic-

ular robust formulation, which we call "Reliability-Based Inversion" (RBI), in which the objective is to solve an inversion problem for a quantile of the output with respect to the uncertain variables of the model (or, equivalently, for the probability of exceeding a given threshold). In this setting, we use the Stepwise Uncertainty Reduction (SUR) paradigm, which has proved fruitful for simpler variants of the inversion problem, to construct efficient sequential sampling strategies for the RBI problem.

Romain Ait Abdelmalek-Lomenech
Université Paris-Saclay, CNRS, CentraleSupélec, L2S
91190, Gif-sur-Yvette, France
romain.ait@centralesupelec.fr

Julien Bect
Université Paris-Saclay, CNRS, CentraleSupélec
Laboratoire des signaux et systèmes
julien.bect@centralesupelec.fr

Emmanuel Vazquez
CentraleSupelec
emmanuel.vazquez@centralesupelec.fr

MS67

Calibration of Neutronic Models based on Bayesian Linear Regression and Sobol-Based Sensitivity Analysis

When comparisons are made between experimental measurements and predictions using a numerical model, differences due to the uncertainty of the model input variables may arise. Consequently, when the numerical discrepancy of the simulator is estimated, an erroneous bias may be obtained. Therefore, to achieve an uncertainty quantification in better agreement with reality, the calibration of the uncertain input variables is necessary. However, the high CPU cost of the simulator used prevents from a direct calibration. A widely accepted approach consists of approximating the simulator by CPU-time inexpensive emulator. The input uncertainties here come from tiny manufacturing tolerances in the assembly and core composition. Therefore, an approach based on the small perturbations regime can be considered to build a simplified physical model on which a linear model is estimated in a Bayesian framework. In addition, a sensitivity analysis based on Sobol's indices is analytically performed to assess the impact of emulator parameter uncertainty. It shows that it is possible to use the maximum a posterior of the emulator parameters in a plug-in approach without introducing additional bias in the calibration process. Built upon this approximation, a posterior distribution associated with each of the uncertain core parameters is obtained from Bayesian calibration with experimental results.

Louis Berry
French Alternative Energies and Atomic Energy Commission
CEA, DES, DER, Cadarache Center, France
louis.berry@cea.fr

Amandine Marrel
CEA Cadarache, France
amandine.marrel@cea.fr

Fabrice Gamboa
Institut de Mathématiques de Toulouse. AOC Project
University of Toulouse
fabrice.gamboa@math.univ-toulouse.fr

MS67

Adaptive Yield Optimization with Mixed Gradient Information

Uncertainties in the manufacturing process of electromagnetic components may lead to rejections due to malfunctioning of the device. The uncertain design parameters can be modelled as random variables. Then, the yield is the probability of success, i.e., the probability that a device fulfills all pre-defined performance requirements under consideration of manufacturing uncertainties. A straightforward approach for yield estimation is the Monte Carlos analysis (MC). However, since the performance requirements typically include solving partial differential equations (PDEs) numerically, MC becomes computationally prohibitive. Therefore, a surrogate model based on Gaussian process regression (GPR) is used for the underlying PDEs. Further, we modify the design of the device in order to maximize the yield, i.e., to improve the reliability of the manufacturing process. In this setting, the partial derivatives with respect to some optimization variables can be obtained easily, for others not. Thus, for the optimization we propose a modified version of Powells BOBYQA algorithm, using gradient information if available. During the optimization process, the MC sample size is increased adaptively. *This work is supported by the Excellence Initiative of the German Federal and State Governments and the Graduate School of Computational Engineering at TU Darmstadt.*

Mona Fuhrländer, Sebastian Schöps
Computational Electromagnetics Group (CEM), TU Darmstadt
Centre for Computational Engineering (CCE), TU Darmstadt
fuhrlaender@temf.tu-darmstadt.de, schoeps@gsc.tu-darmstadt.de

MS67

The ICSCREAM Methodology: Identification of Penalizing Configurations in Computer Experiments using Screening and Metamodel Applications in Thermal-Hydraulics

In the framework of risk assessment in nuclear accident analysis, best-estimate computer codes, associated to a probabilistic modeling of the uncertain input variables, are used to estimate safety margins. A first step is often to identify the critical configurations (or penalizing, in the sense of a prescribed safety margin) of several input parameters (called "scenario inputs"), under the uncertainty on the other input parameters. However, the large CPU-time cost of most of the codes involves to develop highly efficient strategies. To achieve it with a very large number of inputs and from a small-size sample of simulations, a specific and original methodology, called ICSCREAM (Identification of penalizing Configurations using SCREening And Metamodel), has been proposed. The screening of influential inputs is based on an advanced global sensitivity analysis indices, namely the Hilbert-Schmidt Independence Criterion. Then, a Gaussian process metamodel is sequentially built and used to estimate, within a Bayesian framework, the conditional probabilities of exceeding a high-level threshold, according to the scenario inputs. The efficiency of this methodology is illustrated on a high-dimensional (hundred inputs) use case simulating an accident of primary coolant loss in a pressurized water reactor. The study fo-

cuses on the peak cladding temperature (PCT) and critical configurations are defined by exceeding the 90%-quantile of PCT.

Amandine Marrel
CEA Cadarache, France
amandine.marrel@cea.fr

Bertrand Iooss, Vincent Chabridon
EDF R&D
bertrand.iooss@edf.fr, vincent.chabridon@edf.fr

MS68

Spatialised Generalized Lambda Distribution for Risk-Averse Bayesian Optimisation

We consider here the global optimisation of a stochastic black-box system, which objective function (output) given a set of decision parameters (inputs) is a random variable of unknown distribution. The standard approach is to optimize the function expectation and assume that the distribution around the mean is identical for all inputs (i.i.d. noise). In this work, we focus on the case where the shape as well as the amplitude of the distribution may change significantly within the input space. Further, we wish to design risk-averse strategies to account for the variability in the objective. Our contribution is threefold: 1- We propose a new flexible surrogate model to provide an approximation of the entire distribution of the objective at any input point. To do so, we "spatialise' the so-called generalised lambda distribution, by modelling its parameters as latent Gaussian processes. Inference is made possible by the use of variational approaches. 2- We review the stochastic dominance concepts to compare distributions and characterise a set of incomparable inputs (i.e. optimal in the Pareto sense). 3- Using sequential Monte-Carlo concepts, we propose an algorithm to estimate this set while accounting for the GP uncertainties, and define a sequential sampling strategy based on this estimation. Our proposition is illustrated on several challenging toy problems.

Victor Picheny
Secondmind
victor@secondmind.ai

MS68

Inference via Surrogate Data

Inference through data and mathematical modeling is particularly challenging for dynamical systems with noisy data, model uncertainties, and unknown mechanisms. Here, parameter and uncertainty estimation problems are typically ill-posed, meaning solutions do not exist, are not unique, or do not depend continuously on the data. Inference depends on the proper inclusion of prior knowledge. In this talk, we propose and discuss surrogate data to regularize ill-posed problems. We present various examples to illustrate the advances gained by surrogate data.

Hayden Ringer
Virginia Tech
Department of Mathematics
hringer@vt.edu

Matthias Chung
Department of Mathematics
Virginia Tech
mcchung@vt.edu

MS68

Ranking and Selection Trade-Offs in Multiobjective Simulation Optimization

In multiobjective optimization, several conflicting objectives are optimized simultaneously. The goal is to find or approximate a (discrete) set of Pareto-optimal solutions that reveal the essential trade-offs between the objectives, where optimality means that no objective can be improved without deteriorating the quality of any other objective. We consider a setting where the objectives have to be observed through stochastic simulation, which may lead to two possible errors: solutions that are actually Pareto-optimal can be wrongly considered dominated, or solutions that are truly dominated are wrongly considered Pareto-optimal. We propose a Bayesian method to reduce the number of misclassification errors when identifying the solutions with the true best expected performance. We use stochastic kriging metamodels to build reliable predictive distributions of the objectives, and exploit this information in two sampling criteria: one looks at the distance between the sample mean and the predicted mean, and the other one looks at the expected change in hypervolume. We use these criteria in a sequential sampling algorithm to decide how to allocate samples. Experimental results show that the proposed method only requires a small fraction of samples compared to the standard allocation method, and it's competitive against the state-of-the-art, with the exploitation of the correlation structure being the dominant contributor to the improvement.

Sebastian Rojas Gonzalez
Ghent University
sebastian.rojasgonzalez@ugent.be

Juergen Branke
Warwick Business School
The University of Warwick
juergen.branke@wbs.ac.uk

Inneke Van Nieuwenhuyse
Quantitative Methods
Hasslet University
inneke.vannieuwenhuyse@uhasselt.be

MS68

Learning and Deploying Active Subspaces on Black-Box Simulators

Surrogate modeling of computer experiments via local models, which induce sparsity by only considering short range interactions, can tackle huge analyses of complicated input-output relationships. However, narrowing focus to local scale means that global trends must be relearned over and over again. We first demonstrate how to use Gaussian processes to efficiently perform a global sensitivity analysis on an expensive black box simulator. We next propose a framework for incorporating information from this global sensitivity analysis into the surrogate model as an input rotation and rescaling preprocessing step. We further discuss applications to derivative free optimization. Numerical experiments on observational data and benchmark test functions provide empirical validation.

Nathan B. Wycoff
Virginia Tech, U.S.
nathw95@vt.edu

Mickael Binois
INRIA
mickael.binois@inria.fr

Robert Gramacy
Virginia Tech
rbg@vt.edu

Stefan Wild
Argonne National Laboratory
LANS, Mathematics and Computer Science
wild@anl.gov

MS69

Optimal Experiment Design for Hyperpolarized MRI Measurements

Traditional magnetic resonance imaging (MRI) plays a significant role in the diagnosis of cancer and treatment decisions; however, MRI is insufficient to estimate the aggressiveness of the disease, defined as the rate of growth of cancer cells. A new emerging MRI measurement, Hyperpolarized (HP) MRI, provides enhanced insights into the tissue's metabolism and a way to identify the aggressive and less aggressive tumor colonies. HP MRI is based on the pyruvate-to-lactate reaction, and the rate of pyruvate-to-lactate exchange informs about cells' aggressiveness; a higher rate is an indication of a high-grade tumor. In HP MRI, the intensities of magnetization of polarized pyruvate and lactate are recorded multiple scan times. These intensities are used to fit a model parameter, rate of pyruvate-to-lactate exchange, using either the least square or Bayesian method. The control or design parameters in HP MRI experiments include the time intervals between consecutive scans and the flip angles at different scans. Our goal is to formulate an optimal experimental design problem to recover model parameters from the measurements accurately. First, we will present OED results using the low-fidelity HP MRI model based on ordinary differential equations. Next, we will describe the OED formulation using the PDE-based model, where model parameter of interest is treated as a spatially varying field. Finally, we will highlight key challenges and present numerical results near end of the talk.

Prashant Jha
Oden Institute
University of Texas, Austin
pjha.sci@gmail.com

Keyi Wu
University of Texas Austin
keyiwu@math.utexas.edu

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

David Fuentes
University of Texas MD Anderson Cancer Center
dtfuentes@mdanderson.org

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

J. Tinsley Oden
The University of Texas at Austin
ICES
oden@ices.utexas.edu

MS69

Transport Map-Based Bayesian Optimal Experimental Design

The Bayesian optimal experimental design is essential in many fields of science and engineering, especially when each experiment is expensive and resources are limited. Given a prior and a design-dependent likelihood function, we would like to choose the design that maximizes the expected information gain (EIG) in the posterior. Efficient and accurate estimation of EIG therefore becomes crucial. We introduce a flexible transport-map based framework that enables fast estimation of EIG by solving only convex optimization problems. This framework is also compatible with implicit models, where one can simulate from the likelihood but the conditional probability density function of the data is unknown. Several estimators naturally appear within our framework—in particular, positively and negatively biased estimators that provide bounds for the true EIG. We explore the bias and variance of our estimators and study the optimal allocation between the training and the evaluation samples given a fixed number of samples. We then demonstrate the performance of our approach using both a linear and a nonlinear example.

Fengyi Li
Massachusetts Institute of Technology
fengyil@mit.edu

Ricardo Baptista
MIT
rsb@mit.edu

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

MS69

Exploring Bayesian Optimal Experimental Designs for High-Dimensional Combustion Systems

An optimal experimental design can reduce the overall time and expense needed to acquire valuable experimental data. This is achieved by focusing research efforts on experimental conditions that are most informative based on the analysis of a computer model. In this talk, we consider a Bayesian optimal experimental design for inferring parameters of a chemically reacting system subject to both parametric and model uncertainty. A physics-based model was developed to simulate the gas-phase reactions occurring between highly reactive intermediate species in a high-pressure photolysis reactor coupled to a mass spectrometer. This model depends on a high-dimensional input parameter representing the uncertain physics, instrument, and reaction rate parameters. We employ Bayesian optimization to search across a constrained design space, identifying the experimental condition that maximizes the expected information gain. Computational challenges in this optimization procedure and evaluating the expected information gain are discussed in the context of the high-dimensional physics model. We present our key findings and discuss strategies and trade-offs for efficiently computing optimal designs for high-dimensional models.

James Oreluk
Sandia National Laboratories

joreluk@sandia.gov

Leonid Sheps
Sandia National Labs
lsheps@sandia.gov

Habib N. Najm
Sandia National Laboratories
Livermore, CA, USA
hnnajm@sandia.gov

MS69

Projected Neural Network for Nonlinear Bayesian Optimal Experimental Design

We present a deep learning framework for optimal experimental design (OED). OED seeks to design an optimal data acquisition such that uncertainties in the parameters of a Bayesian inverse problem are minimized. The resulting optimization problem requires numerous evaluations of a suitable optimality criterion, for which inner Bayesian inverse problems must be solved. This is computationally prohibitive for large-scale forward models (e.g. discretizations of PDEs) and high-dimensional parameters. To reduce the computational cost, we develop a deep learning framework that constructs a projected neural network (PNN) to parsimoniously learn the parameter-to-observable map. We use the PNN as a surrogate to accelerate the evaluation of the optimality criterion and the OED optimization. Numerical results are presented to demonstrate the accuracy and efficiency of this framework.

Keyi Wu
University of Texas Austin
keyiwu@math.utexas.edu

Thomas O'Leary-Roseberry
The University of Texas at Austin
Oden Institute for Computational and Engineering Sciences
tom@oden.utexas.edu

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS70

Understanding the Mis-Match Between Generating and Model-Estimated Dynamics When Estimating Models with Sparse Data

Using data assimilation to estimate models from sparsely measured, oscillatory, non-stationary systems is difficult. Several solutions have been posed to mitigate this data sparsity problem, but every solution tends to induce new problems without robustly estimating data. In the context of model-based estimationdata assimilationof biomedical data from the glucose-insulin system, solutions that have been used to mitigate the data sparsity problem will be presented. These solutions induce unexpected problems, and while several of these problems will be discussed, the focus of the talk will be on a problem that has been particularly vexing: the induced mismatch between generating dynamics and inferred-model dynamics. The talk will conclude with a discussion of proposed solutions to these problems.

David Albers
University of Colorado Anschutz Medical Campus
david.albers@cuanschutz.edu

MS70

Dynamics-Aware Measure Transport for Bayesian Filtering of Chaotic Dynamical Systems

Data assimilation in high-dimensional chaotic systems is a recurring challenge across disciplines, from meteorology to aerospace engineering. Despite theoretical results on the dynamics of the Bayesian update, including connecting its stability to observability, the ergodic properties as well as the local tangent space decomposition of the underlying chaotic dynamics have not been rigorously exploited in Bayesian filtering algorithms. Here, we aim to connect the concentration properties of the filtering recursion to observations on the unstable manifold. Moreover, we exploit this connection to improve the computational feasibility of the transport-based stochastic map filter of Spantini et al 2019, which, in principle, provides a consistent approximation to the filtering distribution. Specifically, we propose computation of transport maps in which both the state and observations are projected on the unstable manifold, which is approximated using particle evolutions and the tangent linear dynamics. Such a dimension reduction technique in the fully Bayesian setting parallels the extensive development of rank reduction using the unstable subspace in algorithms based on Kalman updates, such as the extended Kalman filter.

Nisha Chandramoorthy, Youssef M. Marzouk
Massachusetts Institute of Technology
nishac@mit.edu, ymarz@mit.edu

MS70

Probabilistic Gradients for Fast Calibration of Differential Equation Models

Calibration of large-scale dynamical system models to observational or experimental data is a widespread challenge throughout applied sciences and engineering. A crucial bottleneck in state-of-the art calibration methods is the calculation of local sensitivities, which often necessitates several numerical solves of the underlying system of partial or ordinary differential equations. In this talk I present a new probabilistic approach to computing local sensitivities. The proposed method has several advantages over classical methods. Firstly, it operates within a constrained computational budget and provides a probabilistic quantification of uncertainty incurred in the sensitivities from this constraint. Secondly, information from previous sensitivity estimates can be recycled in subsequent computations, reducing the overall computational effort for iterative gradient-based calibration methods.

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

Jon Cockayne
The Alan Turing Institute

jcockayne@turing.ac.uk

MS70

Subsampling in Ensemble Kalman Inversion

Inverse problems seek to find parameters that generate a set of observed data in a mathematical model. Several methods have been proposed to solve inverse problems such are optimisation and Bayesian approaches. In this talk, we consider the Ensemble Kalman Inversion (EKI) which has been recently introduced as an efficient, gradient-free method. Based on the continuous-time Ensemble Kalman Filter, it uses an ensemble of particles and a linearisation technique to essentially estimate the posterior distribution in an underlying Bayesian inverse problem. Unfortunately, the algorithm becomes inefficient or even computationally infeasible if the considered data set is too large. A similar problem appears in large-scale optimisation with gradient descent algorithms in, e.g., machine learning. Here, randomised algorithms like stochastic gradient descent (SGD) have become popular: those use only a random subset of the data in each iteration. These are so-called subsampling techniques. Based on a recent analysis of a continuous-time representation of SGD, we propose, analyse, and apply a subsampling-technique within EKI. Indeed, we propose two different subsampling techniques: either every particle observes the same data subset or every particle observes a different data subset. We present convergence results of the method (and some variants) in the setting of linear inverse problems. Then we illustrate our results in PDE-based inverse problems and image reconstruction.

Matei Hanu
University of Mannheim
mhanu@mail.uni-mannheim.de

Jonas Latz
Heriot-Watt University
j.latz@hw.ac.uk

Claudia Schillings
University of Mannheim
Institute of Mathematics
c.schillings@uni-mannheim.de

MS71

Deep Reinforcement Learning for Stochastic Optimal Control of Complex Systems under Constraints

Life-cycle stochastic control for intervention planning in deteriorating engineering environments entails a series of complex mathematical analysis steps. At every decision step, (i) the value of noisy data has to be assessed so that inspection and monitoring decisions are guided; (ii) model and state updating has to be conducted based on efficient probabilistic inference from available measurements; (iii) state-altering actions have to be optimized based on their long-term effects and availability of resources; and (iv) uncertainty propagation has to be performed so that future risks are quantified. The above essential methodological components have been mostly approached fragmentarily by existing methods. In this talk, we discuss how recent developments in multi-agent constrained Partially Observable Markov Decision Processes and Deep Reinforcement Learning can provide a broad unified framework. Emphasis is given on original actor-critic DRL architectures recently developed by the authors, tailored to engineering environments under partial observability and various deterministic or stochastic constraints. It is shown that immense state spaces can be successfully supported, whereas the consideration of factored multi-actor architectures prevents the emergence of combinatorial action spaces. Applications to several stochastically deteriorating engineering systems exemplify the efficiency of the developed algorithms compared to existing state-of-the-art techniques.

Kostas Papakonstantinou
Pennsylvania State University
kpapakon@psu.edu

Charalampos Andriotis
Delft University of Technology
c.andriotis@tudelft.nl

MS71

On-the-Fly Reduced Order Modeling with Time Dependent Basis

Parametric uncertainty quantifications (UQ) and sensitivity analysis of multi-scale time-dependent problems are cost-prohibitive especially in cases with a large number of parameters. On the other hand, many of these high-dimensional problems have a much lower intrinsic dimensionality, that if discovered, can mitigate the curse of dimensionality. This calls for techniques that extract and exploit correlated structures directly from the partial differential equations (PDE). We present a matrix/tensor reduced-order modeling framework, in which the correlated structures are extracted directly from the multi-dimensional PDE. This bypasses the need to generate data as it is required in data-driven dimension reduction techniques. These structures are exploited by building on-the-fly reduced-order models (ROM). The correlated structures are represented by a set of time-dependent orthonormal bases and their evolution is prescribed by the physics of the problem. We present several demonstration cases including reduced-order modeling of reactive species transport equation in turbulent combustion as well as sensitivity analysis and uncertainty quantification in fluid dynamics problems.

Hessam Babaee
University of Pittsburgh
h.babaee@pitt.edu

MS71

Riemannian Manifold Hamiltonian Monte Carlo Based Subset Simulation for Reliability Analysis InNon-Gaussian Space

In this talk, we present the Riemannian Manifold Hamiltonian Monte Carlo-based subset simulation (RMHMC-SS) method. The RMHMC-SS has been developed to overcome the limitations of existing Monte Carlo methods in solving reliability problems defined in highly-curved non-Gaussian spaces. RMHMC generates an optimal trajectory for Markov chain evolutions in a Hamiltonian constructed on the Riemannian manifold. Compared to the Hamiltonian Monte Carlo-based subset simulation (HMC-SS) approach, the RMHMC-SS performs better in handling highly-curved probability distributions. The talk introduces the theory, principles, and limitations of RMHMC-SS and shows a series of benchmark examples.

Marco Broccardo
ETH Zurich
marco.broccardo@unitn.it

Weiming Chen
Earthquake Engineering Research & Test Center, Guangzhou Uni
303805224@qq.com

Ziqi Wang
University of California, Berkeley
ziqiwang@berkeley.edu

Junho Song
Seul National University
junhosong@snu.ac.kr

MS71

Deep Bayesian Spline Learning for Closed-Form Equation Discovery of Nonlinear Dynamics with Quantified Uncertainty

Nonlinear dynamics are ubiquitous in scientific and engineering applications. Discovering closed-form governing equations in these fields can help us understand and predict the behavior of complex dynamic systems. In the past few years, extensive work has been done in this field. However, accounting for considerable data noise and quantifying the identified system's uncertainty from noisy data is challenging, and relevant literature is still limited. To bridge this gap, we develop a Bayesian deep spline learning framework to identify closed-formed governing equations from corrupted data with quantified uncertainty. The proposed method includes (1) Gaussian-mean fields variational inference to approximate the intractable posterior distribution for the learning parameters, (2) physics-informed Spline-based network to accurately estimate the local derivatives from noisy data, (3) dictionary-based sparse regression for constructing the relevant physical terms, 4) alternative direction optimization (ADO) for systematically approximating L0 sparsity constraints. The proposed algorithm will be evaluated on a group of classic nonlinear dynamical systems in terms of predictive accuracy and uncertainty quantification.

Luning Sun
University of Notre Dame
luning.sun.73@nd.edu

Hao Sun
Northeastern University
h.sun@northeastern.edu

Jian-Xun Wang
University of Notre Dame
jwang33@nd.edu

MS72

Fast Surrogate Model for Predicting Spatiotemporal Physics in Mesh-Reduced Space based on Temporal Attention and Graph Neural Network

Developing surrogate models that enable fast state predictions is of great importance for uncertainty quantification (UQ) analysis. It is extremely challenging to build an accurate surrogate for predicting physics with spatiotemporal behavior. Graph-based next-step prediction models have recently been very successful in modeling complex high-dimensional physical systems on irregular meshes, but suffer from error accumulation and drift, due to their short temporal attention span. In this paper, we present a new mesh-reduced graph neural network (GNN) that leverages the multi-attention mechanism in temporal space. We use a GNN to locally summarize features and create coarsened, compact mesh representation of the system state, onto which we apply a transformer-style temporal attention module. A second GNN decodes these predictions back to a full-sized graph and performs fine-scale updates. Our method outperforms a competitive GNN baseline on three complex fluid dynamics prediction tasks, from sonic shocks to vascular flow. We demonstrate stable rollouts without the need for training noise, and show perfectly phase-stable predictions even for very long sequences. More broadly, we believe our approach paves the way to bringing the benefits of GNNs and attention-based sequence models to surrogate modeling of complex physics for UQ purposes.

Han Gao
University of Notre Dame
hgao1@nd.edu

Xu Han
Tufts University
xu.han@tufts.edu

Tobias Pfaff
DeepMind
tob.pfaff@gmail.com

Liping Liu
Tufts University
liping.liu@tufts.edu

Jian-Xun Wang
University of Notre Dame
jwang33@nd.edu

MS72

Spectral Deep Operator Networks

DeepONets have recently become a popular method of approximating non-linear operators using neural networks, which has many applications to the field of numerically solving PDEs and more. This talk proposes an extension of DeepONets named Spectral DeepONets, where the functions are inputted to the network using a spectral decomposition as opposed to its values on fixed sensors as is commonly done. We showcase a Universal Approximation Theorem for this new type of network and present some case studies of its efficiency. Spectral DeepONets not only broaden the possibilities where DeepONets can be applied, they can also be used as a mean of dimensionality reduction, and can improve accuracy when noise is present in the data.

Leonardo Guilhoto
University of Pennsylvania
guilhoto@sas.upenn.edu

MS72

Learning Solution Operators of Elliptic PDEs from Matrix Vector Products

In this work, we show that solvers of elliptic boundary value problems in d dimensions can be approximated to accuracy ϵ from only $\mathcal{O}\left(\log(N)\log^d(N/\epsilon)\right)$ matrix-vector products with carefully chosen vectors (right-hand sides). The solver is only accessed as a black box, and the underlying operator may be unknown and of an arbitrarily high order. Our algorithm (1) has complexity $\mathcal{O}\left(N\log^2(N)\log^{2d}(N/\epsilon)\right)$ and

represents the solution operator as a sparse Cholesky factorization with $\mathcal{O}\left(N\log(N)\log^d(N/\epsilon)\right)$ nonzero entries, (2) allows for embarrassingly parallel evaluation of the solution operator and the computation of its log-determinant, (3) allows for $\mathcal{O}\left(\log(N)\log^d(N/\epsilon)\right)$ complexity computation of individual entries of the matrix representation of the solver that in turn enables its recompression to an $\mathcal{O}\left(N\log^d(N/\epsilon)\right)$ complexity representation. We include rigorous proofs of these results, and to the best of our knowledge, the proposed algorithm achieves the best tradeoff between accuracy ϵ and the number of required matrix-vector products of the original solver. From a statistical learning point of view, our results provide an example where incorporating geometric information in the learning algorithm leads to an exponentially reduced sample complexity for learning PDEs.

Florian Schaefer
California Institute of Technology
florian.schaefer@cc.gatech.edu

Houman Owhadi
Applied Mathematics
Caltech
owhadi@caltech.edu

MS72

Learning Low-Dimensional Models from Noisy State Trajectories with Operator Inference and Re-Projection

This work investigates learning reduced models from noisy data of polynomially nonlinear dynamical systems. In the absence of noise, operator inference with data sampling via re-projection guarantees the recovery of the very same reduced operators that are obtained with traditional intrusive model reduction. In this work, we consider trajectories polluted with zero-mean Gaussian noise and show that the recovery guarantee holds in expectation. Furthermore, building on matrix concentration inequalities, we derive upper bounds in terms of the signal-to-noise ratio for the expected error of predictions made with the inferred model. The error analysis motivates a design of experiments to control the noise-to-signal ratio by judiciously selecting inputs at which to query the high-dimensional system. Numerical experiments demonstrate the derived error bounds and show that the proposed design of experiments reduces the expected error by several factors compared to a uniform sampling of the high-dimensional systems. This is joint work with Yuepeng Wang and Yuxiao Wen.

Wayne Isaac T. Uy
Courant Institute of Mathematical Sciences
wayne.uy@cims.nyu.edu or wtu4@cornell.edu

Benjamin Peherstorfer
Courant Institute of Mathematical Sciences
New York University
pehersto@cims.nyu.edu

MS73

Investigating Atmospheric Carbon Dioxide and Solar-Induced Chlorophyll Fluorescence (SIF) using Functional ANOVA

Data sampled densely in space and time have become increasingly abundant as a result of advances in modern technology, particularly in the field of remote sensing. However, the presence of complex dependence, large data volumes, and current computational limitations have made many classical inferential approaches practically infeasible. In this work, we develop an analysis of variance (ANOVA-type) method for functional data that allows comparisons among groups of time series with complex spatio-temporal dependence. We utilize this method to study the joint behavior of atmospheric carbon dioxide and solar-induced chlorophyll fluorescence (SIF), an indicator of photosynthetic activity. Satellite-based estimates of these variables are available from NASAs Orbiting Carbon Observatory-2 (OCO-2) mission. The proposed method provides robust and interpretable tests which are used to investigate carbon cycle dynamics, and the relationship between atmospheric carbon dioxide and SIF.

Manju Johny, Jonathan Hobbs, Vineet Yadhav, Margaret Johnson
Jet Propulsion Laboratory
manju.m.johny@jpl.nasa.gov,
jonathan.m.hobbs@jpl.nasa.gov,
vineet.yadav@jpl.nasa.gov, maggie.johnson@jpl.nasa.gov

Amy Braverman, Hai Nguyen
Jet Propulsion Laboratory
California Institute of Technology
Amy.Braverman@jpl.nasa.gov, hai.nguyen@jpl.nasa.gov

Petruta Caragea, Kieran Liming
Iowa State University
pcaragea@iastate.edu, kliming@iastate.edu

MS73

Forward Model Emulator for Orbiting Carbon Observatory-2 Satellite's Co2 Retrieval Inverse Problem

In recent years, satellite-based observations of atmospheric carbon dioxide (CO_2) concentrations have emerged as a means of providing data with global coverage and high spatial resolution. Satellite instruments like Orbiting Carbon Observatory (OCO-2 & OCO-3) measure absorbed solar radiation as radiances, from which a retrieval algorithm is used to infer atmospheric CO2 concentrations. Many further use cases, such as Carbon Flux Inversion, impose strict requirements for the accuracy of this data, and thus applying rigorous Uncertainty Quantification (UQ) is extremely important for assessing how well the data product can be trusted. A major computational bottleneck in conducting UQ experiments, such as Observing System Uncertainty Experiments and Markov Chain Monte Carlo sampling, is the repeated evaluation of a computationally expensive atmospheric radiative transfer physics model. To remedy this computational problem, we propose and implement a Gaussian Process based statistical emulator for the Full Physics forward model used in OCO-2 satellites CO_2 retrieval algorithm. The computationally inexpensive emulator accurately predicts the outputs of the Full Physics forward model, providing additionally an estimate of uncertainty for the predictions.

Otto Lamminpää, Jonathan Hobbs
Jet Propulsion Laboratory
otto.m.lamminpaa@jpl.nasa.gov,
jonathan.m.hobbs@jpl.nasa.gov

Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology

Amy.Braverman@jpl.nasa.gov

MS73

Optimizing Confidence Intervals for Satellite-Based Carbon Flux Inversion

Steadily increasing atmospheric carbon dioxide (CO2) concentration is largely responsible for the observed radiative forcing in Earth's climate system over the last century. Inference of land-air carbon fluxes from satellite-based CO2 observations helps uncover the responsible locations and mechanisms. Doing this is an ill-posed, likelihood-free inverse problem for which a Bayesian data assimilation is the usual approach to obtain regularized estimation and uncertainty quantification. From a frequentist perspective, such estimates can have nonnegligible bias and coverage issues. Using GEOS-Chem and the GEOS-Chem Adjoint, we directly obtain confidence intervals for linear spatio-temporal functionals via convex optimization such that frequentist coverage is guaranteed. This approach is particularly well-suited for this problem since we do not have access to an explicit forward model. Using a variety of constraints, we optimize a collection of intervals and compare and contrast them against conventional Bayesian intervals obtained through a Monte Carlo method.

Michael Stanley, Mikael Kuusela
Carnegie Mellon University
mcstanle@andrew.cmu.edu, mkuusela@andrew.cmu.edu

MS74

Verification Strategies for High-Order Eulerian Vlasov Codes

For many fields of applied science and engineering, high-order accurate numerical methods are potentially orders-of-magnitude more efficient than their low-order counterparts. This efficiency opens many interesting possibilities for simulation including larger ensemble calculations, larger system sizes, or more complex and coupled physics models. Verification of such a high-order accurate code is a critical step in the development of a practical tool. In this talk, we verification strategies employed in the high-order accurate Vlasov code LOKI. A unique approach to manufactured solutions demonstrates high-order convergence (in this case 4th and 6th order accuracy) for problems that are near to physically relevant cases. Subtleties in extracting the decay rate and frequency for Landau Damping are also discussed, and high-order accuracy is obtained.

Jeffrey W. Banks, Jennifer Gorman
Rensselaer Polytechnic Institute
banksj3@rpi.edu, gormaj4@rpi.edu

MS74

Introducing the Method of Rotated Solutions to Streamline Code Verification

The purpose of code verification is to provide mathematical evidence that a scientific computing software platform is free of bugs in the source code and that the numerical algorithms are consistent. The Method of Exact Solutions (MES) and the Method of Manufactured Solutions (MMS) are the currently accepted approaches for calculating the observed order of accuracy, which is the most stringent evidence that the numerical algorithms are error-free. And while both methods are well-established, inherent drawbacks are limiting their utilization. This presentation will outline the Method of Rotated Solutions (MRS), which relies on coordinate transformation(s) of traditional engineering problems to generate analytical solutions with non-zero values for all equation terms in all required coordinate directions. The example of pressure-driven flow in a circular pipe will be used to demonstrate the MRS procedure. This solution has traditionally been viewed as having limited applicability in code verification studies because all inertial terms and all but one of the viscous terms in the Navier-Stokes are identically zero. The example will illustrate how MRS transforms this uni-axial flow solution into a set of equations that can be used to fully verify a three-dimensional, laminar CFD code.

Marc Horner
ANSYS
marc.horner@ansys.com

MS74

Finding Confidence and Meaning in Verification

Verification seems like a fairly simple prospect, but this simplicity belies some harsh realities. Conduct mathematically well-defined tests of a code to show correctness or estimate error. Mesh convergence is utilized in both cases. For code verification this requires an analytic solutions to be used. For solution verification a sequence of convergent calculations are used to estimate errors. At the core of both activities is the Lax Equivalence Theorem, but its applicability is highly limited. First, it only applies to linear equations and virtually all our applications are nonlinear. It states that the combination of consistency and stability implies convergence. In the process of verification demonstrated stability is implied, and instead consistency is demonstrated via convergence. Its application is essentially. an act of faith. More broadly the idiosyncrasies of mathematical theory plays a key role in verification and proper interpretation of the results. A broad range of results from hyperbolic PDEs can be used to demonstrate a host of important theory explaining results and the proper expectations. When examining the results the theory should be integrated into the assessment. This is particularly true in solution verification where the practicalities of modeling realistic phenomena often rule out regularity in the model or the solution. The result are lower rates of convergence and large errors that are predictable given the theory.

William J. Rider
Sandia National Laboratories
wjrider@sandia.gov

MS74

Verification and Validation Method for An Acoustic Mode Prediction Code for Turbomachinery Noise

As airport noise limitations become more restrictive over time, reducing aircraft takeoff and landing noise remains a prominent issue in the aviation community. One popular method to reduce aircraft noise is using acoustic liners placed on the walls of the engine inlet and exhaust ducts. These liners are designed to reduce the amplitude of acoustic modes emanating from the bypass fan as they propagate through the engine. The SWIRL code is a frequency-domain linearized Euler equation solver that is designed to predict the effect of acoustic liners on acoustic modes propagating in realistic sheared and swirling mean flows, guiding the design of more efficient liner configurations.

The purpose of this study is to validate SWIRL using the Method Of Manufactured Solutions (MMS). This study also investigated the effect of the integration and spatial differencing methods on the convergence for a given Manufactured Solution. In addition, the effect of boundary condition implementation was tested. The improved MMS convergence rates shown for these tests suggest that the revised SWIRL code provides more accurate solutions with less computational effort than the original formulation.

Jeffrey Severino
The University of Toledo
jeffrey.severino@rockets.utoledo.edu

MS75

Physics-Informed and Data-Driven Predictive Models with Quantified Uncertainty

Solving partial differential equations (PDEs) is the canonical approach for understanding the behavior of physical systems. In this work, we proposed a general physics-constrained neural network (NN) approach to solve PDEs without labels, where the loss functions of NNs are expressed in terms of the discretized residual of PDEs. We studied both deterministic and probabilistic NNs, with Bayesian NNs (BNNs) for the latter to further quantify the epistemic uncertainty from model parameters and the aleatoric uncertainty from noise in the data. In our approach, both problem domains and boundary conditions (BCs) are specified as inputs to NNs, which allows a well-trained NNs to readily make predictions for new BCs and new domains orders faster than traditional numerical methods. We demonstrate the capacity and performance of the proposed framework by applying it to different steady-state and equilibrium boundary value problems (BVPs) with physics that spans diffusion, linear and nonlinear elasticity. The proposed method is shown to work for both small and large datasets, where the number of BVPs solved by a single NN ranges from one to hundreds of thousands. The trained NN solvers demonstrate a degree of success at interpolated/extrapolated predictions for new BCs and new domains that they were not exposed to during training. This framework is important for problems where high-throughput solutions of PDEs are desired in support of design and decision-making.

Xiaoxuan Zhang
University of Michigan
xxzh@umich.edu

Krishna Garikipati
Mechanical Engineering
University of Michigan
krishna@umich.edu

MS75

Efficient Derivative-Free Bayesian Inference for Large-Scale Inverse Problems

We consider Bayesian inference for large scale inverse problems, where computational challenges arise from the need for repeated evaluations of an expensive forward model. This renders most of Markov chain Monte Carlo methods infeasible, since they typically require $O(10^4)$ model runs. Moreover, the forward model are often given as a black box and may be impractical to differentiate. Therefore derivative-free algorithms are highly desirable. We propose a framework, which is built on Kalman methodology, to efficiently calibrate and provide uncertainty estimations of such models with noisy observation data. Theoretical guarantees for linear inverse problems are provided. Practical strategies, including low-rank approximation and a bi-fidelity approach, to further reduce the computational and memory cost of Kalman methodology, are presented. The effectiveness of the framework is demonstrated on several numerical experiments, including proof-of-concept linear/nonlinear examples and two applications: learning of permeability parameters in subsurface flow; and learning subgrid-scale parameters in a general circulation model from time-averaged statistics.

Daniel Huang
California Institute of Technology
dzhuang@caltech.edu

Jiaoyang Huang
New York University
jjh4427@nyu.edu

Sebastian Reich
Universität Potsdam, Germany
sereich@uni-potsdam.de

Andrew Stuart
Computing + Mathematical Sciences
California Institute of Technology
astuart@caltech.edu

MS75

Model-Constrained Bayesian Neural Networks

While Bayesian Neural Networks facilitate uncertainty quantification for neural network prediction, the uncertainty is questionable using Gaussian priors on the weights and biases. Clearly, weights/biases are artificial quantities/parameter and thus Gaussian priors are the matter of convenience instead of rationale. In this talk we present of one of the first attempts to construct a meaningful priors on the weights/biases. Theoretical results will be presented and various numerical results will be provided to justify our approach.

Tan Bui-Thanh
The University of Texas at Austin
tanbui@ices.utexas.edu

Russell Philley
The Oden Institute for computational engineering and science
rsphilley@gmail.com

Krishanunni Giri
The Department of Aerospace Engineering and Engineering
Mechanics
cg.krishnanunni@gmail.com

MS75

Multi-Variance Replica Exchange Sgmcmc for Inverse and Forward Problems via Bayesian Pinn

Physics-informed neural network has been successfully applied in solving a variety of nonlinear non-convex forward and inverse problems. However, the training is challenging because of the non-convex loss functions and the multiple optima in the Bayesian inverse problem. We propose a multi-variance replica exchange stochastic gradient

Langevin dynamics method to tackle the challenge. Two chains with different temperatures are designed where the low temperature chain aims for the local convergence, and the target of the high temperature chain is to explore the whole loss function entropy landscape. However, it may not be efficient to use the vanilla replica method since the it doubles the computational cost in evaluating the forward solvers. To address this issue, we propose to make different assumptions on the energy function estimation and one can use solvers of different fidelities in the likelihood function evaluation. Our proposed method lowers the computational cost in the high temperature chain, meanwhile preserves the accuracy and converges fast. We give an unbiased estimate of the swapping rate and give an estimation of the discretization error of the scheme. To verify our idea, we solve four inverse problems which have multiple modes. The proposed method is also employed to train the Bayesian PINN to solve the forward and inverse problems; faster and more accurate convergence has been observed when compared to the vanilla replica exchange methods.

Zecheng Zhang
Purdue University
zecheng.zhang.math@gmail.com

MS76

Training Reduced Deep Networks for Nonlinear Model Reduction

Deriving reduced models is challenging if the latent dynamics of the high-dimensional system are dominated by nonlinear behavior that renders linear approximations in classical low-dimensional approximation spaces inefficient (slowly decaying Kolmogorov n-width). Transport-dominated problems describing wave-type phenomena, strong convection, and phase transitions with sharp gradients typically lead to such latent dynamics. This presentation discusses model reduction methods for constructing nonlinear reduced models that seek approximations on manifolds, rather than in subspaces, and so lead to efficient dimensionality reduction even for transport-dominated problems. The proposed reduced models are compositions of multiple layers of approximations and can be interpreted as deep neural networks with reduced layers. Numerical results demonstrate that the proposed models can be efficiently trained from few data points and approximate well transport-dominated dynamics with fewer degrees of freedom than traditional, linear reduced models.

Nitin Shyamkumar, Yuxiao Wen, Benjamin Peherstorfer
Courant Institute of Mathematical Sciences
New York University
nhs9171@nyu.edu, yw3210@nyu.edu, pehersto@cims.nyu.edu

MS76

Trends in Model Reduction for Flow Control Problems with Random Inputs

This talk focuses on tailored reduced order techniques for parametrized Optimal Control Problems (OCP(μ)s) governed by Partial Differential Equations (PDE(μ)s) under the action of random inputs. Stochastic PDE(μ)s describe uncertainty in the considered setting and, together with statistical analysis, are an asset to reach more reliable simulations. Moreover, OCP(μ)s increase the model accuracy, filling the gap between equations and collected data. A deep analysis of such complicated problems relies on an unbearable amount of simulations. Thus, we propose weighted reduced order methods (w-ROMs) to solve the stochastic OCP(μ)s to accelerate Monte Carlo simulations. Namely, a low dimensional framework is built, exploiting the probability distribution of the underlying parameters. Besides this information, we also study the influence of several quadrature rules on the numerical approximation of the problem. The methodology is validated with numerical tests in environmental sciences, a field where stochastic OCP(μ)s are a tool of utmost usefulness to reach accurate predictions. Two numerical experiments are presented: a pollutant control in the Gulf of Trieste and a model forecast of the North Atlantic ocean dynamic [G. Carere, M. Strazzullo, F. Ballarin, G. Rozza, and R. Stevenson, A weighted POD-reduction approach for parametrized PDE-constrained Optimal Control Problems with random inputs and applications to environmental sciences. Submitted,2021].

Gianluigi Rozza
SISSA, International School for Advanced Studies,
Trieste, Italy
grozza@sissa.it

Maria Strazzullo
mathLab, Mathematics Area, SISSA International School for Advanced Studies
mstrazzu@sissa.it

MS76

Stabilization in Time Series Learning for Transport Dominated Problems by Neural Networks

In this talk, we consider the use of residual neural networks for the approximation of time series arising from reduced order modeling of hyperbolic systems. It is well understood that such residual networks are quite similar to time stepping methods for ODEs/PDEs. For hyperbolic problems, the latter require a careful choice of the discretization in order to obtain stable methods. In the talk, we observe that such instabilities can manifest as ill conditioned loss functions for the training of corresponding neural networks, requiring excessively small learning rates and complicating the learning process. While hand-crafted penalty terms may alleviate the problem, we consider a more automated way to learn a stabilization for a class of interrelated problems that eases the loss function and thus simplifies the training process.

Gerrit Welper
University of Central Florida
gerrit.welper@ucf.edu

MS77

All-at-Once Model Tuning for Multifidelity Sampling

Computational simulation continues to advance in its predictive capability through the development of high-fidelity multi-physics/multi-scale simulation models, with unprecedented resolution enabled by the latest high-performance computers. Uncertainty quantification (UQ) methodologies are challenged in this environment, both by the prohibitive cost of computing high-fidelity ensembles and by the increasing number of uncertainty sources that is often induced by this model complexity. One strategy to address these challenges is to harness the utility that exists within an ensemble of model forms and resolution levels in order to control multiple sources of error while optimiz-

ing the allocation of simulation resources. An important challenge that is commonly encountered in practice is the need to identify predictive low fidelity approximations that can be effective within a multifidelity setting. Given a set of hyper-parameters that govern the balance of accuracy versus cost for one or more low fidelity models, one can formulate an optimization problem to tune these models in order to provide the greatest utility. Here, we propose performing this optimization within the context of a particular statistical estimator and employing an all-at-once approach to minimizing the estimator's variance. Our initial demonstrations will focus an sampling approaches, especially multifidelity Monte Carlo (MFMC) and approximate control variates (ACV).

Michael S. Eldred
Sandia National Laboratories
Optimization and Uncertainty Quantification Dept.
mseldre@sandia.gov

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

Alex Gorodetsky
University of Michigan
goroda@umich.edu

MS77

Analysis of Multifidelity Monte Carlo Applied to Chaotic Systems

Multifidelity Monte Carlo methods rely upon correlations between models to generate high accuracy estimators of statistics of interest at minimal computational cost. When the forward model is chaotic or stochastic, the effects of finite sampling in each forward run tend to reduce the correlations between model below what they would be with infinite sampling. While the reduction in these correlations tends to reduce the effectiveness of multifidelity methods, it also introduces additional parameters of the forward simulation that may be optimized to improve performance. Examples of such parameters include the averaging time when approximating the expectation in simulating an ergodic chaotic system and the number of particles in particle-in-cell-based simulations. In general, these parameters control the finite-sampling error in each forward simulation, enabling increased correlations between models at different levels at the cost of more expensive forward simulations. In this talk we explore these trade-offs both analytically and computationally for model problems based on the Lorenz equations and the Vlasov-Poisson equations.

Todd A. Oliver
PECOS/ICES, The University of Texas at Austin
oliver@ices.utexas.edu

Robert Moser
UT Austin
Oden Institute
rmoser@oden.utexas.edu

MS77

Efficient Multifidelity Strategies for Uncertainty Quantification of Non-Deterministic Models

Multifidelity (MF) approaches for uncertainty quantification aim to improve the predictivity of high-fidelity computational models by leveraging information from different sources which vary in both accuracy and cost. Common across engineering disciplines, non-deterministic solvers (ones with an uncontrollable source of stochasticity) are candidates for both low- and high-fidelity information sources. Examples are found in turbulent flow simulations, particle-in-cell methods for plasmas, and stochastic modeling of material microstructures. MF methods exploit correlations between models to learn the response of quantities of interest to changes in the uncertain, controllable input space in a more computationally efficient manner. The intrinsically uncontrollable stochasticity of a non-deterministic model weakens these correlations. The stochastic noise can be mitigated by, e.g., running multiple simulations at the same point in input space or a longer time-average of a chaotic system, but this increases cost. Therefore, for a fixed computational budget, there is a trade-off between exploring the input space and improving the correlations at any one point. This leads to a more complicated optimization problem whose solution is both the allocation of samples amongst models as well as the required averaging over the stochastic response. In this talk, we explore the efficacy of common MF techniques for various model hierarchies which include non-deterministic models.

Bryan Reuter
Sandia National Laboratories
bwreute@sandia.gov

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

Tim Wildey
Optimization and Uncertainty Quantification Dept.
Sandia National Laboratories
tmwilde@sandia.gov

MS77

Automating Model Selection and Tuning for Multifidelity UQ

This work focuses on developing strategies for optimal model selection and tuning for multifidelity UQ methods. In general, multifidelity estimators provide improved accuracy and efficiency for estimating statistics of computationally expensive simulations by leveraging an ensemble of low-fidelity models with lower cost. Existing work has made significant progress on determining optimal sample allocations that lead to maximum variance reduction for a given computational budget and a *fixed* collection of models. However, the efficacy of multifidelity methods is critically dependent on the characteristics of the models themselves, specifically the degree of correlation of the low-fidelity models with the high-fidelity model, the correlation amongst the low-fidelity models, and their relative computational cost. Furthermore, the cost/accuracy of a low-fidelity model can often be easily tuned via model-specific hyperparameters (e.g., the element size for a finite element mesh or time step for a dynamic model). This work investigates the open problem of optimal model selection for multifidelity UQ methods by including model tuning/selection within the sample allocation optimization problem. The advantages of automating model selection in this manner are demonstrated in the context of trajectory

simulation for entry, descent, and landing applications.

James Warner
Nasa Langley Research Center
james.e.warner@nasa.gov

Geoffrey Bomarito, Patrick Leser, William Leser
NASA Langley Research Center
geoffrey.f.bomarito@nasa.gov, patrick.e.leser@nasa.gov, william.p.leser@nasa.gov

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

Michael S. Eldred
Sandia National Laboratories
Optimization and Uncertainty Quantification Dept.
mseldre@sandia.gov

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

MS78

Computationally Efficient Methods for Large-Scale Atmospheric Inverse Modeling with Mean Estimation

Atmospheric inverse modeling describes the process of estimating greenhouse gas fluxes or air pollution emissions at the Earth's surface using observations of these gases collected in the atmosphere. In this talk we describe generalized hybrid projection methods, which are iterative methods for estimating surface fluxes. These algorithms confer several advantages. They are efficient, in part because they converge quickly, they exploit efficient matrix-vector multiplications, and they do not require inverting any matrices. These methods are also also robust because they can accurately reconstruct surface fluxes, they are automatic since regularization or covariance matrix parameters and stopping criteria can be determined as part of the iterative algorithm, and they are flexible because they can be paired with many different types of atmospheric models. We demonstrate the benefits of generalized hybrid methods with a case study from NASA's Orbiting Carbon Observatory 2 (OCO-2) satellite. We then address the more challenging problem of solving the inverse model when the mean of the surface fluxes is not known a priori; we do so by reformulating hybrid projection methods using hierarchical priors. We further show that by exploiting mathematical relations provided by the generalized hybrid method, we can efficiently calculate approximate posterior uncertainties.

Julianne Chung
Department of Mathematics
Virginia Tech
jmchung@vt.edu

MS78

Infemus: An Infinite-Dimensional Method for Marginal Likelihood Estimation

We investigate applications of the Eigenvector Method for Umbrella Sampling (EMUS) to the exploration of the marginal likelihood in Bayesian models. This method is more general, hence more efficient than popular algorithms in this context, such as bridge sampling. We formulate inference for the marginal likelihood as a function estimation problem and appropriately combine EMUS with a Gaussian process such that EMUS provides a data model with a structured error distribution and the Gaussian process prior imposes smoothness. Inference for the marginal likelihood proceeds by building on Gaussian process regression ideas. Importantly, our framework allows for a dynamic experimental design and the refinement of the lattice on which EMUS is applied.

Timothée Stumpf-Fetizon
University of Warwick
tim.stumpf-fetizon@warwick.ac.uk

MS78

On the Frequentist Accuracy of Bayesian Uncertainty Quantification Procedures in Imaging Problems

One of the main strengths of Bayesian imaging strategies is their ability to quantify uncertainty in the solutions delivered and support inferences such as hypothesis tests. If the marginal distribution the unknown image were known and used as a prior distribution in a Bayesian model, the probabilities from a Bayesian analysis would also be valid in a frequentist sense (i.e. Bayesian probabilities would coincide with the frequencies obtained over many repetitions of the same experiment). For example, if a large number of replicate experiments were performed, the 95% Bayesian credible interval would cover the true image in 95% of the experiments. However, the prior distributions used in Bayesian imaging models are not specified this way. We conducted a study in which the coverage properties, under a TV prior, a TGV prior, and a Plug-and-Play (PnP) prior, were estimated through Monte-Carlo. Using large image data-sets as a sample from the distribution of the images (or class of images) of interest, credible intervals were obtained from samples of the posterior distribution using the MYULA (or PnP-ULA) algorithm. The regularization parameters (for the TV and TGV priors) were estimated using an empirical Bayes approach. The methodology is illustrated on a range of examples, such as non-blind deblurring and MRI reconstruction from a limited number of radial lines. We find that the credible intervals are conservative but highly reliable, in the TGV and TV cases, and include the true image with high probability. With PnP priors the credible intervals were found to be overconfident.

David Thong
Heriot-Watt University
d.thong@hw.ac.uk

Marcelo A. Pereyra
Maxwell Insitute for Mathematical Science
Heriot-Watt University
m.pereyra@hw.ac.uk

MS78

Learning Regularization Parameters via Stochastic Bilevel Optimization: Consistency Analysis

One fundamental problem when solving inverse problems is how to find regularization parameters. In this talk, we consider solving this problem by using a data-driven approach via bilevel optimization. The regularization parameter will be adaptively learned from data by means of stochastic optimization. Therefore, we formulate the corresponding empirical risk minimization problem and analyze its perfor-

mance in the large sample size limit for general nonlinear problems. In order to reduce the associated computational costs, we derive online numerical schemes using stochastic gradient descent. Under suitable assumptions on the forward problem we prove convergence of these numerical schemes and demonstrate the applicability and efficiency in numerical examples for various linear and nonlinear inverse problems.

Simon Weissmann
University of Mannheim
simon.weissmann@uni-heidelberg.de

MS79

Machine Learning Failure Probabilities from Multi-Fidelity Models

Despite the use of efficient sampling methods, reliability analysis for complex engineering systems remains a challenge, primarily due to the high computational cost of evaluating the system response at each iteration. One way around this challenge is to use Machine Learned surrogates in tandem with multi-fidelity modeling to reduce computation times while maintaining acceptable levels of accuracy. Adopting this approach, our proposed framework functions by estimating the high-fidelity (HF) model predicted system state (safe/fail), using multiple pairs of low-fidelity (LF) models and machine-learned corrections (MLCs). These correction terms are trained using small sets of HF and LF evaluations, and the information from all the LF-MLC pairs are combined through a simple probabilistic formulation to estimate the state of the system response at any input point, i.e., whether the system lies in the safety or the failure zone. No assumptions are made on LF model types or correlations with the HF model. Using this compound state estimate, in conjunction with active learning functions that check for acceptability of the estimate and call the HF model when the estimates are unacceptable, we conduct reliability analysis using Subset Simulation (a variance-reduced MCMC method). Lastly, our proposed framework is tested on several analytical case studies to demonstrate its efficiency.

Promit Chakroborty
Johns Hopkins University
pchakro1@jhu.edu

Som L. Dhulipala
Idaho National Laboratory
Som.Dhulipala@inl.gov

Michael D. Shields
Johns Hopkins University
michael.shields@jhu.edu

MS79

Quantified Uncertainty for Safe Operation of Particle Accelerators

In order to achieve reliable deployment of deep learning tools on safety-critical systems, such as particle accelerators, the ability to make informed decisions under uncertainty is crucial. Particle accelerators are noisy systems that are difficult to control and operate safely and efficiently. Characterizing the electron beam properties is key. In this work, we aim to accurately and confidently predict the electron beam properties by using various compositions of neural networks to explore and enhance prediction uncertainty and robustness. The results show that the deep learning approach helps in situations that are beyond the capabilities of conventional tools. We also show its robustness against out-of-distribution inputs, and extract information from the latent space representations. Providing shot-to-shot estimates of beam properties will also help the scientific users with their data analysis.

Adi Hanuka
SLAC Stanford
hanukaadi@gmail.com

Owen Convery
SLAC National Accelerator Laboratory
owenco@slac.stanford.edu

Lewis Smith, Yarin Gal
University of Oxford
lsgs@robots.ox.ac.uk, yarin.gal@cs.ox.ac.uk

MS79

Uncertainty-Aware Human-Machine Teaming in Scientific AI and Simulation

Simulation has become an indispensable tool for researchers across the sciences to explore the behavior of complex, dynamic systems under varying conditions, including hypothetical or extreme conditions. Recent advances in AI and ML applied to simulations help alleviate significant bottlenecks in computation, such as physics-infused learning for accelerated surrogate models in climate science and nuclear energy, or *in silico* optimization of *in situ* experiment- or sensor-design in particle physics and materials science. However, on top of the existing challenges practitioners face working with simulations, scientific computation and data, the addition of AI/ML injects more complexities and uncertainties in already risk-averse environments. It is thus more critical than ever to reliably estimate uncertainties in data, models, and broader system components such as sensors and feedback loops involving downstream tasks and users. Uncertainty reasoning with AI and simulation systems is paramount for usability and trust, and can further be leveraged to improve human-machine teaming – uncertainty-aware systems lead to human-machine synergies. This talk will explore the current challenges and opportunities in the context of real-world examples at the forefront of Earth systems and life sciences.

Alex Lavin
Institute for Simulation Intelligence
lavin@simulation.science

MS79

Hierarchical Multiscale Uncertainty Quantification in Material Modeling

The macroscopic properties of materials and structures that we observe and exploit in engineering applications result from complex interactions between physics at multiple length and time scales: electronic, atomistic, defects, domains, etc. Multiscale modeling seeks to understand the interactions between these physics across scales. However, assessing such interactions can be challenging due to the complex nature of material properties and the prohibitive computational cost of integral calculations. This talk will focus on how to quantify the propagation of material uncertainties across multiple scales. To this end, we exploit the multiscale and hierarchical nature of material response and develop a framework to quantify the overall uncer-

tainty of material response induced by the uncertainties at finer scales without the need for integral calculations. Specifically, we bound the uncertainty at each scale and then combine the partial uncertainties in a way that provides a bound on the overall or integral uncertainty. The bound provides a conservative estimate on the uncertainty. Importantly, this approach does not require integral calculations that are prohibitively expensive. Finally, the developed framework has been employed to investigate how material uncertainties propagate from the single-crystal properties of magnesium to the ballistic performance of armor structures.

Xingsheng Sun
Department of Mechanical Engineering
University of Kentucky
xingsheng.sun@uky.edu

Burigede Liu
University of Cambridge
BL377@cam.ac.uk

Kaushik Bhattacharya
Howell N. Tyson Sr. Professor of Mechanics
California Institute of Technology
bhatta@caltech.edu

Michael Ortiz
California Institute of Technology
ortiz@aero.caltech.edu

MS80

Petrov-Galerkin Methods for the Construction of Non-Markovian Dynamics Preserving Nonlocal Statistics

A common observation in the learning reduced model of multiscale problems is the non-Markovian behavior, primarily due to the lack of scale separations. To learn a high-fidelity reduced model that gives rise to the correct non-Markovian statistical properties, we propose a Galerkin projection approach, which transforms the exhausting effort of finding an appropriate model to choosing appropriate subspaces in terms of the derivatives of the coarse-grained variables, and at the same time, provides an accurate approximation to the non-Markovian memory term. We introduce the notion of fractional statistics that embodies nonlocal properties. More importantly, we show how to pick subspaces in the Galerkin projection so that those statistics are automatically matched.

Huan Lei
Michigan State University
leihuan@msu.edu

MS80

Majorization Minimization-Based Laplace Approximation for Bayesian Inverse Problems with Arbitrary Prior and Noise Models

We develop sampling based iterative methods for computing the solutions to large scale Gaussian prior Bayesian inverse problem when the data are contaminated by the presence of noise whose statistical properties are a combination of different distributions. We consider the cases when the data are contaminated by Poisson, Laplace or a mixture of Poisson and Gaussian noise. The solution of the non-linear optimization problem can be found by solving a sequence of linear least square problems that arise from the linear approximation of the nonlinear problem at each iteration by a majorization-minimization (MM) strategy with quadratic tangent majorant, which allows the resulting least squares problem to be solved by a Krylov subspace method. Numerical examples from various applications and with arbitrary prior and noise, illustrate the effectiveness of the described framework.

Mirjeta Pasha
Arizona State University
mpasha3@asu.edu

Felipe Uribe
Technical University of Denmark
Department of Applied Mathematics and Computer Science
furca@dtu.dk

MS80

Dimension Reduction of the Posterior Distribution from a Bayesian Inverse Problem

In this work, we develop a deep-learning technique to address the dimension reduction for high-dimensional distributions. We combine the canonical variational autoencoder (VAE) and a transport map called KRnet, where VAE is used as a dimension reduction technique to capture the latent space, and KRnet is used to model the distribution of the latent variables. The variational Bayes approaches are usually based on the minimization of the Kullback-Leibler (KL) divergence between the model and the posterior, which often underestimates the variance if the model capability is not sufficiently strong. To alleviate the underestimation of variance, we include into the loss the maximization of the mutual information between the latent random variable and the original one. Numerical experiments have been presented to demonstrate the effectiveness of our model.

Xiaoliang Wan
Louisiana State University
Department of Mathematics
xlwan@math.lsu.edu

MS80

An Autoencoder Based Deeponet for Dimension Reduction, Operator Learning and Uncertainty Quantification

Deep operator network(DeepONet) has recently been proposed to deal with the problem of learning operators through deep neural networks(DNNs). The network consists of a branch net to encode the input functions and a trunk net to encode the domain of the output functions. DeepONet has small generalization error and the training and testing errors decay quickly with the size of training data. In this paper, we focus on the problem of learning operators in the form of stochastic differential equations(SDEs). For this kind of problem, it can be challenging by using DeepONet especially when the dimension of the operator is very high. Thus we combine the autoencoder and DeepONet to build a new network: auto-DeepONet. A convolutional encoder is designed to reduce the dimensionality as well as discover the hidden features of high-dimensional inputs. The decoder consists of two DeepONets and they share a common branch net which takes the hidden features as input and output a basis. The first DeepONet is designed to reconstruct the input function involving randomness and the second DeepONet is

used to find an approximation of the solution of desired equations. By adding L_1 regularization to the two trunk nets, we found both the coefficients(outputs of trunk nets) and the basis(outputs of branch net) are sparse. We conduct several numerical experiments to illustrate the effectiveness of our proposed auto-DeepONet model with uncertainty quantification.

Jiahao Zhang, Shiqi Zhang, Guang Lin
Purdue University
jiahzhang2@outlook.com, zhan2585@purdue.edu, guanglin@purdue.edu

MS81

The Representation and Optimization of Monotone Triangular Transports

We propose a general framework to robustly characterize probability distributions by estimating triangular transport maps. Transport maps deterministically couple two distributions via a bijective transformation. Yet, learning the parameters of such constrained transformations in high dimensions is challenging given few samples from the unknown target distribution, and structural choices for these transformations can have a significant impact on their performance and the optimization procedure for finding these maps. Here we present a framework for representing and learning monotone triangular maps, via invertible transformations of smooth functions, and demonstrate that the associated minimization problem has no spurious local minima, i.e., all local minima are global minima.Given a hierarchical basis for the appropriate function space, we propose a sample-efficient adaptive algorithm that estimates a sparse approximation for the map. We demonstrate how this framework can be applied for joint and conditional density estimation, likelihood-free inference, and structure learning of directed graphical models with stable generalization performance across a range of sample sizes.

Ricardo Baptista
MIT
rsb@mit.edu

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

Olivier Zahm
Univ. Grenoble Alpes, Inria
olivier.zahm@inria.fr

MS81

Diffusion Schrodinger Bridges - From Generative Modeling to Posterior Simulation

Denoising diffusion models, also known as score-based generative models, have recently emerged as a powerful class of generative models. They provide state-of-the-art results, not only for unconditional simulation, but also when used to sample from complex posterior distributions arising in a wide range of inverse problems such as image inpainting or deblurring. A limitation of these models is that they are computationally intensive as obtaining each sample requires simulating a non-homogeneous diffusion process over a long time horizon. We show here how a a Schrodinger bridge formulation of generative modeling leads to a theoretically grounded algorithm shortening generation time which is complementary to other proposed acceleration techniques. We further extend the Schrodinger bridge framework to perform posterior simulation. We demonstrate this novel methodology on various applications including image super-resolution and optimal filtering for state-space models.

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

MS81

Covariance-Modulated Optimal Transport

We study the properties of a generalized Wasserstein metric space, where the L^2 inner product of the dynamical formulation is weighted by the covariance metric at this point in probability space. Such a metric arises for example as the mean-field limit of certain Ensemble Kalman methods such as the Kalman-Bucy filter and the Ensemble Kalman Sampler, providing a gradient flow structure. We introduce a splitting into shape and moments that allows to generalize a number of results for the classical Wasserstein metric.

Franca Hoffmann
University of Bonn
franca.hoffmann@hcm.uni-bonn.de

Burger Martin
FAU
martin.burger@fau.de

Matthias Erbar
University of Bonn
erbar@iam.uni-bonn.de

Daniel Matthes
Department of Mathematics
Technische Universitaet Muenchen
matthes@ma.tum.de

André Schlichting
Institut für Angewandte Mathematik
Universität Bonn
schlichting@iam.uni-bonn.de

MS81

Towards Practical Estimation of Transport Maps

Given two probability distributions in R^d, a transport map is a function which maps samples from one distribution into samples from the other. For absolutely continuous measures, Brenier proved a remarkable theorem identifying a unique canonical transport map, which is "monotone" in a suitable sense. We study the question of whether this map can be efficiently estimated from samples. The minimax rates for this problem were recently established by Hutter and Rigollet (2021), but the estimator they propose is computationally infeasible in dimensions greater than three. We propose two new estimators—one minimax optimal, one not—which are significantly more practical to compute and implement. The analysis of these estimators is based on new stability results for the optimal transport problem and its regularized variants.

Jonathan Niles-Weed
New York University
jnw@cims.nyu.edu

Tudor Manole, Sivaraman Balakrishnan, Larry Wasserman
Carnegie Mellon University
tmanole@andrew.cmu.edu, siva@stat.cmu.edu, larry@stat.cmu.edu

Aram-Alexandre Pooladian
New York University
aram-alexandre.pooladian@nyu.edu

MS82

Learning the Probability Distribution of Random Initial States in Nonlinear Dynamical Systems Using Deep Neural Networks

We propose a new method to infer the joint probability density (PDF) of random parameters and initial states in nonlinear dynamical systems based on noisy measurements of quantities of interest. The key idea is to minimize the Kantorovich-Rubinstein dual of the Wasserstein distance between the PDF of the systems state at a finite number times and the PDF of the measurements. The evaluation of this distance involves appending the dynamical flow to a deep neural network architecture that allows us to sample the systems state efficiently. The resulting problem, is solved using off the shelf optimization methods, and it yields an accurate estimation of the joint PDF of the dynamical systems initial state and random parameters using only partial state space measurements. We demonstrate the new framework in prototype applications.

Panos Lambrianides
University of California, Santa Cruz
panos@soe.ucsc.edu

Daniele Venturi
Department of Applied Mathematics
University of California Santa Cruz
venturi@ucsc.edu

Qi Gong
University of California, Santa Cruz
qigong@soe.ucsc.edu

MS82

Newton-Type Methods for Risk-Averse PDE-Constrained Optimization Problems

This talk presents modifications of Newton-Conjugate Gradient (CG) methods for risk-averse optimization problems with PDE constraints, which substantially reduce the computational cost. Risk-averse optimization formulations use a risk measure penalize high-cost, rare events, but are computationally difficult to solve as many risk measures such as the Conditional Value-at-Risk (CVaR) or buffered probability of failure are non-smooth and require sampling in the tail of a complex, unknown probability distribution. Several optimization approaches for risk-averse optimization, including direct smoothing, augmented Lagrangian, or log-barrier methods lead to subproblems with smoothed risk measures, which can be solved using Newton-CG methods. The first modification overcomes difficulties resulting from (near) rank-difficient Hessians, which arise from the structure of smoothed risk measures. The other modification exploits the structure of the risk measure to use only a small number of samples to compute good approximations of gradients and Hessians. We present convergence result and apply our modified Newton-CG methods to model problems arising from optimal control of PDEs under uncertainty. In many examples, the modifications reduce the solution cost by a factor of six.

Mae Markowski
Rice University
mae.markowski@rice.edu

Matthias Heinkenschloss
Department of Computational and Applied Mathematics
Rice University
heinken@rice.edu

MS82

Multilevel Monte Carlo Estimators for Derivative-Free Optimization under Uncertainty

In the field of optimization under uncertainty, we consider nonlinear constrained optimization formulations that include stochastic parameters. The problems are based on PDEs where we assume that derivatives are unavailable and the underlying model is a black box. To handle such stochastic optimization problems, we use measures of robustness like the expected value or a combination of expected value and standard deviation, to ensure that the optimal solutions are robust with respect to the uncertainties. Those measures can be estimated by using a single level Monte Carlo method which, however, quickly becomes computationally expensive due to its slow convergence. We can improve the estimator efficiency by employing a multilevel Monte Carlo approach, if a hierarchy of levels is available. In this work we show how a multilevel Monte Carlo estimator has to be adapted to these problem formulations and the respective optimization problems. We present multilevel Monte Carlo estimators for the standard deviation and for the combination of mean and standard deviation. We show that by using these estimators, which optimally allocate resources for these target statistics, we can retain the required accuracy, while also reducing the computational cost. We present the implementation for these estimators in the software Dakota coupled with the optimization software SNOWPAC and show results for benchmark problems as well as more challenging numerical tests.

Friedrich Menhorn
Department of Informatics
Technical University of Munich
menhorn@in.tum.de

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

Daniel T. Seidl
Sandia National Laboratories
dtseidl@sandia.gov

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

Michael S. Eldred
Sandia National Laboratories
Optimization and Uncertainty Quantification Dept.
mseldre@sandia.gov

Hans-Joachim Bungartz
Technical University of Munich, Department of Informatics

Chair of Scientific Computing in Computer Science
bungartz@in.tum.de

MS82

Solving and Learning Nonlinear PDEs with Gaussian Processes

We introduce a simple, rigorous, and unified framework for solving nonlinear partial differential equations (PDEs), and for solving inverse problems (IPs) involving the identification of parameters in PDEs, using the framework of Gaussian processes. The proposed approach: (1) provides a natural generalization of collocation kernel methods to nonlinear PDEs and IPs; (2) has guaranteed convergence for a very general class of PDEs, and comes equipped with a path to compute error bounds for specific PDE approximations; (3) inherits the state-of-the-art computational complexity of linear solvers for dense kernel matrices. Most traditional approaches to IPs interleave parameter updates with numerical solution of the PDE; our algorithm solves for both parameter and PDE solution simultaneously. This is a joint work with Yifan Chen, Bamdad Hosseini and Andrew Stuart. The corresponding paper can be found at Journal of Computational Physics Volume 447, 2021 and arXiv:2103.12959

Houman Owhadi
Applied Mathematics
Caltech
owhadi@caltech.edu

Yifan Chen
Caltech
yifanc@caltech.edu

Bamdad Hosseini
California Institute of Technology
bamdadh@uw.edu

Andrew Stuart
Computing + Mathematical Sciences
California Institute of Technology
astuart@caltech.edu

MS83

Entropy-Based Adaptive Design for Contour Finding and Estimating Reliability Conservation

In reliability analysis, methods used to estimate failure probability are often limited by the costs associated with model evaluations. Many of these methods, such as multifidelity importance sampling (MFIS), rely upon a computationally efficient, surrogate model like a Gaussian process (GP) to quickly generate predictions. The quality of the GP fit, particularly in the vicinity of the failure region(s), is instrumental in supplying accurately predicted failures for such strategies. We introduce an entropy-based GP adaptive design that, when paired with MFIS, provides more accurate failure probability estimates and with higher confidence. We show that our greedy data acquisition strategy better identifies multiple failure regions compared to existing contour-finding schemes. We then extend the method to batch selection, without sacrificing accuracy. Illustrative examples are provided on benchmark data as well as an application to an impact damage simulator for National Aeronautics and Space Administration (NASA) spacesuits.

Austin Cole
GlaxoSmithKline, Inc
1250 S Collegeville Rd, Collegeville, PA 19426
austin.cole8@vt.edu

Robert Gramacy
Virginia Tech
rbg@vt.edu

James Warner
Nasa Langley Research Center
james.e.warner@nasa.gov

Geoffrey Bomarito, Patrick Leser, William Leser
NASA Langley Research Center
geoffrey.f.bomarito@nasa.gov, patrick.e.leser@nasa.gov, william.p.leser@nasa.gov

MS83

Exploring the Building Blocks of Adaptive Sampling Methods for Kriging

Metamodels aim to approximate characteristics of functions or systems from the knowledge extracted on only a finite number of samples. In recent years Gaussian process regression or Kriging has emerged as a widely applied surrogate technique for resource-intensive computational experiments due to its statistical fundamentals. However its prediction quality is highly dependent on the size of its dataset and the distribution of the given training points. Hence, in order to build proficient Kriging models with as few samples as possible, adaptive sampling strategies have gained considerable attention in the last 30 years. These techniques aim to find relevant sample points in an iterative manner based on information extracted from the current metamodel. In this talk we present an overview of the basic ideas and goals of adaptive sampling with a focus on Kriging. We do this by exposing characterizing features of existing methods which allow us to identify the building blocks of commonly applied adaptive sampling techniques. We show how these blocks can be combined to build efficient and reliable sampling approaches and highlight this based on examples found in the literature. Finally, we offer insights on best practices and compare the performances of different state-of-the-art techniques on benchmark problems.

Jan Niklas Fuhg
Cornell University
414 Upson Hall,124 Hoy Rd, Ithaca, NY 14850
jf853@cornell.edu

Amelie Fau
Laboratoire de Mécanique et Technologie
ENS Paris-Saclay
amelie.fau@ens-paris-saclay.fr

Udo Nackenhorst
Leibniz Universität Hannover
nackenhorst@ibnm.uni-hannover.de

MS83

Active Learning Approaches to Simulation-Driven Design in Marine Engineering

The design of complex industrial systems like aerial, ground, and water-borne vehicles demands the use of high-fidelity numerical solvers with fine computational grids to accurately assess the design performance and make sound

design decisions. The latter can be achieved by combining computational fluid dynamics simulation solvers with a shape/design modification tool and an optimization algorithm into an automatic simulation-driven design optimization (SDDO) framework. Since the optimization algorithm may require a large number of function evaluations to converge to the final solution, especially if a global optimum is desired, the computational cost of SDDO could become unaffordable. To reduce the computational cost of the SDDO, supervised active learning (SAL) methods have been developed and successfully applied in several engineering fields. Specifically, herein SAL is intended as surrogate models (supervised learning; e.g., radial basis functions and Gaussian process) with adaptive sampling procedures or infill criteria (active learning). The use of four active learning approaches is presented and discussed for single- and multi-fidelity surrogate models, applied to analytical benchmarks and SDDO problems in marine engineering.

Andrea Serani, Riccardo Pellegrini
Consiglio Nazionale delle Ricerche
Istituto di Ingegneria del Mare (CNR-INM)
andrea.serani@cnr.it, riccardo.pellegrini@inm.cnr.it

Jeroen Wackers
Ecole Central de Nantes
jeroen.wackers@ec-nantes.fr

Matteo Diez
CNR-INM
National Research Council-Institute of Marine Engineering
matteo.diez@cnr.it

MS83

Batch-Sequential Design and Heteroskedastic Surrogate Modeling for Delta Smelt Conservation

Delta smelt is an endangered fish species in the San Francisco estuary that have shown an overall population decline over the past 30 years. Researchers have developed a stochastic, agent-based simulator to virtualize the system, with the goal of understanding the relative contribution of natural and anthropogenic factors suggested as playing a role in their decline. However, the input configuration space is high-dimensional, running the simulator is time-consuming, and its noisy outputs change nonlinearly in both mean and variance. Getting enough runs to effectively learn input–output dynamics requires both a nimble modeling strategy and parallel supercomputer evaluation. Recent advances in heteroskedastic Gaussian process (HetGP) surrogate modeling helps, but little is known about how to appropriately plan experiments for highly distributed simulator evaluation. We propose a batch sequential design scheme, generalizing one-at-a-time variance-based active learning for HetGP surrogates, as a means of keeping multi-core cluster nodes fully engaged with expensive runs. Our acquisition strategy is carefully engineered to favor selection of replicates which boost statistical and computational efficiencies when training surrogates to isolate signal in high noise regions. Design and modeling performance is illustrated on a range of toy examples before embarking on a large-scale smelt simulation campaign and downstream high-fidelity input sensitivity analysis.

Boya Zhang
Computational Engineering Department, Lawrence Livermore Nat
218-2-288, 700 East Avenue, Livermore, CA, 94550
boya66@vt.edu

Robert Gramacy, Leah Johnson
Virginia Tech
rbg@vt.edu, lrjohn@vt.edu

Kenneth Rose
University of Maryland Center for Environmental Science, Point Laboratory, Cambridge, MD
krose@umces.edu

Eric Smith
Virginia Tech
epsmith@vt.edu

MS84

High-Dimensional Bayesian Optimization at Scale

Bayesian optimization is a powerful framework that leverages uncertainty quantification to efficiently solve black-box optimization problems. Much work in this area has focused on optimization problems where the objective function is highly expensive to evaluate, where one is limited to a few hundred evaluations at most. However, many of the most interesting optimization problems today are challenging for entirely different reasons. For example, a scientist that wants to optimize the properties of a molecule may have access to libraries of millions of known molecules and their properties, but may still find the problem difficult due to the extreme dimensionality and structured nature of the input space. In this talk, I will discuss an adaptation of Bayesian optimization to these large scale structured optimization problems. To accomplish this, we develop an approach that scales to millions of observations, optimizes high dimensional problems efficiently, and even handles structured or discrete inputs. We demonstrate that Bayesian optimization can achieve state-of-the-art performance on several popular molecule optimization benchmark problems by large margins, and show similar performance on a variety of other tasks.

Jacob Gardner
University of Pennsylvania
jacobrg@seas.upenn.edu

MS84

A Probabilistic, Data-Driven Framework for Interpretable Surrogates of Dynamical Systems

Despite recent successes from applications of data-driven methods in the field of computational physics, significant challenges remain especially in the context of multiscale systems and in Small-Data regimes. Given high-dimensional time-series data from a multiscale dynamical system, we present a probabilistic framework that simultaneously addresses the tasks of dimensionality reduction and model compression by identifying interpretable representations, which capture slow-varying features and whose dynamics are guaranteed to be stable. These representations are beneficial as they enable fully probabilistic reconstructions of the high-dimensional system into the future with the help of a suitable coarse-to-fine map and deliver interpretable insights about the structure and dynamics of the system. The coarse-to-fine map is parametrized by deep neural networks and incorporates an intermediate layer of latent variables that can reconstruct the full high-dimensional system with a linear mapping. By applying a sparsity prior at this intermediate layer we make

sure to identify the smallest amount of variables needed to linearly reconstruct the system. We apply our method to small amounts of simulation data of high-dimensional physical systems and are able to characterize the systems by the learned features, to generate extrapolative predictions and to capture the predictive uncertainty due to the information loss because of dimension and model-order reduction.

Sebastian Kaltenbach, Phaedon-Stelios Koutsourelakis
Technical University of Munich
sebastian.kaltenbach@tum.de,
p.s.koutsourelakis@tum.de

MS84

Output-Weighted Sampling for Multi-Armed Bandits with Extreme Payoffs

We present a new type of acquisition functions for online decision making in multi-armed and contextual bandit problems with extreme payoffs. Specifically, we model the payoff function as a Gaussian process and formulate a novel type of upper confidence bound (UCB) acquisition function that guides exploration towards the bandits that are deemed most relevant according to the variability of the observed rewards. This is achieved by computing a tractable likelihood ratio that quantifies the importance of the output relative to the inputs and essentially acts as an attention mechanism that promotes exploration of extreme rewards. We demonstrate the benefits of the proposed methodology across several synthetic benchmarks, as well as two realistic examples involving noisy sensor network data (specifically, temperature and air quality measurements). Finally, we provide a JAX library for efficient bandit optimization using Gaussian processes.

Yibo Yang
University of Pennsylva
ybyang@seas.upenn.edu

MS85

Importance Sampling-Based Reliability Assessment of Dynamic Systems

Computing the failure probability of dynamic systems using direct Monte Carlo simulation is computationally expensive, particularly for low probability values $lessthan 1e-4, forinstance$. We present a computationally efficient, importance sampling-based methodology for reliability assessment of dynamic systems. The proposed approach addresses a general random process inputs, including non-Gaussian, non-stationary random processes, and b combinations of random process and random variable inputs. The proposed approach is presented as a generalization of Girsanov transformation, which is intended for dynamic systems subjected to Gaussian white noise excitation. We demonstrate the proposed methodology using a a numerical illustration on a multi-physics model of a flexible panel subjected to hypersonic flow, and b a real-world implementation on safety assessment of aircraft within a sector using recorded flight data.

Abhinav Subramanian, Sankaran Mahadevan
Vanderbilt University
abhinav.subramanian@vanderbilt.edu,
sankaran.mahadevan@vanderbilt.edu

MS85

The Ensemble Kalman Filter for Rare Event Estimation

We present a novel sampling-based method for estimating probabilities of rare or failure events. Our approach is founded on the Ensemble Kalman filter (EnKF) for inverse problems. Therefore, we reformulate the rare event problem as an inverse problem and apply the EnKF to generate failure samples. To estimate the probability of failure, we use the final EnKF samples to fit a distribution model and apply Importance Sampling with respect to the fitted distribution. To handle multi-modal failure domains, we localise the covariance matrices in the EnKF update step around each particle and fit a mixture distribution model in the Importance Sampling step. For afine linear limit-state functions, we investigate the continuous-time limit and large time properties of the EnKF update. We prove that the mean of the particles converges to a convex combination of the most likely failure point and the mean of the optimal Importance Sampling density if the EnKF is applied without noise. We provide numerical experiments that demonstrate the performance of the EnKF in a variety of problem settings.

Fabian Wagner
TUM
fabian.wagner@ma.tum.de

Iason Papaioannou
Engineering Risk Analysis Group
TU M{ü}nchen
iason.papaioannou@tum.de

Elisabeth Ullmann
TU München
elisabeth.ullmann@tum.de

MS86

Bamcafe: A Bayesian Machine Learning Advanced Forecast Ensemble Method for Complex Turbulent Systems with Partial Observations

Ensemble forecast based on physics-informed models is one of the most widely used forecast algorithms for complex turbulent systems. A major difficulty in such a method is the model error that is ubiquitous in practice. Data-driven machine learning (ML) forecasts can reduce the model error but they often suffer from the partial and noisy observations. In this paper, a simple but effective Bayesian machine learning advanced forecast ensemble (BAMCAFE) method is developed, which combines an available imperfect physics-informed model with data assimilation (DA) to facilitate the ML ensemble forecast. In the BAMCAFE framework, a Bayesian ensemble DA is applied to create the training data of the ML model, which reduces the intrinsic error in the imperfect physics-informed model simulations and provides the training data of the unobserved variables. Then a generalized DA is employed for the initialization of the ML ensemble forecast. In addition to forecasting the optimal point-wise value, the BAMCAFE also provides an effective approach of quantifying the forecast uncertainty utilizing a non-Gaussian probability density function that characterizes the intermittency and extreme events. A two-layer Lorenz 96 model and a nonlinear conceptual model

will be used as numerical illustrations.

Nan Chen, Yingda Li
University of Wisconsin-Madison
chennan@math.wisc.edu, yli678@wisc.edu

MS86

ML-Assisted Resampling for Stochastic Parameterization with Memory

For parameterization of unresolved processes in multiscale dynamical systems, data-based methods relying on machine learning (ML) techniques are rapidly gaining ground. Ususally the ML-based parameterization is deterministic and ignores uncertainty in the feedback from the small-scale (unresolved) to the large-scale processes. By considering stochastic rather than deterministic parameterization, this uncertainty can be taken into account. In this talk I will discuss recent work constructing data-based stochastic parametrizations with memory, using resampling. A straightforward approach to implement resampling is by binning. In case of long memory, resampling by binning is hampered by curse of dimension. To overcome this, a neural network for probabilistic classification can be used in combination with resampling. I will discuss both approaches and show their performance on several test problems.

Daan Crommelin
CWI Amsterdam
Daan.Crommelin@cwi.nl

MS86

Machine Learning Reduced Parameterizations for Turbulent Flow

It is well known that the wide range of spatial and temporal scales present in turbulent flow problems represents a (currently) insurmountable computational bottleneck, which must be circumvented by a coarse-graining procedure. The effect of the unresolved fluid motions enters the coarse-grained equations as an unclosed forcing term. Traditionally, the system is closed by approximate deterministic closure models. Instead, we focus on creating a stochastic, data-driven surrogate model from a (limited) set of reference data. Since the unclosed forcing term is a dynamically evolving field, a surrogate should be able to mimic the complex spatial patterns of this term. Rather than creating such a spatially extended surrogate, we propose to precede the surrogate construction step by a procedure that replaces the unclosed forcing term with a new source term which is tailor-made to capture spatially integrated quantities of interest, and which significantly reduces the amount of training data that is needed. Instead of creating a surrogate model for an evolving field, we now only require a much smaller (stochastic) surrogate model for one scalar time series per quantity-of-interest. We derive the new source terms for a simplified ocean model of two-dimensional turbulence, and we discuss the challenges of training data-driven surrogates models which are coupled to physical models of the (macroscopic) flow quantities.

Wouter Edeling
CWI
wouter.edeling@cwi.nl

MS86

Physics-Constrained Data-Driven Subgrid-Scale Parameterization of 2D Turbulence in the Small-Data Regime

In this work, we develop a data-driven subgrid-scale (SGS) model for large eddy simulation (LES) of 2D turbulence using a fully convolutional neural network (CNN). In the small-data regime, the LES-CNN generates artificial instabilities and thus leads to unphysical results. We propose four remedies for the CNN to work in the small-data regime: (1) data augmentation (DA), (2) group equivariant convolution neural network (GCNN), leveraging the rotational equivariance of the SGS term, (3) incorporating a physical constraint on the SGS enstrophy transfer, and (4) variational autoencoder providing stochasticity and uncertainty quantification. The rotational equivariance of SGS terms can be accounted for by either including rotated snapshots in the training data set (DA) or by a GCNN that enforces rotational equivariance as a hard constraint. Additionally, the SGS enstrophy transfer constraint can be implemented in the loss function of the CNN. Stochasticity can be crucial in modeling backscattering (energy transferred from subgrid scales to resolved scales) of SGS terms. A priori and a posteriori analyses show that the proposed approaches enhance the SGS model and allow the data-driven model to work stably and accurately in a small-data regime.

Yifei Guan, Adam Subel
Rice University
yg62@rice.edu, as170@rice.edu

Ashesh K. Chattopadhyay
Mechanical Engineering
Rice University
akc6@rice.edu

Pedram Hassanzadeh
Mechanical Engineering and Earth Science
Rice University
pedram@rice.edu

MS87

Variational Bayesian Approximation of Inverse Problems using Sparse Precision Matrices

Inverse problems involving partial differential equations are widely used in science and engineering. Although such problems are generally ill-posed, different regularisation approaches have been developed to ameliorate this problem. Among them is the Bayesian formulation, where a prior probability measure is placed on the quantity of interest. The resulting posterior probability measure is usually analytically intractable. The Markov Chain Monte Carlo (MCMC) method has been the go-to method for sampling from those posterior measures. MCMC is computationally infeasible for large-scale problems that arise in engineering practice. Lately, Variational Bayes has been recognised as a more computationally tractable method for Bayesian inference, approximating a Bayesian posterior distribution with a simpler trial distribution by solving an optimisation problem. In this work, we argue, through an empirical assessment, that VB methods are a flexible and efficient alternative to MCMC for this class of problems. We propose a natural choice of a family of Gaussian trial distributions parametrised by precision matrices, thus taking advantage of the inherent sparsity of the inverse problem encoded in its finite element discretisation. We utilise stochastic optimisation to efficiently estimate the variational objective and assess not only the error in the solution mean but also the ability to quantify the uncertainty of the estimate. We

test this on PDEs based on the Poisson equation.

Jan Povala
Department of Mathematics, Imperial College London
jan.povala11@imperial.ac.uk

Ieva Kazlauskaite
University of Cambridge
ik394@cam.ac.uk

Eky Febrianto
Department of Engineering, University of Cambridge
efebrianto@turing.ac.uk

Fehmi Cirak
bDepartment of Engineering, University of Cambridge
f.cirak@eng.cam.ac.uk

Mark Girolami, Carl Henrik Ek
University of Cambridge
mag92@cam.ac.uk, che29@cam.ac.uk

MS87

Physically-Inspired Gaussian Processes with Application to Biology

Physically-inspired Gaussian processes (GPs) provide a flexible stochastic framework where linear differential equations are encoded into covariance functions (kernels). As data-driven approaches, they can be established without specifying all the physical interactions from mechanistic processes. By enforcing GPs with physical knowledge, accurate predictions are provided even in regions where data are not available. In this talk, we focus on GPs physically inspired by a reaction-diffusion model where both the decay and diffusion rate constants are encoded as parameters of the kernels. Two types of GP-based models are studied where the main difference lies in where the prior is placed. On a biological application describing the post-transcriptional regulation of Drosophila, we demonstrate the capability and versatility of the models to capture the dynamics of spatio-temporal interactions between mRNAs and gap proteins.

Andrés Lopez
Université Polytechnique Hauts-de-France (UPHF)
anfelopera@utp.edu.co

Mauricio Alverez
University of Sheffield
mauricio.alvarez@sheffield.ac.uk

Durrande Nicolas
SecondMind
durrande@gmail.com

MS87

End to End GP-Based Inversion of a Mass Density Field from Gravimetric Measurements

The use of GP priors in Bayesian inverse problems is well established but resulting implementations are seldom discussed. We provide a detailed description of the whole inversion process on a test case from volcano geophysics, demonstrating in turn how to overcome various practical problems. The goal in our motivating inverse problem is to reconstruct the mass density field inside a bounded region from measurements of the gravitational field on the outside, i.e. on the surface of the Stromboli volcano in our test case. Those measurements may be viewed as linear forms of the field, a neat framework for GP modelling. Also, considering the GP induced on the surface by the prior mass density field and the forward operator turns out to be fruitful in several respects. We show how the proposed approach enables enjoying MLE for hyper-parameters such as classically used in GP modelling for computer experiments, yet at the cost of numerical instabilities due to the specific form for the covariance matrix. Using chunking, automatic differentiation and GPU-based libraries, we demonstrate how GP-based inversion scales to grids of several hundreds of thousands of cells. We also appeal to fast k-fold cross-validation and explore how associated outputs can help balance numerical instabilities. We finally highlight how this approach can be used to guide new measurements towards an efficient reconstruction of the mass density field.

Cédric Travelletti
University of Bern
cedric.travelletti@stat.unibe.ch

Niklas Linde
University of Lausanne
niklas.linde@unil.ch

David Ginsbourger
University of Bern
david.ginsbourger@stat.unibe.ch

MS87

Kernel-based Statistical Methods for Functional Data

Kernel-based statistical algorithms have found wide success in statistical machine learning in the past ten years as a non-parametric, easily computable engine for reasoning with probability measures. The main idea is to use a kernel to facilitate a mapping of probability measures, the objects of interest, into well-behaved spaces where calculations can be carried out. This methodology has found wide application, for example two-sample testing, independence testing, goodness-of-fit testing, parameter inference and MCMC thinning. Most theoretical investigations and practical applications have focused on Euclidean data. This talk will outline work that adapts the kernel-based methodology to data in an arbitrary Hilbert space which then opens the door to applications for functional data, where a single data sample is a discretely observed function, for example time series or random surfaces. Such data is becoming increasingly more prominent within the statistical community and in machine learning.

George Wynne
Imperial College London
g.wynne18@imperial.ac.uk

MS88

UQ on Point-Like Emission Source Estimates by Satellite Data

We propose a statistical approach to estimate emissions of point-like sources from satellite observations. We employ a 'gray-box' approach where data-driven estimation is combined with rudimentary physics and chemistry. The cross-sectional flux method is used, where the 2d plume data of a single emission source is integrated orthogonal to the wind direction. We formulate the outcome as a

1D convection-reaction system. The approach consists of several steps: aligning the coordinates to the main wind direction, constructing smooth representations of the satellite observations using Gaussian random field (GRF) modelling, performing the cross-sectional integrations using the obtained continuous models, and finally estimating the parameters of the convection-reaction system. The point-emission flux is the parameter of interest, while the reaction terms are considered as nuisance parameters. At every step the uncertainties - the background concentrations, the wind speed and direction, the cross-sectional integrals by the GRF models, are sampled by MCMC methods. The approach is verified by synthetic data, and applied to plumes measured by instruments such as the Tropospheric Monitoring Instrument (TROPOMI). The example cases include large power plants in Belchatw, Poland and Wuhan, China, as well as moving cargo ships. The results are compared against available in-situ data.

Heikki Haario
Lappeenranta University of Technology
Department of Mathematics and Physics
heikki.haario@lut.fi

Teemu Härkönen
LUT University
teemu.harkonen@lut.fi

Anu-Maija Sundström
Finnish Meteorological Insitute
anu-maija.sundstrom@fmi.fi

MS88

Simulation-Based Uncertainty Quantification for Infrared Sounder Atmospheric Retrievals

An ever-changing constellation of Earth-orbiting satellites continues to provide large volumes of data with comprehensive spatial and temporal coverage. These remote-sensing observations provide indirect information on numerous surface and atmospheric quantities of interest, and the operational data processing pipeline typically involves multiple levels of data products that are used in scientific inference for the Earth system. A critical step involves inferring surface and atmosphere states from satellite spectra through an inverse method known as a retrieval. We will present an overview of the remote sensing observing system and retrieval for a class of instruments known as hyperspectral infrared sounders. The sounder observation record spans multiple decades and includes the Atmospheric Infrared Sounder (AIRS) and Cross-track Infrared Sounder (CrIS). We illustrate a simulation-based framework for uncertainty quantification (UQ) for the AIRS retrieval of atmospheric temperature and humidity. The approach includes a flexible statistical model for the conditional distribution of application-relevant quantities of interest given the retrieval that is trained via simulation and applied to operational AIRS products. Validation of results for near-surface temperature over the continental United States will be highlighted.

Jonathan Hobbs
Jet Propulsion Laboratory
jonathan.m.hobbs@jpl.nasa.gov

Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology
Amy.Braverman@jpl.nasa.gov

Berlin Chen
University of Cambridge
berlinchen7@gmail.com

Eric Fetzer, Kyo Lee
Jet Propulsion Laboratory
eric.j.fetzer@jpl.nasa.gov, huikyo.lee@jpl.nasa.gov

Hai Nguyen
Jet Propulsion Laboratory
California Institute of Technology
hai.nguyen@jpl.nasa.gov

Joaquim Teixeira
Jet Propulsion Laboratory
joaquim.p.teixeira@jpl.nasa.gov

MS88

Statistical Data Fusion for Multispectral Surface Reflectance Products

Daily, high spatial resolution (¡100m) visible to shortwave infrared surface reflectance datasets are needed to estimate and monitor geophysical processes that exhibit rapid changes over space and time, e.g., agricultural harvest, deforestation, and disease. However, no single, non-commercial space mission currently provides such data due to limitations of the spatial, temporal and spectral resolutions of individual instruments. In this talk, we propose a flexible, spatiotemporal data fusion methodology to combine multispectral surface reflectance measurements from three remote sensing instruments (Suomi NPP VIIRS, Sentinel 2A/B and Landsat 8) to produce daily, 70m resolution products with associated uncertainty estimates. The methods rely on space-time dynamic linear models to incorporate spatial change-of-support and to leverage spatial and temporal dependence for gap-filling between high resolution images. A computationally scalable, moving-window Kalman filter is developed to facilitate integration into downstream science processing algorithms.

Margaret Johnson, Gregory Halverson
Jet Propulsion Laboratory
maggie.johnson@jpl.nasa.gov,
gregory.h.halverson@jpl.nasa.gov

Kerry Cawse-Nicholson
Jet Propulsion Laboratory
California Institute of Technology
kerry-anne.cawse-nicholson@jpl.nasa.gov

Joshua Fisher
University of California, Los Angeles
joshbfisher@gmail.com

Jouni Susiluoto, Glynn Hulley
Jet Propulsion Laboratory
jouni.i.susiluoto@jpl.nasa.gov, glynn.hulley@jpl.nasa.gov

MS88

Large Scale Data Fusion of Remote Sensing and in Situ Observations by Dimension Reduction

To efficiently utilize available environmental observations from various sources such as Earth observing satellites and in-situ measuring instruments, we need spatio temporal data fusion methods. Kalman smoother techniques provide practical tools to do retrospective multi-dimensional time

series analysis by high dimensional dynamical linear models. Proper uncertainty quantification can be done with Bayesian statistics and by hierarchical state space description of the data, the processes, and the parameters defining the data fusion system. This allows us to carefully consider the representativeness and uncertainties of data and models as well as the natural variability of the phenomena of interest in different spatial and temporal scales. This talk will demonstrate a data fusion system that utilizes different dimension reduction techniques to tackle large modelling domains and large amounts of data. The system is used to analyze water quality in the Baltic Sea using.

Marko Laine
Finnish Meteorological Institute
marko.laine@fmi.fi

MS89

Rovering Dynamical Systems from Noisy Measurements Using Constrained Optimization

Even if we know the physical laws that govern a dynamical system, we may seek to infer a simplified set of dynamics from measurements. Here we build a nonlinear dynamical system by identifying the coefficients attached to each term in a sum of nonlinear functions of the state. Many standard approaches like SINDy identify these coefficients by minimizing the least squares mismatch in this differential equation at each time point. Unfortunately this amplifies the effect of noise emerging from both derivative estimation and the nonlinear basis functions yielding suboptimal estimates of the dynamical system. Here we correctly accounts for the noise, yielding a maximum likelihood estimate for the dynamical system. Our approach introduces a variable for the noiseless state, adds an equality constraint corresponding to the dynamical system, and then minimizes the least squares mismatch between the noiseless state estimate and the measured state. We efficiently solve this large scale optimization problem using sequential quadratic programming on an augmented Lagrangian formulation with a specialized preconditioner. Regularization, such as promoting sparsity in coefficients using an 1-norm, can easily be included in this formulation using iteratively reweighted least squares. As a result of our proper accounting of noise, we obtain models that are two to ten times more accurate than standard approaches.

Jeffrey M. Hokanson
University of Colorado Boulder
jeffrey.hokanson@colorado.edu

Alireza Doostan
Department of Aerospace Engineering Sciences
University of Colorado, Boulder
Alireza.Doostan@Colorado.EDU

MS89

A Framework for Variational Embedding of Data in Physics-Based Models

This paper presents a physics-based data embedding framework that weakly enforces the balance laws via variational multiscale approach. The physics informed method is augmented with high-fidelity sparse data through fine-scale variational embedding of the variationally derived loss function. The structure of the loss function is discussed in the context of variational correction to the coarse-scale model that weakly enforces physical constraints, wherein the loss function penalizes the difference of the response of the physics-based model from the prescribed data that represents the local behavior of the system. In the context of forward simulations, the proposed approach can be seen as inducing inductive biases that exploits the difference between the computed and measured quantities in the parametric space. With the help of a model problem, we show that the proposed method learns from the sparse data sets while preserving the balance laws. This framework results in physics-informed data-driven technique, and at the same time, has an impact on forward simulations that are driven by boundary and initial conditions. Method is applied to a mathematically non-smooth problem and the attributes of the formulation are investigated in the light of the high-fidelity sparse data sets.

Arif Masud
University of Illinois at Urbana-Champaign
amasud@uic.edu

Shoaib Goraya
University of Illinois Urbana-Champaign
Department of Civil and Environmental Engineering
sgoraya2@illinois.edu

MS89

Neural Implicit Flow: a Representation Learning Paradigm beyond POD and CNN

Most of the physical science and engineering problems exhibit complexities arise from non-linear partial differential equations (PDE) in three folds: multi-scale in space, chaos in time, and bifurcation in parameters. Among these challenges, spatial complexity is the major complexity of fluid dynamics, which motivates the need of dimensionality reduction. Existing paradigms, e.g., POD and CNN, both struggle to accurately and efficiently represent flow structures for problems requiring variable geometry, non-uniform grid resolution (e.g., wall-bounded flows, flow phenomenon induced by small geometry features), adaptive mesh refinement, or parameter-dependent meshes. To resolve these difficulties, we propose a general framework called Neural Implicit Flow (NIF) that enables a compact and flexible dimension reduction of large-scale, parametric, spatial-temporal data into mesh-agnostic fixed-length representations. This work complements existing mesh-less methods, e.g., physics-informed neural networks, and we focus specifically on obtaining effective reduced coordinates where modeling and control tasks may be performed more efficiently. In several challenging examples, we demonstrate the utility of NIF for data-fit parametric surrogate modeling, mesh-agnostic interpretability of spatio-temporal fields, efficiency for many-spatial-query, improved performance for sparse reconstruction, and compressed representation of turbulence.

Shaowu Pan, Steve Brunton
University of Washington
shawnpan@uw.edu, sbrunton@uw.edu

J. Nathan Kutz
University of Washington, Seattle
Dept of Applied Mathematics
kutz@uw.edu

MS89

PDE-Preserved Network Architecture for Predicting Spatiotemporal Dynamics

The recent development of physics-informed deep learn-

ing (PiDL) that leverages domain knowledge and physical prior has shown promise in solving physical problems in many science and engineering fields. In most existing PiDL works (e.g., physics-informed network, or PINN), physics prior knowledge is mainly utilized to inform or constrain the training process of the network by incorporating known governing equations into the loss function in a soft manner. However, the mathematical understanding of the system has not been fully leveraged, for example, in network architecture design. In this work, we develop a novel physics-informed deep learning architecture for predicting spatiotemporal dynamics. The network construction, layer connections, and output constraints are based on the discretized structure of the governing partial differential equations (PDEs). The learned model can be used as a parametric surrogate model for the fast prediction of spatiotemporal dynamics, facilitating UQ analysis in high-dimensional complex systems.

Jian-Xun Wang, Xin-Yang Liu, Han Gao
University of Notre Dame
jwang33@nd.edu, xliu28@nd.edu, hgao1@nd.edu

MS90

Prediction of Failure Locations in Porous Metals Using a Bayesian Convolutional Neural Network

We consider the problem of predicting locations where mechanical failure occurs in porous metal specimens under tension. Porosity is a feature of additively manufactured materials and determines failure locations through nonlinear mechanics that exhibits sensitivity to the initial pore locations. While traditional viscoplastic damage models provide an accurate model of the evolution of damage in porous specimens, they are computationally expensive. In this work, we study the use of convolutional neural networks as surrogate models for predicting failure locations. The binary classification problem of categorizing failed voxels is first regularized by recasting it as a regression problem for the continuous damage field subjected to pre-processing transformations. The damage fields display a relatively small number of voxels close to failure leading to a form of class imbalance for regression that can cause the optimizer to converge to a poor local minimum. We address this through a re-weighting of the loss function which accounts for the relative frequencies of damage values. Furthermore, a sensitivity analysis of the viscoplastic model is carried out and indicates that there are multiple regions of high damage competing for failure. This motivates the use of Bayesian neural networks to capture sensitivities in the prediction through uncertainty quantification.

Reese Jones
Co-author
rjones@sandia.gov

Krishna Garikipati
Mechanical Engineering
University of Michigan
krishna@umich.edu

Gregory Teichert
University of Michigan
greght@umich.edu

Mohammad Khalil
Sandia National Laboratories
mkhalil@sandia.gov

Xiaoxuan Zhang
University of Michigan
xxzh@umich.edu

Wyatt H. Bridgman
Sandia National Laboratories
whbridg@sandia.gov

MS90

Probabilistic Neural Network Surrogates for Reduced-Order Dynamics in Reacting Flows

Many dynamical systems of physical interest are characterized by high-dimensional state spaces. For conservation laws governing complex fluid systems, e.g. chemically reacting flows and plasmas, often many hundreds or thousands of components are required to represent the thermochemical state and capture detailed kinetic processes. Solving such conservation equations numerically thus involves considerable expense for detailed simulations. In this talk we describe approaches to mitigate this expense by learning low-dimensional manifolds (e.g. using principle component analysis) to define reduced bases for representing the thermochemical state, where the transformed state can then be evolved using the original governing system of partial differential equations. The use of neural network surrogate models to accelerate source term evaluation or integration, stiffness mitigation, and uncertainty quantification within the reduced construction will be discussed in the context of model problems relevant to gas phase combustion and low-temperature plasmas.

Tiernan Casey
Sandia National Laboratories
tcasey@sandia.gov

Simone Venturi
Sandia NAtional Laboratories
sventur@sandia.gov

MS90

Bayesian Calibration of Interatomic Potential Models for Binary Alloys

Developing reliable interatomic potential models with quantified predictive accuracy is crucial for the successful application of atomistic simulation. Commonly used potentials, such as those constructed through the embedded atom method (EAM), are typically derived from semi-empirical considerations and contain unknown parameters that must be properly fit before use. This fitting is often performed by comparing to training data generated with a more expensive first principles physics model, e.g., based on density functional theory (DFT). In this presentation, we explore Bayesian calibration as a means of fitting EAM potentials for binary alloys. Here, probabilistic assertions about the model parameters, model error, and other uncertain quantities are updated by including a training set of DFT-simulated data. The outcome is a posterior distribution over the aforementioned items, which can then be carried forward to quantify predictive uncertainties and assess the quality of fit. We apply these techniques to investigate an EAM potential for a family of gold-copper alloys in which the training data consists of physical properties such as lattice parameters, mixing enthalpies, and elastic constants. We conclude by discussing the models predictive performance on interatomic forces, a quantity

not included in the training set.

Arun S. Hegde
Sandia National Laboratories
ahegde@sandia.gov

Elan Weiss, Wolfgang Windl
The Ohio State University
weiss.443@buckeyemail.osu.edu, windl.1@osu.edu

Habib N. Najm
Sandia National Laboratories
Livermore, CA, USA
hnnajm@sandia.gov

Cosmin Safta
Sandia National Laboratories
csafta@sandia.gov

MS90

Use of a Machine Learning Model for a Constitutive Chemistry Model within a Groundwater Flow and Transport Application Modeling Nuclear Fuel Degradation in a Waste Repository

Physics-informed surrogate models are often constructed to represent a computational simulation model for an application of interest. In this work, we extend the surrogate framework further: we replace one component of a large, multiphysics model of a subsurface nuclear waste repository with a machine-learned surrogate. The component for which the ML surrogate is constructed is the Fuel Matrix Degradation (FMD) constitutive chemistry model, which represents UO2 spent fuel degradation rates as a function of radiolysis, redox reactions, and other electrochemical reactions. This is a major advance in fidelity relative to the previously-employed fractional dissolution rate model, which is computationally more tractable but less accurate. The prohibitive cost of the FMD model for full-system simulations motivates its replacement with a machine-learned surrogate which can handle the varying inputs to the FMD model. We examine the effect the surrogate FMD model has on uncertainty analyses for a realistic repository application which also considers the treatment of parameter uncertainty and spatial heterogeneity in the subsurface. We summarize challenges with the use of a surrogate model as a constitutive model within large, multiphysics applications.

Laura Swiler
Sandia National Laboratories
Albuquerque, New Mexico 87185
lpswile@sandia.gov

Teresa Portone
Sandia National Laboratories
tporton@sandia.gov

Paul Mariner
Sandia National Laboratory
pmarine@sandia.gov

Rosie Leone, Dusty Brooks, Daniel T. Seidl
Sandia National Laboratories
rleone@sandia.gov, dbrooks@sandia.gov,
dtseidl@sandia.gov

Bert J. Debusschere
Chemistry, Combustion and Materials Center
Sandia National Laboratories, Livermore CA
bjdebus@sandia.gov

Timothy M. Berg
Sandia National Laboratory
tberg@sandia.gov

MS91

Uncertainty Quantification of Tsunami Currents and Heights

Realistic tsunami hazard assessments require a large number of predictions, typically hundreds of thousands, for comprehensively sweeping through the earthquake magnitude-frequency distribution. This enables confident probabilistic quantification of hazard, especially for high-impact low-frequency events. Further, actionable hazard assessments utilize area-wide regional hazard maps which require hundreds of thousands of emulators, each emulator corresponding to a point in the region of interest. We build workflows to individually address these large-scale challenges one million predictions showcased for the port of Karachi, and hundreds of thousands of emulators for Vancouver. The parameters of the tsunami source act as inputs for the emulator. The computer code simulates the coupled numerical models of deformation of sea floor due to the earthquake source and the consequent propagation of the tsunami. Multiple emulators are constructed to approximate the functional dependence of the hazard intensities, viz. wave heights and velocities, on the earthquake source parameters. Apart from obtaining probabilistic quantification for the hazard, we observe high velocities, and spatial patterns that correspond to meaningful geophysical behaviour in the harbour. We conclude that statistical emulators underpinned by high resolution numerical simulations and realistic bathymetry data show promise for large-scale probabilistic tsunami hazard assessment.

Devaraj Gopinathan
Department of Statistical Science
University College London
d.gopinathan@ucl.ac.uk

Serge Guillas
University College London
s.guillas@ucl.ac.uk

MS91

Multiscale Emulation of Tsunami Waves

Building an emulator of a computer model, using a small design of experiments, greatly alleviates the computational burden to carry out uncertainty quantification. We introduce a design strategy that allows us to efficiently allocate limited computational resources over simulations of different levels of fidelity combined with a sequential design at each level. The multilevel adaptive sequential design of computer experiments (MLASCE) makes use of reproducing kernel Hilbert spaces as a new tool for our GP approximations of the increments across levels. We theoretically prove the validity of our approach in some settings, and compare with existing models of multi-fidelity Gaussian process emulation. Compared to a single level approach, gains of orders of magnitudes in accuracy for medium-size computing budgets are demonstrated in numerical examples. We provide an illustration to tsunami hazard assessment for the city of Cilacap in Indonesia. By running a few tsunami simulations at high resolution and many more simulations at lower resolutions we provide realistic assess-

ments whereas, for the same budget, using only the high resolution tsunami simulations is not satisfactory. Hence, using our multi-level method, tsunami hazard assessments can be achieved with higher precision using the highest spatial resolutions, and for impacts over larger regions.

Serge Guillas
University College London
s.guillas@ucl.ac.uk

MS91

A Combined Physical-Statistical Approach for Estimating Storm Surge Risk

Storm surge is an abnormal rise of seawater caused by a storm. It poses the most severe threat to property and life in a coastal region. Thus, it is crucially important to assess the storm surge risk, typically summarized by r-year surge return level with return period r ranging from 10, 50, 100, or even much longer along a coastline. However, it is challenging to reliably estimate this quantity due to the limited storm surge observations in space and time. This talk presents an approach to integrate physical and statistical models to estimate extreme storm surge. Specifically, A physically-based hydrodynamics model is used to provide the needed interpolation in space and extrapolation in both time and atmospheric conditions. Statistical modeling is needed to 1) estimate the input distribution for running the computer model, 2) develop a statistical emulator in place of the computer simulator, and 3) estimate uncertainty due to input distribution, statistical emulator, missing/unresolved physics.

Whitney K. Huang
Clemson University
wkhuang@clemson.edu

Emilty Tidwell
Dynetics
emilytidwell99@gmail.com

MS91

Uncertainty Quantification in Assessing Storm Surge

Storm surge is one of the most severe natural hazards that can lead to significant flooding in coastal areas and severe damages to the life and property from a hurricane. To assess storm surge hazards, current coastal flood hazard studies are often performed through a synthesis of computer modeling, statistical modeling, and extreme-event probability computation. Since post-Katrina coastal flood hazard studies, a technique called Joint Probability Method (JPM) with its improvements has become the gold standard to compute annual exceedance probability (AEP) levels at certain frequencies by federal agencies such as Federal Emergency Management Agency, private sectors, and academic researchers in coastal engineering. However, the JPM suffers several disadvantages including excessive usage of computing resources, inappropriate uncertainty quantification, and lack of optimal statistical modeling, which make the JPM based coastal flood hazard studies prohibitively costly and unrealistic. To address these issues, we employ a new risk assessment framework to assess storm surges hazards in Southwest Florida (SWFL), where we develop an emulator - a fast approximation to the surge prediction model, based on which, we use an efficient sampling technique to enable fast computation of AEP. Our methodology is capable of handling massive number spatial locations efficiently and provides rigorous uncertainty quantification for risk assessment of storm surges in SWFL.

Pulong Ma
Clemson University
plma@clemson.edu

MS92

Estimation of Global Reliability-Oriented Sensitivity Indices under Epistemic Uncertainty

Uncertainty quantification (UQ) is a global methodology relying on a probabilistic modeling of uncertain input variables of a computer model. Usually, the joint input probability distribution is supposed to be known and is set so as to represent the intrinsic stochastic behavior and the dependence of the inputs. However, if the type of the marginal distributions can be assessed by rational considerations (e.g., based on the best available statistical information or according to standards or expert judgment), the distribution parameters are often tainted with a residual statistical uncertainty. In the typical context of reliability assessment, this second uncertainty level does affect the robustness of the numerical results obtained in the rest of the UQ study and has to be taken into account in the failure probability estimation. This robustness can be investigated by performing reliability-oriented sensitivity analysis. The present work aims at presenting a set of dedicated reliability-oriented Sobol indices taking the bi-level input uncertainty into account. The separation between aleatory (irreducible) and epistemic (reducible) uncertainties is proposed via a disaggregated version of the input random variables. An efficient estimation strategy based on a splitting algorithm and an adapted kernel density estimation is proposed. The methodology is applied on a representative test-case about safety evaluation of a flood protection dike.

Vincent Chabridon
EDF R&D
vincent.chabridon@edf.fr

MS92

Incorporating Distribution-Form and Parameter Uncertainty into Simulation-Based Reliability Analysis

How confident can we be in reliability estimates? We try to answer this question by quantifying the uncertainty in probabilities of failure that stem from lack of sufficient data to precisely identify probability distributions for the model input parameters. We propose an imprecise Subset simulation (SuS) method that utilizes Bayesian/information theoretic multi-model inference to estimate probabilities of probability of failures. Although SuS is very efficient when the probability models for input random variables are precisely defined, the method cannot be directly applied when distributions for the random variables are uncertain. The proposed work consists of two steps: First, Bayesian/information multi-model inference is used to identify model probabilities for each candidate distribution and their associated joint parameter probability densities. From inference, we obtain a set of candidate distributions, all of which are equally probable of representing the limited data. Then, SuS is performed using an optimal sampling density analytically and the resulting conditional probabilities are re-weighted according to each probability distribution in the set using importance sam-

pling. It is shown that the uncertainty in probability of failure estimates may be very large, especially when data sets are small. The outcome of the proposed method is an empirical probability distribution of failure probabilities that allows us to assess confidence in reliability estimates.

Dimitrios Giovanis, Michael D. Shields
Johns Hopkins University
dgiovan1@jhu.edu, michael.shields@jhu.edu

MS92

Enabling and Interpreting Hyper-Differential Sensitivity Analysis for Bayesian Inverse Problems

Inverse problems constrained by partial differential equations (PDEs) play a critical role in model development and calibration. In many applications, there are multiple uncertain parameters in a model which must be estimated. Computational cost and high dimensionality frequently prohibit a thorough exploration of the parametric uncertainty. A common approach is to reduce the dimension by fixing some parameters (called auxiliary parameters) to a best estimate and using techniques from PDE-constrained optimization to approximate properties of the Bayesian posterior distribution. For instance, the maximum a posteriori probability (MAP) and the Laplace approximation of the posterior covariance can be computed. We propose using hyper-differential sensitivity analysis (HDSA) to assess the sensitivity of the MAP point to changes in the auxiliary parameters and establish an interpretation of HDSA as correlations in the posterior distribution. Foundational assumptions require satisfaction of the optimality conditions which are not always feasible as a result of ill-posedness in the inverse problem. We introduce novel theoretical and computational approaches to justify and enable HDSA for ill-posed inverse problems by projecting the sensitivities on likelihood informed subspaces and defining a posteriori updates. Our proposed framework is demonstrated on an infinite dimensional nonlinear multi-physics inverse problem.

Joseph Lee Hart
Sandia National Laboratories
joshart@sandia.gov

MS92

Optimal Uncertainty Quantification of a Risk Measurement on Moment Class

In uncertainty quantification study, we model the uncertain input parameters as random variables. The choice of the probability distributions usually come from expert judgment and/or statistical inference. Therefore, their establishment lack accuracy and result in a second level uncertainty. In this work we gain robustness on the quantification of a risk measurement by accounting for all sources of uncertainties tainting the inputs of a computer code. To that extent, we evaluate bounds on the quantity of interest over a class of bounded distributions satisfying constraints on their moments, called moment class. The problem can be reformulated as the optimization of a quantity of interest over a compact convex set of probability measures. This set is infinite dimensional and nonparametric, so that the optimization of the quantity of interest is generally computationally intractable. However, when the quantity of interest is a lower semi-continuous and quasi-convex functional (for instance a probability of failure, a quantile, a moment, a Sobol index ...), the optimum can be computed only on the extreme points of the set of probability measure. We identify a well-suited parameterization of the extreme points of the moment class based on the theory of canonical moments. It allows an effective, free of constraints, optimization of the quantity of interest. This methodology is applied on a representative test-case about safety evaluation of a flood protection dike.

Merlin Keller
EDF R&D
merlin.keller@edf.fr

Jérôme Stenger
Université Toulouse
jer.stenger@gmail.com

Fabrice Gamboa
Institut de Mathématiques de Toulouse. AOC Project
University of Toulouse
fabrice.gamboa@math.univ-toulouse.fr

Bertrand Iooss
EDF R&D
bertrand.iooss@edf.fr

MS93

Adapting Verification and Validation Principles to a Credibility Process for Scientific Machine Learning

This presentation will provide a summary of the current perspectives, opportunities, and gaps in adapting verification and validation (V&V) principles developed for computational simulation (CompSim) models to a credibility process for scientific machine learning (SciML). The discussion will be framed in the context of maturity readiness levels that allow us to communicate the believability of a model with the intended use as the key source of information for high consequence decision-making. While this approach is readily available for CompSim models, the extendibility and limitations in application to SciML models used in lieu of, complementary to, or as surrogates for CompSim will be presented. To motivate the discussion, examples will be provided that highlight uncertainty quantification (UQ) as one of the key elements to credibility. Noting that the validity of any ML model is predicated on the quality of the data used to train it, emphasis will be made on how the data was sourced and its relationship to the SciML training objectives. The presentation will close with a meta proposal for a SciML V&V/UQ framework and the understanding that there are more research opportunities in this area than well-defined methods.

Erin Acquesta
Sandia National Laboratories
eacques@sandia.gov

MS93

Optimizing Machine Learning Decisions with Prediction Uncertainty

Proper scoring rules are used in decision analysis to elicit probability estimates that reflect an individuals beliefs. In this talk, we consider a decision framework in which the probability estimates come from a statistical or machine learning model. We regard the probability estimates themselves as random variables with respect to the underlying distribution of the data set from which the estimates were derived, and we show that proper loss functions (the neg-

ative of a proper scoring rule) can be used to quantify the variability of the estimator via a generalized bias-variance decomposition. Furthermore, we derive a decision-tailored proper loss function, and we show that the generalized variance of the estimator under this loss quantifies the uncertainty in the estimator in a way that is relevant to the decision problem. In particular, points with low decision-tailored generalized variance correspond to points whose optimal decisions are robust with respect to the distribution of the probability estimator at that point, whereas high decision-tailored generalized variance corresponds to higher decision disagreement under the distribution of the probability estimator.

J. Eric Bickel
University of Texas, Austin
ebickel@mail.utexas.edu

Jason W. Boada, Zachary Smith
UT Austin
jwboada@utexas.edu, zack.smith@utexas.edu

Michael Darling, Justin Doak, Richard Field, Mark Smith
Sandia National Laboratories
mcdarli@sandia.gov, jedoak@sandia.gov,
rvfield@sandia.gov, masmit@sandia.gov

David Stracuzzi
Sand
djstrac@sandia.gov

MS93

Learning Not to Answer: Training a Multi-Task Abstaining Classifier Using PID Control

Deep neural networks (DNN) often provide the only way to build machine capable of making decisions in the face of real-world complexity. In many situations, these decisions are not actionable unless they reach very high certainty. Often, in these cases, the input samples are a mixture of 'easy' instances of the problem where the DNN performance is acceptable, and 'harder' instances where the training turns out to be inadequate. Moreover, the uncertainty is patterned—there are often features that correlate not with the response variable, but with higher prediction errors. We have previously shown that abstaining classifiers can be used to learn such features effectively, but have a tunable parameter that controls the penalty of incorrect answers over abstention. Real-world requirements, however, are stated in terms of a balance between a desired maximum rate of abstention and the accuracy required on the non-abstained set; this would usually need expensive hyperparameter tuning to achieve. Here we show that we can borrow ideas from control theory to tune the abstention penalty parameter using a proportionalintegralderivative (PID) controller during training to reach the desired goal smoothly. We demonstrate the method on a multi-task NLP problem, cancer report classification, simultaneously tuning several tasks with different levels of inherent uncertainty.

Jamaludin Mohd-Yusof
CCS-2, Computational Physics and Methods
Los Alamos National Laboratory
jamal@lanl.gov

Sunil Thulasidasan, Sayera Dhaubhadel, Cristina Garcia Cardona
Los Alamos National Laboratory
sunil@lanl.gov, sayeradbl@lanl.gov, cgarciac@lanl.gov

Tanmoy Bhattacharya
Los Alamos National Lab
tanmoy@lanl.gov

MS94

When Abstention Is Really the Best Policy

Deep Neural Networks often provide the only practical means of training a system that can deal with complexities of the real world. In many decision tasks, however, the high cost of incorrect responses makes it advantageous to design systems that choose to abstain when the uncertainty is large. In multi-task settings, this cost may not be additivethe figure of merit might be the simultaneous case-level correctness of the predictions of all the tasks. We previously developed an abstaining classifier in such a multitask setting that can triage the low-uncertainty cases at both task- and case-level. A benefit to the use of such abstaining classifiers is that they can be coupled with standard methods like Local Interpretable Model-agnostic Explanations (LIME) to look not only for features that correlate with classification, but also those that explain the abstentions. Here, we will present results of our studies using LIME in this setting to compare the correlates of case-level abstention with those of task-level abstentions. Together, the techniques we developed, provide a powerful tool to partially automate real world problems.

Sayera Dhaubhadel
Los Alamos National Laboratory
sayeradbl@lanl.gov

Jamaludin Mohd-Yusof
CCS-2, Computational Physics and Methods
Los Alamos National Laboratory
jamal@lanl.gov

Sunil Thulasidasan, Kumkum Ganguly, Cristina Garcia Cardona, Benjamin McMahon
Los Alamos National Laboratory
sunil@lanl.gov, kumkum.ganguly@lanl.gov,
cgarciac@lanl.gov, mcmahon@lanl.gov

Tanmoy Bhattacharya
Los Alamos National Lab
tanmoy@lanl.gov

MS94

Partially Bayesian Neural Networks: Low-Cost Bayesian Uncertainty Quantification for Deep Learning in Medical Image Segmentation

Radiologists identify and segment tumors based on their appearance in medical images and indicate their confidence about their segmentation. While deep learning (DL) algorithms greatly accelerate the segmentation process, they lack transparency in communicating the uncertainty in the model and the results. Traditional Bayesian Uncertainty Quantification (UQ) methods such as MCMC are prohibitively costly for large, million-dimensional deep neural networks (DNN) used in medical image segmentation. In this talk, we discuss a computationally efficient approach for computing uncertainty of automated DL-based tumor segmentation, via the partially Bayesian neural networks (pBNN). In pBNN, a single strategically chosen layer is used for targeted Bayesian inference while the rest of the

network can be trained using less-expensive deterministic methods. Sensitivity Analysis is employed to guide the selection of the layer for Bayesian inference. We illustrate the benefits of our method, including computational efficiency, and how practitioners and model developers can use this approach to understand the models uncertainty with lowered computational resources.

Snehal Prabhudesai, Jeremiah Hauth, Dingkun Guo, Arvind Rao, Nikola Banovic, Xun Huan
University of Michigan
snehalbp@umich.edu, hauthj@umich.edu, guodk@umich.edu, ukarvind@med.umich.edu, nbanovic@umich.edu, xhuan@umich.edu

MS94

Uncertainty Quantification in Machine Learning-Driven Decision Making

We propose to incorporate uncertainty and performance into AI-advised human decision making to boost user trust in AI. Motivated by the trial-to-trial approach in the adjustment of decision boundaries by humans in decision making, this paper presents a novel framework of human-AI teaming for trustworthy AI under uncertainty. The framework employs the trial-to-trial approach to simulate the trial to trial process for adjusting decision boundaries and provide an estimation of uncertainty in AI. An Uncertainty-Performance Interface (UPI) in the framework is then proposed to allow users access the quality and performance of machine learning models in AI at the same time. The proposed framework allows human and AI collaborate in a teaming environment for trustworthy decisions under uncertainty.

Jianlong Zhou
University of Technology Sydney
jianlong.zhou@uts.edu.au

MS96

Drawing Independent Samples is the Key to a Fast Bayesian Sampler

Bayesian sampling algorithms are crucial to providing uncertainty on model parameters inferred from data and prior assumptions. Because the space of models compatible with data and prior assumptions is not known a priori and many samples may need to be drawn to estimate model uncertainty, Bayesian sampling can be costly in terms of both computational resources and total run time. This is especially true if samples must be drawn in sequence (a serial algorithm) and/or are highly correlated with one another. Here we show that the key to a fast Bayesian sampler is to draw independent samples, enabling parallelization of the algorithm and requiring far fewer samples to estimate uncertainty. We will show how randomized regularized inversion can draw independent samples from the Bayesian posterior, using regularization as prior. We will also show how this prior can be made flexible by regularizing using covariance matrices instead of derivative smoothing. The parameters of covariance regularizationa length scale and a kernelcan be selected hierarchically by the data during the inversion.

Daniel Blatter
Scripps Institution of Oceanography
University of California, San Diego
dblatter@ucsd.edu

Matthias Morzfeld
University of California, San Diego
Scripps Institution of Oceanography
mmorzfeld@ucsd.edu

MS96

Continuous Learning by Integrating Reinforcement Learning and Data Assimilation to Individualise Drug Treatments

In many application areas, there is a need to determine a control variable that optimizes a prespecified objective. This problem is particularly challenging when knowledge on the underlying dynamics is subject to various sources of uncertainty and access to observations is severely limited. A scenario such as that for example arises in the context of therapy individualization to improve the efficacy and safety of medical treatment. Mathematical models describing the pharmacokinetics and pharmacodynamics of a drug together with data on associated biomarkers can be leveraged to support decision-making by predicting therapy outcomes. We present a continuous learning strategy that allows us to sequentially update the model parameters and states via a particle-based data assimilation scheme and combine it with reinforcement learning to tailor the dosing policy to the specific patient. We explore different schemes to parametrize the policy as well as the expected long-term reward in order to deal with the high dimensional action and state space. As these parameterizations may increase the overall uncertainty we explore how they are reflected in the approximated control variable.

Katherine Briceno Guerrero
Institute for Mathematics
University of Potsdam
bricenoguerrero@uni-potsdam.de

Jana de Wiljes
University of Potsdam
wiljes@uni-potsdam.de

MS96

A Nonlinear Dimension Reduction Method for Bayesian Inverse Problems Using Gradient Evaluations

A high dimensional Bayesian inverse problem has a low effective dimension when the data are informative only on a low-dimensional subspace. Identifying the parametric directions where the forward model does not vary significantly is a key preprocessing step to reduce the dimension of the problem, and thus to facilitate the construction of a reduced order model. In this talk, we introduce a methodology to detect and exploit such a low-dimensional structure using gradients of the model. The method consists in minimizing an upper-bound of the approximation error obtained using Poincar-type inequalities. This provides a certified bound on the error caused by the reduction of the parametric dimension. We then show how the method naturally extends to *nonlinear* dimension reduction, i.e. when the variations of the model is essentially contained on a low-dimensional manifold. Exploiting this kind of low-dimensional structure yields approximations under the form of compositions of functions. Within this framework, there is considerable flexibility and we show on various numerical examples the benefit of this approach.

Olivier Zahm
Univ. Grenoble Alpes, Inria

olivier.zahm@inria.fr

Clémentine Prieur
Grenoble Alpes University
Jean Kunzmann Lab, INRIA project/team AIRSEA
clementine.prieur@univ-grenoble-alpes.fr

Youssef M. Marzouk, Daniele Bigoni
Massachusetts Institute of Technology
ymarz@mit.edu, dabi@mit.edu

MS96

A Data-Driven and Model-Based Accelerated Hamiltonian Monte Carlo Method for Bayesian Elliptic Inverse Problems

We propose a data-driven and model-based approach to accelerate the Hamiltonian Monte Carlo (HMC) method in solving large-scale Bayesian inverse problems. The key idea is to exploit (model-based) and construct (data-based) the intrinsic approximate low-dimensional structure of the underlying problem which consists of two components – a training component that computes a set of data-driven basis to achieve significant dimension reduction in the solution space, and a fast-solving component that computes the solution and its derivatives for a newly sampled elliptic PDE with the constructed data-driven basis. Hence we achieve an effective data and model-based approach for the Bayesian inverse problem and overcome the typical computational bottleneck of HMC – repeated evaluation of the Hamiltonian involving the solution (and its derivatives) modeled by a complex system, a multiscale elliptic PDE in our case. We present numerical examples to demonstrate the accuracy and efficiency of the proposed method.

Zhiwen Zhang
Department of Mathematics,
The University of Hong Kong,
zhangzw@hku.hk

MS97

Lagrangian Uncertainty Quantification in Stochastic Flows

We develop a systematic information-theoretic framework for quantification and mitigation of error in probabilistic path-based (Lagrangian) predictions which are obtained from dynamical systems generated by uncertain (Eulerian) vector fields. This work is motivated by the desire to improve Lagrangian predictions in complex dynamical systems based either on analytically simplified or data-driven models. We derive a hierarchy of general information bounds on the uncertainty in estimates of statistical observables $E^{\nu}[f]$, evaluated on trajectories of the approximating dynamical system, relative to the true observables $E^{\mu}[f]$ in terms of certain φ-divergencies $D(\mu||\nu)$ which quantify discrepancies between probability measures μ associated with the original dynamics and their approximations ν. We then derive bounds on $D(\mu||\nu)$ itself in terms of the Eulerian fields. This framework provides a rigorous way for quantifying and mitigating uncertainty in Lagrangian predictions due to Eulerian model error. Links to uncertainty quantification in Data Assimilation techniques will also be mentioned.

Michal Branicki
University of Edinburgh
m.branicki@ed.ac.uk

MS97

Deep Learning-Enhanced Ensemble-Based Data Assimilation for High-Dimensional Nonlinear Systems

Data assimilation (DA) is an indispensable component of many prediction systems, particularly weather forecasting. Accurate DA with sequential algorithms such as ensemble Kalman filter (EnKF) requires generating a large ensemble of forecasts using an often-expensive dynamical model for calculating the background covariance matrix. In practice, only a small ensemble is generated and usually ad-hoc techniques are used to remove spurious correlations in the covariance matrix, but thus can also lead to the removal of physical correlations. Here, we leverage recent advances in deep learning-based data-driven forecast models and build a computationally inexpensive framework, named Hybrid EnKF (H-EnKF), to generate large data-driven ensembles and compute accurate background covariance matrices. A much smaller number of numerically generated ensembles are used to provide the background forecast, which will be used along with the covariance matrix and noisy observations to produce the analysis state. The performance of H-EnKF is demonstrated on a two-layer quasi-geostrophic turbulent system. Without the need for localization and for the same computational cost, H-EnKF substantially outperforms EnKF. The approach can be readily used for other ensemble-based DA methods, such as particle filters.

Pedram Hassanzadeh
Mechanical Engineering and Earth Science
Rice University
pedram@rice.edu

Ashesh K. Chattopadhyay
Mechanical Engineering
Rice University
akc6@rice.edu

Ebrahim Nabizadeh
Rice University
en10@rice.edu

Eviatar Bach
University of Maryland
eviatarbach@protonmail.com

MS97

An Efficient Continuous Data Assimilation Algorithm for the Sabra Shell Model of Turbulence

Complex nonlinear turbulent dynamical systems are ubiquitous in many areas of research. Recovering unobserved state variables is an important topic for the data assimilation of turbulent systems. We introduce an efficient continuous in time data assimilation scheme, which exploits closed analytic formulae for updating the unobserved state variables. Therefore, it is computationally efficient and accurate. The new data assimilation scheme is combined with a simple reduced order modeling technique that involves a cheap closure approximation and a noise inflation. In such a way, many complicated turbulent dynamical systems can satisfy the requirements of the mathematical structures for the proposed efficient data assimilation scheme. The new data assimilation scheme is then applied to the Sabra shell

model, which is a conceptual model for nonlinear turbulence. The goal is to recover the unobserved shell velocities across different spatial scales. It is shown that the new data assimilation scheme is skillful in capturing the nonlinear features of turbulence including the intermittency and extreme events in both the chaotic and the turbulent dynamical regimes. It is also shown that the new data assimilation scheme is more accurate and computationally cheaper than the standard ensemble Kalman filter and nudging data assimilation schemes for assimilating the Sabra shell model with partial observations.

Evelyn Lunasin
United States Naval Accademy
lunasin@usna.edu

Nan Chen, Yuchen Li
University of Wisconsin-Madison
chennan@math.wisc.edu, yli966@wisc.edu

MS97

Unbiased Filtering for a Class of Partially Observed Diffusion Process

In this article we consider a Monte Carlo-based method to filter partially observed diffusions observed at regular and discrete times. Given access only to Euler discretizations of the diffusion process, we present a new procedure that can return online estimates of the filtering distribution with no discretization bias and finite variance. Our approach is based upon a novel double application of the randomization methods of Rhee & Glynn (2015) along with the multilevel particle filter (MLPF) approach of Jasra et al (2017). A numerical comparison of our new approach with the MLPF, on a single processor, shows that similar errors are possible for a mild increase in computational cost. However, the new method scales strongly to arbitrarily many processors with high parallel efficiency. On 1024 cores, the new method demonstrates a 300 times speed-up compared to MLPF for similar errors.

Fangyuan Yu
KAUST
fangyuan.yu@kaust.edu.sa

Kody Law
University of Manchester
kodylaw@gmail.com

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

MS98

Spatial Logistic Gaussian Process for Density Field Modelling: Application to Stochastic Inverse Problems

Nowadays, stochastic simulators are extensively used to accurately model both natural and artificial systems. The response distribution can not only vary in mean and/or variance but also in other features including for instance shape or uni-modality versus multi-modality. Complex simulations come to a high computational cost, and it is common to rely on meta-modelling to perform statistical inference on the systems at hand. We consider a class of models for non-parametric estimation of the thereby induced fields of probability distributions based on scattered samples of heterogeneous sizes. The Spatial Logistic Gaussian Process (SLGP) models deliver probabilistic predictions of distributions at candidate points. Having a generative model enables us to perform uncertainty quantification on our predictions, and we leverage it to efficiently select additional simulation locations. We demonstrate applicability of our approach on stochastic inverse problems and show that relying on SLGP models allows for speeding-up Approximate Bayesian Computation methods. We illustrate it on applications from natural sciences.

Athénaïs Gautier
University of Bern, Switzerland
athenais.gautier@stat.unibe.ch

David Ginsbourger
University of Bern
david.ginsbourger@stat.unibe.ch

Guillaume Pirot
The University of Western Australia
guillaume.pirot@uwa.edu.au

MS98

Separating Intrinsic from Extrinsic Randomness with Gillespie's Algorithm for Epidemic Models

In mathematical modeling of infectious diseases spread, compartmental models are often used. These models can rely on random processes and they can also depend on many unknown parameters. So, they include two sources of variability: intrinsic randomness and parameter uncertainty. Therefore, one important issue related to the exploration of the global variability of model outputs, in the context of sensitivity analysis, concerns the separation of these two sources of input variability. In this work, an approach of separating intrinsic randomness from parameter uncertainty is proposed for two types of epidemic models: continuous-time Markov chain based models and non-Markovian models. Exact representations of model outputs as deterministic functions of uncertain parameters and some random variables that represent the intrinsic randomness are provided. These representations are built on exact stochastic simulation algorithms of random process corresponding to the two types of models considered: the Gillespie algorithm and the Sellke algorithm, respectively. An application to a model of SARS-CoV-2 spread is included to illustrate the practical impact of our approach, by calculating Sobol sensitivity indices for uncertain input parameters and intrinsic randomness.

Henri Mermoz Kouye
INRAE and Université Paris-Saclay
henri-mermoz.kouye@inrae.fr

Gildas Mazo
INRAE and Université Paris-Saclay, France
gildas.mazo@inrae.fr

Clémentine Prieur
Grenoble Alpes University
Jean Kunzmann Lab, INRIA project/team AIRSEA
clementine.prieur@univ-grenoble-alpes.fr

Elisabeta Vergu
INRAE and Université Paris-Saclay, France

elisabeta.vergu@inrae.fr

MS98

Surrogating Stochastic Simulators from Sample Trajectories Using a Non-Gaussian Random Field Approach

Stochastic simulators are a class of computational models that give a different response each time they are run, even if the same input parameters are used. Such a simulator can be viewed as a random field, indexed by the space of its input parameters. We focus on a class of stochastic simulators for which it is possible to generate trajectories, i.e., evaluations of the simulator throughout the space of input parameters for which the latent variables that induce the stochasticity of the simulator are held fixed (e.g., by fixing the random seed). Stochastic simulators are typically highly complex and expensive to run, which makes uncertainty analysis and optimization costly. These costs can be alleviated by replacing the simulator with a suitable surrogate model, which captures the essential characteristics of the original model while being much cheaper to evaluate. We propose a surrogate model that combines sparse polynomial chaos expansion, extended Karhunen-Love expansion, and parametric inference of joint distributions in the marginal-copula framework to represent the stochastic simulator based on a number of model evaluations. The resulting surrogate model has an analytical form that can easily be used to compute moments and to sample new trajectories. In this talk, we demonstrate the performance of our surrogate model on a real-world engineering application, and show how it can be utilized to perform conditional prediction.

Nora Lüthen
ETH Zürich
luethen@ibk.baug.ethz.ch

Stefano Marelli
Chair of Risk, Safety and Uncertainty Quantification
Institute of Structural Engineering, ETH Zurich
marelli@ibk.baug.ethz.ch

Bruno Sudret
ETH Zurich
sudret@ibk.baug.ethz.ch

MS98

Extension of Polynomial Chaos Expansions to the Metamodeling of Stochastic Simulators

Performing uncertainty quantification requires repetitive evaluations of computational models, which is intractable for high-fidelity expansive models. This problem exacerbates for stochastic simulators, the output of which is a random variable: each model evaluation produces different values of the model response even when using the same input parameters. To alleviate the computational burden associated with the use of stochastic simulators in optimization or uncertainty quantification, surrogate models have gained attention in recent years. In this field, many methods have been focused on representing some summary statistics of the random response (such as mean, variance, and quantiles, etc.) as deterministic functions of the input variables. However, much less effort has been devoted to the emulation of the entire probability distribution of the model response conditional to the input parameters. In this contribution, we extend the classical polynomial chaos expansion to emulating the response distribution of stochastic simulators. To reproduce the stochasticity in the model output, we include a latent variable in the expansion on top of the well-defined input parameters and introduce an additive Gaussian noise. We propose to combine the maximum likelihood estimation with cross-validation to fit the surrogate model from data. We compare the performance of the novel approach with other state-of-the-art methods on various examples from engineering and epidemiology.

Xujia Zhu
ETH Zürich
zhu@ibk.baug.ethz.ch

Bruno Sudret
ETH Zurich
sudret@ibk.baug.ethz.ch

MS99

Edge-Preserving Bayesian Optimal Experimental Design in Inverse Problems

While computational resources are growing rapidly, the data acquisition in many large-scale inverse problems will remain restricted or expensive due to fundamental physical or economical limitations related to the task. As common examples of such restrictions consider, e.g., astronomical or seismic imaging. Accordingly, efficient methods are needed to computationally design e fficient experimental settings to maximize the information content of experimental data. Bayesian optimal experimental design (OED) provides a principled approach for this task firmly based on statistics. In this talk we discuss novel edge-preserving sequential Bayesian OED methods inspired by so-called lagged diffusivity iteration. Our method is computationally efficient for large-scale problems as it relies on iterative approximations of the posterior by Gaussian distributions.

Tapio Helin
LUT University
Tapio.Helin@lut.fi

Nuutti Hyvönen
AaltoUniveristy
nuutti.hyvonen@aalto.fi

Juha-Pekka Puska
Aalto University
juha-pekka.puska@aalto.fi

MS99

Optimal Bayesian Design of Sequential Experiments Using Deep Deterministic Policy Gradient

Experiments are indispensable for learning and developing models in engineering and science. When experiments are expensive, a careful design of these limited data-acquisition opportunities can be immensely beneficial. Optimal experimental design, while leveraging the predictive capabilities of a simulation model, provides a rigorous framework to systematically quantify and maximize the value of experiments. We focus on designing a finite sequence of experiments, seeking fully optimal design policies (strategies) that can (a) adapt to newly collected data during the sequence (i.e. feedback) and (b) anticipate future changes (i.e. lookahead). We cast this sequential decision-making problem in a Bayesian setting with information-based utilities, and solve it numerically via policy gradient methods from reinforcement learning. In particular, we directly pa-

rameterize the policies and value functions by neural networksthus adopting an actor-critic approachand improve them using gradient estimates produced from simulated design and observation sequences. The overall method is demonstrated on an algebraic benchmark and a sensor movement application for source inversion. The results provide intuitive insights on the benefits of feedback and lookahead, and indicate substantial computational advantages compared to previous numerical approaches based on approximate dynamical programming.

Wanggang Shen, Xun Huan
University of Michigan
wgshen@umich.edu, xhuan@umich.edu

MS99

Convex Relaxation for Sensor Selection in the Presence of Correlated Measurement Noise

Best sensor selection is of paramount importance when monitoring spatiotemporal phenomena using large-scale sensor networks. Our main aim is to develop a technique of this type for maximizing the parameter estimation accuracy when the system in question is modeled by a partial differential equation and the measurement noise is correlated. The weighted least squares method is supposed to be used for estimation and the determinant of the covariance matrix of the resulting estimator is adopted as the measure of estimation accuracy. This design criterion is to be maximized by choosing a set of spatiotemporal measurement locations from among a given finite set of candidate locations. To make this combinatorial problem computationally tractable, its relaxed formulation is considered. Optimal solutions are found using extremely efficient simplicial decomposition. The sequence of iterates monotonically increases the value of the original concave design criterion. As the resulting relaxed solution is a measure on the set of candidate measurements and not a specific subset, randomization and a restricted exchange algorithm are used to convert it to a nearly-optimal subset of selected sensors. A simulation experiment is reported to validate the proposed approach. The generality of the proposed technique makes it suitable for other measurement selection problems for least-squares estimation subject to correlated observations.

Dariusz Ucinski
Institute of Control and Computation Engineering
University of Zielona Góra
d.ucinski@issi.uz.zgora.pl

MS100

Multilevel Ensemble-Based Data Assimilation for Subsurface Flow Problems

The process of conditioning a model to observations data assimilation - is a crucial step when applying models describing subsurface flow in an operational setting. Within reservoir engineering, data assimilation using Kalman-filter type estimation methods have been successfully applied in a wide range of applications. Based on an ensemble of models, Monte-Carlo estimates of mean and covariance are applied. However, it is well known that Monte-Carlo errors can significantly decrease the accuracy of the results. Multi-level Monte-Carlo methods have shown large potential for efficient estimation of statistical quantities. Moreover, multi-level estimators have been applied in data assimilation methods. The reservoir engineering problem is described by a system of partial differential equations with poorly known coefficients. The simulation models are typically very large, and black-box solvers are typically applied. Hence, it is difficult to define a multi-fidelity simulator. In this talk, I discuss the main challenges when applying multi-level ensemble-based data assimilation methods for subsurface flow problems. I will present multiple ways of obtaining multi-fidelity simulators and demonstrate why the standard multi-level schemes fail. Moreover, I will introduce and demonstrate some alternative multilevel methods and compare them to single fidelity methods.

Kristian Fossum
NORCE
kristian.fossum@norceresearch.no

MS100

Smc2 for Financial Volatility Models

SMC^2 algorithms, which consist of two nested particle filters, can be used for sequential joint parameter and state estimation for non-linear and non-Gaussian state-space models. We first showcase applicability of the SMC^2 algorithm with simple Ornstein-Uhlenbeck and Heston stochastic volatility models. Then we generalise to non-Gaussian financial rough volatility models, and demonstrate the applicability of the algorithm. Briefly, we also discuss how to use the methodology to physical state-space models driven by Cauchy noise. Finally, we apply the rough volatility models and SMC^2 methodology to the January 2015 Swiss Franc-Euro de-pegging event.

Lassi Roininen
Lappeenranta-Lahti University of Technology
lassi.roininen@lut.fi

MS100

Ensemble Kalman Filter (enkf) for Reinforcement Learning (rl)

This talk is concerned with the problem of representing and learning the optimal control law for the linear quadratic Gaussian (LQG) optimal control problem. In recent years, there is a growing interest in re-visiting this classical problem, in part due to the successes of reinforcement learning (RL). The main question of this body of research (and also of our paper) is to approximate the optimal control law *without* explicitly solving the Riccati equation. For this purpose, a novel simulation-based algorithm, namely an ensemble Kalman filter (EnKF), is introduced. The algorithm is used to obtain formulae for optimal control, expressed entirely in terms of the EnKF particles. For the general partially observed LQG problem, the proposed EnKF is combined with a standard EnKF (for the estimation problem) to obtain the optimal control input based on the use of the separation principle. The theoretical results and algorithms are illustrated with numerical experiments.

Anant Joshi
University of Illinois at Urbana Champaign
anantaj2@illinois.edu

Amirhossein Taghvaei
University of Washington
amirtag@uw.edu

Prashant G. Mehta
Department of Mechanical and Industrial Engineering

University of Illinois at Urbana Champaign
mehtapg@illinois.edu

MS100

Sparse Online Variational Bayesian Inference

This work aims to study variational Bayesian inference for sparse regression. Sparsity promoting priors have proven to be very successful in regression scenarios since it helps to select meaningful features and avoid overfitting (e.g. LASSO regression or total variation (TV) regularization in imaging). We focus on a general class of shrinkage priors that can be represented as a scale mixture of normal distributions with a generalized inverse Gaussian distribution and includes such priors as Laplace, Generalized Jeffrey's, Student-t and others. However, since shrinkage priors are non-Gaussian, a fully Bayesian solution becomes very expensive, requiring MCMC methods. To alleviate this, we employ a variational approach that leverages a generalization of the expectation-maximization algorithm to recover the best Gaussian approximation to the sparsity-promoting posterior. This approach turns out to be especially fast and scalable in the case of linear models, where it provides approximate UQ for a substantially smaller cost than fully Bayesian approaches yet keeping comparable accuracy. Besides, the proposed approach supports online inference to process the data in batches and strategies for online hyperparameter estimation. The high performance in terms of the variable selection and UQ is demonstrated for complex real and simulated data examples where it competes against MCMC based methods as well as the other approximate approaches.

Vitaly Zankin
University of Manchester
The Alan Turing Institute
vitalii.zankin@manchester.ac.uk

Kody Law
University of Manchester
kodylaw@gmail.com

MS101

A Quantile Conserving Particle Filter with Likelihood Localization

Recent work has shown that a number of deterministic ensemble filter variants can be unified as specific examples of a more general quantile conserving ensemble filter (QCEF) framework. The QCEF algorithm begins by finding an appropriate continuous probability density function (PDF) and associated cumulative distribution function (CDF) corresponding to the prior marginal ensemble for each scalar state variable. It then uses Bayes and a continuous likelihood to compute a continuous posterior PDF and CDF. The posterior ensemble is selected so that the quantiles of the posterior ensemble with respect to the posterior CDF are identical to the quantiles of the prior ensemble with respect to the prior CDF. A deterministic particle filter algorithm is developed using the QCEF framework. Combined with localization of the likelihood it can produce posterior ensemble samples of the marginal for all state variables in a prediction model. However, successful cycling data assimilation in low-order dynamical system observing system simulation experiments also requires a marginal adjustment step in which the rank order of the different marginal samples is adjusted to be consistent with results from a more traditional multivariate ensemble filter. The algorithm will be described along with some results from Lorenz-96 model experiments

Jeffrey Anderson
National Center for Atmospheric Research
Institute for Math Applied to Geophysics
jla@ucar.edu

MS101

Formulation of the Variational Fokker-Planck Filters and Smoothers

Particle flow filters aim to smoothly transform particles from being samples of a prior distribution to becoming samples of a posterior distribution. The particle dynamics in state space is described by a stochastic differential equation. We discuss several aspects of the formulation of Variational Fokker-Planck method for filtering and smoothing, including the formulation of optimal drift of particle dynamics equations, and the regularization of the flow, based on information theoretic arguments, such as to avoid particle collapse.

Adrian Sandu
Virginia Polytechnic Institute and State University
Computational Science Laboratory
sandu@cs.vt.edu

MS101

A Variational Fokker-Planck Method for Data Assimilation

Particle filters are used for state estimation while dealing with uncertainty in model dynamics and observations. Unlike traditional filters, particle filters are constrained by fewer assumptions. Variational particle filtering techniques aim to minimize the KL Divergence between the forecast prior and a target posterior that combines the noisy observed data with the prior states. One drawback of these methods is the excessive tuning in the face of high dimensionality. Our formulation via the Fokker-Planck equation performs the above minimization successfully even in higher dimensions. We have developed methods for both filtering and smoothing. Our experiments show that out formulation is competitive with many state-of-the-art methods.

Amit Subrahmanya
Virginia Tech
amitns@vt.edu

Andrey A. Popov
Virginia Polytechnic Institute and State University
apopov@vt.edu

Adrian Sandu
Virginia Polytechnic Institute and State University
Computational Science Laboratory
sandu@cs.vt.edu

MS101

Particle Flow Filters and Smoothers for High-Dimensional Geophysical Problems

Fully nonlinear Bayesian Inference in high-dimensional systems is hard to achieve, and progress in the last 50 years is perhaps disappointing. Variational methods suffer from a Gaussian prior. They target the mode of the posterior pdf and are not efficient to provide uncertainty on that

estimate. Monte-Carlo methods such as Particle Filters suffer from weight collapse, even when only local updates are performed, simply because the number of independent observations is too high in these local areas. Markov Chain methods are inefficient because the algorithm to generate samples is essentially sequential. Recently optimal transportation methods have gained popularity, especially so-called Particle Flow Filters. Particle Flow Filters move samples from the prior into samples from the posterior by solving an SPDE for each particle, where the flow fields of different particles are coupled. No weighting is needed, so filter degeneracy does not occur. Many variants have been and are being developed, and some can be applied to very high dimensional systems. We will show results for Particle Flow Filters and Smoothers in which the flow field is embedded in a Reproducing Kernel Hilbert Space. Specifically, both filters and smoothers methods have been applied to high-dimensional Lorenz 1996 models and to multilayer quasi-geostrophic models. We specifically concentrate on the ability of the methods to accurately represent multimodal posterior pdfs in high-dimensional systems.

Peter Jan van Leeuwen
Colorado State University and University of Reading
peter.vanleeuwen@colostate.edu

Chih-Chi Hu
Colorado State University
chihchi.hu@colostate.edu

MS102

Nonlinear Model Reduction via Projection onto Optimal Dynamically Orthogonal Subspaces

The success of Petrov-Galerkin projection for the reduction of PDE models relies critically on the existence and identification of a low-dimensional subspace in which the characterizing dynamic features of the high-fidelity model can be captured. While many traditional methods such as POD provide a powerful machinery to identify such subspaces in the form of a time-invariant basis simply from simulation data, they are known to fall short in certain settings. A particularly important shortcoming in the context of space weather applications is that these techniques are challenged in the advection dominated regime; in this regime, capturing the propagation of sharp features such as shocks through time and space is required, however, only seldom possible with a low-dimensional time-invariant subspace. To address this issue, we present a versatile smooth optimal control formulation that seeks to identify a time-variant subspace spanned by dynamically orthogonal evolving modes from off-line generated simulation data. We show how this formulation informs several strategies for building projection-based reduced order models. These strategies range from direct solution of the optimal control problem which allows fitting of ROM trajectories to reference trajectories, over more scalable greedy approaches akin dynamical low rank approximation, to fully data driven approaches. We discuss the necessary computational infrastructure developed in Julia and provide illustrative examples.

Flemming Holtorf
Julia Lab
Massachusetts Institute of Technology
holtorf@mit.edu

Alan Edelman
Department of Mathematics
Massachusetts Institute of Technology and Julia Computing
edelman@mit.edu

Chris Rackauckas
Julia Computing and MIT
chris.rackauckas@juliacomputing.com

MS102

Sobol' Sensitivity for Uncertain Model Parameters in Simulations of Background Solar Wind

The Space Weather Modelling Framework (SWMF) offers efficient and flexible sun-to-earth simulations based on coupled first principles and/or empirical models. This encompasses computing the quiet solar wind, generating a coronal mass ejection (CME), propagating the CME through the heliosphere, and calculating the magnetospheric impact via geospace models. Accurate long-term predictions from these different steps and models are challenging due to the uncertainty and variation of many model inputs and parameters. In this work, we perform uncertainty quantification (UQ) for the quiet solar wind simulations produced by the Alfven Wave Solar atmosphere Model (AWSoM). First, the various sources of parametric uncertainty and their distributions are catalogued. This leads to space-filling designs for high-fidelity simulations that propagate the uncertainty from the inputs to key predictive quantities of interest (QoIs) such as radial velocity and number density. In the next step, Polynomial Chaos Expansion (PCE) surrogates are built for each QoI, enabling Global Sensitivity Analysis (GSA) through calculation of time-varying Sobol indices. Assessing the sensitivity allows retention of only the most impactful parameters from the stochastic space. We summarize our findings for solar maximum and minimum conditions.

Aniket Jivani, Xun Huan, Yang Chen, Bart van der Holst, Shasha Zou, Zhenguang Huang, Nishtha Sachdeva, Daniel Iong, Ward Manchester
University of Michigan
ajivani@umich.edu, xhuan@umich.edu, ychenang@umich.edu, bartvand@umich.edu, shashaz@umich.edu, zghuang@umich.edu, nishthas@umich.edu, daniong@umich.edu, chipm@umich.edu

Gabor Toth
Center for Space Environment Modeling
University of Michigan
gtoth@umich.edu

MS102

Shifted Operator Inference for Data-Driven Modeling of Solar Winds

In this talk, we present a new data-driven reduced-order modeling strategy for advection dominated phenomena, such as solar winds. The methods is based on operator inference, which learns reduced-order model operators directly from data. We augment the method by shifting the basis along with the advection direction. We suggest two strategies to detect the shift, one by analyzing the characteristics of the flow solution (in a situation where the PDE is known, but no access to the code is given) and one by learning the shift from a given data set, which then constitutes a fully non-intrusive framework. This new method is developed for the challenges that arise in modeling solar winds, which consist of a continuous flow of charged particles from the sun caused by coronal heating. Accurate and

fast forecasts of solar winds are important avoid damage to space equipment. Our numerical results show that we can accurately predict solar wind streams in one and two-dimensional models at much reduced computational cost.

Boris Kramer
University of California San Diego
bmkramer@ucsd.edu

Opal Issan
University of California, San Diego
oissan@ucsd.edu

MS102

Modeling Solar Wind Flow with Quantified Uncertainties at the Earth's Orbit

The solar wind (SW) emerging from the Sun is the main driving mechanism of solar events which may lead to geomagnetic storms that are the primary causes of space weather disturbances that affect the magnetic environment of Earth and may have hazardous effects on the space-borne and ground-based technological systems as well as human health. We describe our approach to the development of a new chain of models based on modern computational methods and entirely open source. The new models will provide more accurate solutions and will be scalable on massively parallel systems, including Graphic Processing Units. This will allow the user community to easily experiment with these tools and combine them with other models to produce more advanced space weather modeling capabilities. This suite allows us to perform simulations with uncertainty quantification, which is implemented through ensemble modeling based on predictive metrics and skill analysis. The uncertainties in the input data are considered together with the coronal mass ejection model parameter uncertainties. The application of machine learning techniques is discussed.

Nikolai Pogorelov
University of Alabama at Huntsville
np0002@uah.edu

Charles N. Arge
Goddart Space Flight Center
charles.n.arge@nasa.gov

Ronald Caplan
Predictive Science Inc.
caplanr@predsci.com

Dinesha Vasanta Hegde
Department of Space Science
University of Alabama in Huntsville
dvh0006@uah.edu

Carl Henney
AFRL/RVBXD
carl.henney.1@us.af.mil

Shaela Jones-Mecholsky
Goddard Space Flight Center
shaela.i.jonesmecholsky@nasa.gov

Tae Kim
CSPAR, University of Alabama in Hunstville
tae.kim@uah.edu

Jon Linker
Pedictive Science Inc.
linkerj@predsci.com

Syed Raza, Talwinder Singh
CSPAR, University of Alabama in Hunstville
syed972@uab.edu, talwinder.singh@uah.edu

Lisa Upton
Space Systems Research Corporation
upton.lisa.a@gmail.com

MS103

History Matching: Overview and Challenges

This talk provides an introduction and overview for history matching, covering basic ideas, how they are implemented, the challenges involved in applying the approach to large and complex systems and some of the directions in which the methodology has developed in order to meet these challenges.

Michael Goldstein
Durham University
michael.goldstein@durham.ac.uk

MS103

Design of Physical Experiments for History Matching

History matching aims to find the set of all non-implausible inputs to a computer model, that is, those which are not inconsistent with observed data, given all the sources of uncertainty associated with the model and the measurements. The progress of a history match is often measured in terms of the proportion of the initial input space classed as implausible, although other criteria can be used, such as the reduction in the variance of scientifically important parameters. Analysis of such quantitative features of the non-implausible set is informative for answering current questions about the real-world quantities associated with the input parameters and the links between them. We therefore quantify the expected information gain resulting from performing possible future physical experiments in terms of history matching criteria related to scientific questions of interest, thus allowing the most relevant and informative experiments to be performed. We demonstrate our techniques on an important systems biology model of hormonal crosstalk in the roots of an Arabidopsis plant.

Sam Jackson, Ian Vernon
University of Durham
samuel.e.jackson@durham.ac.uk,
ian.vernon314@gmail.com

MS103

Tackling Persistent Climate Model Uncertainty Using a History Matching Approach

The effects of aerosols (small particles suspended in the air) on the Earths energy balance since pre-industrial times (aerosol radiative forcing) has significantly and repeatedly dominated the uncertainty in reported estimates of global temperature change from the Intergovernmental Panel on Climate Change (IPCC). The magnitude of aerosol radiative forcing of climate over the industrial period is estimated to lie between -2 and -0.4 W m^{-2}, compared to a much better understood forcing of 1.6 to 2.0 W m^{-2} due

to CO_2. In this study we quantify the range of possible aerosol forcings in the HadGEM3-UKCA aerosol-climate model caused by parametric uncertainty and then constrain that forcing uncertainty through a history matching approach, using an extensive set of (9000+) aerosol measurements (including aerosol optical depth, and PM2.5, N50 and sulphate concentrations) from ships, flight campaigns and ground stations. We find that despite a very large reduction in plausible parameter space, and reasonable constraint on global and regional mean aerosol properties, the observational constraint based on this comprehensive set of measurements only mildly reduces the range of aerosol radiative forcings from our model. This work has highlighted several key statistical challenges to address in order to improve the model-observation comparison process for constraint, including better characterisations of representation errors and reducing error compensation effects.

Jill S. Johnson
The University of Sheffield
jill.johnson@sheffield.ac.uk

Leighton Regraye, Ken Carslaw
University of Leeds
l.a.regayre@leeds.ac.uk, k.s.carslaw@leeds.ac.uk

MS104

Model Correction and Validation of Reduced Lotka-Volterra Models

Interacting physical systems often comprise hundreds of species and their complex dynamics, such as ecological communities, chemical reactions, or epidemic populations. Modeling these interactions with standard Lotka-Volterra type models quickly becomes computationally expensive or intractable, and so we often use reduced models involving only interactions between the species of interest. But these reductions can lead to high model error, rendering the model useless for various prediction tasks, e.g., probabilities of species extinctions, contaminant levels, or epidemic outbreaks. Therefore, we explore the reduction of model error through interpretable model correction. We augment partial (reduced) Lotka-Volterra models with an inadequacy operator to form an enriched model, calibrate with hierarchical Bayesian inference, and validate with a posterior predictive assessment. The inadequacy operator is informed by the physical system and contains terms to capture the equilibrium and transient behavior of the species of interest. Results show that the model error-caused by omitting interaction terms beyond the species of interestcan be recovered with an inadequacy operator involving a small fraction of the number of omitted terms. We also identify a surprising trend between the complexity of the partial model and the enriched models error: it is easier to correct partial models involving either a small or large number of species, while the mid-range is the hardest to correct.

Rileigh Bandy
University of Colorado Boulder
rileigh.bandy@colorado.edu

Rebecca Morrison
CU Boulder
rebeccam@colorado.edu

MS104

Physics Constrained Learning for Model Form Uncertainty Quantification

This work proposes techniques to quantify and reduce model form errors by embedding learnable functions within partial differential equation-based models. With a view towards ensuring robustness, the feature space and the features-to-augmentation map are carefully designed. The approach is designed to ingest sparse data from different physical problems, and create models to make predictions in unseen configurations. To promote generalizability, a constrained optimization problem is solved to minimize the mutual information between the observed data and the model parameters while maintaining a threshold on the reconstruction accuracy. Demonstrations are presented on synthetic partial differential equation examples as well as in the prediction of turbulent flows.

Sahil Bhola
University of Michigan, Ann Arbor
sbhola@umich.edu

Karthik Duraisamy
University of Michigan Ann Arbor
kdur@umich.edu

MS104

Efficient Characterization of Model Uncertainty for Ice Sheet Flow Problems

We consider the Bayesian inference of the unknown basal sliding coefficient field in the presence of additional uncertainty in the nonlinear first-order Stokes ice sheet model MALI. To account for the associated model uncertainty (due to an imperfectly known secondary nuisance stiffening factor parameter), we employ the Bayesian Approximation Error (BAE) approach. With BAE the nuisance parameter and measurement noise are approximately premarginalized, which results in a posterior distribution for the basal sliding coefficient field. We discuss an efficient approximation strategy for the first and second order statistical moments of the model discrepancy, which is a critical component of the BAE framework. The resulting (approximate) Gaussian model discrepancy probability density is utilized for the inference of the basal sliding coefficient from noisy surface velocity measurements on the Humboldt glacier. We show that the BAE approach avoids overly confident and erroneous inference such as that which can occur when model uncertainty is neglected entirely.

Tucker Hartland
University of California, Merced
thartland@ucmerced.edu

Mauro Perego
CCR Sandia National Laboratories
mperego@sandia.gov

Georg Stadler
Courant Institute for Mathematical Sciences
New York University
stadler@cims.nyu.edu

Noemi Petra
University of California, Merced
npetra@ucmerced.edu

MS104

Discrepancy Prediction in Dynamic System Models

Using Machine Learning

Simulation models are used to make predictions where there is no experimental data, such as untested input time histories. The reliability of simulation models for these conditions is in question since the model discrepancy in the predictions is unknown. In this work, a probabilistic framework for discrepancy prediction is developed for dynamic system models under untested input time histories. First, model discrepancy is estimated from available experiments using state estimation. Then two surrogate modeling-based methods, namely observation surrogate and bias surrogate, are developed. In the first method, a neural network is trained for the observed experimental output, and the model discrepancy for the untested input is obtained by comparing the output of the observation surrogate with the output of the physics-based model. The second method trains a neural network for the discrepancy in terms of the inputs in the conducted experiments. The discrepancy machine learning model is then used to correct the simulation model prediction at each time step under a predictor-corrector scheme; the corrected prediction is compared with the original prediction to compute the model discrepancy under untested conditions. The two approaches are demonstrated to predict the discrepancy in an air cycle machine computational simulation under an untested time history.

Kyle Neal
Sandia National Laboratories
kneal@sandia.gov

Zhen Hu
University of Michigan-Dearborn
zhennhu@umich.edu

Sankaran Mahadevan
Vanderbilt University
sankaran.mahadevan@vanderbilt.edu

Jon Zumberge
Air Force Research Laboratory
jon.zumberge@us.af.mil

MS105

Low-Rank Conditional Structure in Transport Maps for Bayesian Inverse Problems

Approximate sampling via probability transport maps has recently gained popularity as an alternative and a complement to standard sampling strategies such as MCMC. The computational effort needed to train a transport map depends on the total number of map parameters, making high-dimensional inference problems difficult. We present a new method for encoding sparsity in transport maps by exploiting low-rank conditional structure in the target distribution. We enforce map components to depend on low-dimensional summaries of the input variables we learn by solving a particular eigenvalue problem. We will demonstrate performance improvements in both the accuracy of the approximate posterior and in the training behavior of the map compared to un-structured maps.

Michael C. Brennan, Youssef M. Marzouk
Massachusetts Institute of Technology
mcbrenn@mit.edu, ymarz@mit.edu

MS105

BayesFlow: New Advances From the Frontier of Simulation-Based Inference

Invertible neural networks have proven remarkably powerful in performing amortized simulation-based Bayesian inference. However, important issues, such as Bayesian model comparison, model misspecification, and sources of epistemic uncertainty, have barely been the topic of discussion throughout the recent literature. In this talk, we present novel developments aimed at tackling these issues within a single framework for performing fully Bayesian amortized inference. Further, we briefly showcase the beta version of out BayesFlow library which intends to bridge the gap between custom tailored solutions for individual applications and the general needs of modelers relying on simulation-based inference.

Stefan Radev
Universität Heidelberg
stefan.radev93@gmail.com

Paul-Christian Bürkner, Marvin Schmitt
University of Stuttgart
paul.buerkner@gmail.com,
marvin.schmitt@psychologie.uni-heidelberg.de

Ullrich Köthe
Heidelberg University
ullrich.koethe@iwr.uni-heidelberg.de

MS105

Statistical Inverse Problems and Affine-Invariant Gradient Flow Structures in the Space of Probability Measures

Statistical inverse problems lead to complex optimisation and/or Monte Carlo sampling problems. Gradient descent and Langevin samplers provide examples of widely used algorithms. In my talk, I will discuss recent results on sampling algorithms, which can be viewed as interacting particle systems, and their mean-field limits. I will highlight the geometric structure of these mean-field equations within the, so called, Otto calculus, that is, a gradient flow structure in the space of probability measures. Affine invariance is an important outcome of recent work on the subject, a property shared by Newtons method but not by gradient descent or ordinary Langevin samplers. The emerging affine invariant gradient flow structures allow us to discuss coupling-based Bayesian inference methods, such as the ensemble Kalman filter, as well as invariance-of-measure-based inference methods, such as preconditioned Langevin dynamics, within a common mathematical framework. Applications include nonlinear and logistic regression.

Sebastian Reich
Universität Potsdam, Germany
sereich@uni-potsdam.de

MS106

Multilevel Stein Variational Gradient Descent with Applications to Bayesian Inverse Problems

This work presents a multilevel variant of Stein variational gradient descent to more efficiently sample from target distributions. The key ingredient is a sequence of distributions with growing fidelity and costs that converges to the target distribution of interest. For example, such a sequence of distributions is given by a hierarchy of ever finer discretization levels of the forward model in Bayesian inverse problems. The proposed multilevel Stein variational gradi-

ent descent moves most of the iterations to lower, cheaper levels with the aim of requiring only a few iterations on the higher, more expensive levels when compared to the traditional, single-level Stein variational gradient descent variant that uses the highest-level distribution only. Under certain assumptions, in the mean-field limit, the error of the proposed multilevel Stein method decays by a log factor faster than the error of the single-level counterpart with respect to computational costs. Numerical experiments with Bayesian inverse problems show speedups of more than one order of magnitude of the proposed multilevel Stein method compared to the single-level variant that uses the highest level only.

Terrence Alsup
New York University
Courant Institute of Mathematical Sciences
alsup@cims.nyu.edu

Benjamin Peherstorfer
Courant Institute of Mathematical Sciences
New York University
pehersto@cims.nyu.edu

Luca Venturi
Courant Institute of Mathematical Sciences
New York, United States
lv800@nyu.edu

MS106

A Scalable Sampling Approach for Multilevel Markov Chain Monte Carlo

Model-based simulations are an important tool for predicting physical phenomena. In the presence of uncertain physical parameters, which describe the system of interest, accurately assessing confidence in these results, e.g., performing Bayesian inference, can be computationally expensive, particularly when the uncertain physical parameters are spatially varying quantities. Not only are the corresponding high-fidelity (fine grid) simulations expensive, but commonly used statistical approaches, such as Markov chain Monte Carlo (MCMC), require an exceedingly high number of simulations. In this work, we develop a new scalable sampling method which complements the algorithmic framework of multilevel MCMC, where coarse grid simulations are used to inform the fine level proposal distribution, thereby accelerating MCMC. Specifically, using tools from algebraic multigrid, we form a (scalable) fine grid Gaussian random field realization from a fine grid proposal by combining Gaussian random fields sampled across multiple levels of discretization. In this talk, we describe this new approach, corresponding theory, and numerical results when applied to a 3D subsurface flow application.

Hillary Fairbanks
Lawrence Livermore National Laboratory
fairbanks5@llnl.gov

Umberto Villa
Electrical and Systems Engineering
Washington University in St Louis
uvilla@wustl.edu

Panayot Vassilevski
Lawrence Livermore National Laboratory
vassilevski1@llnl.gov

MS106

Multifidelity-Multilevel Approaches for Solution of Parametric Partial Differential Equations: Application to Topology Optimization

In this talk we present two main topics involving solution of parametric partial differential equations. We begin with the problem of robust topology optimization and present an approach which approximates the high fidelity structural analysis solutions, design sensitivities and subsequently the robust design with a neural network which is trained by the data associated with the map between low resolution images and the low rank approximation coefficients in a bi-fidelity approximation setting [Keshavarzzadeh et al., Robust topology optimization with low rank approximation using artificial neural networks, Computational Mechanics, 2021]. We then briefly discuss multilevel approximation to the solution of linear elliptic parametric PDEs using a novel numerical quadrature rule. We show the application of this approach to the analysis of linear elastic structures with random field elastic modulus which are frequently used in the context of robust design optimization [Keshavarzzadeh et al., Multilevel Designed Quadrature for Partial Differential Equations with Random Inputs, SIAM Journal on Scientific Computing, 2021].

Vahid Keshavarzzadeh
University of Utah
vkeshava@sci.utah.edu

Mike Kirby
University of Utah
School of Computing
kirby@cs.utah.edu

Akil Narayan
University of Utah
akil@sci.utah.edu

MS106

Multifidelity Methods for Covariance Estimation with Applications to Data Assimilation

Ensemble data assimilation (DA) methods perform sequential state estimation in dynamical systems by interleaving applications of a dynamical model with conditioning on observations. The latter involves estimating and applying a prior-to-posterior transformation (P2PT). The ensemble Kalman filter (EnKF) constructs affine P2PTs from covariance matrix estimates. Nonlinear generalizations of the EnKF may instead solve a convex optimization for parameters of a nonlinear transformation. In most practical DA scenarios the ensemble size is small relative to state dimension, rendering P2PT estimation difficult. We thus consider multifidelity schemes in which cheaper, less accurate dynamical models are used alongside their high-fidelity counterpart. To this end, we present three multifidelity covariance estimators. We first extend optimal linear control variate (LCV) multifidelity estimators to matrix-valued random variables on Euclidean space; here we achieve significant variance reduction but find that the resulting estimates may be indefinite. We next present two multifidelity covariance estimators which preserve positive definiteness by construction: one applies LCVs to the covariance matrix logarithm, and the other performs multifidelity estimation on the manifold of positive definite matrices via regression.

Finally, we comment on links to multifidelity convex optimization, where loss of convexity may occur if LCVs are applied to a sample average estimate of the objective.

Aimee Maurais
MIT
maurais@mit.edu

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

MS107

Fast Generation and Enumeration Strategies of Frolov Lattices

In high dimensional integration, lattices rules are highly effective. In particular, prescribed lattices yield reliable error guarantees for many classical smoothness classes. As far as implementation is concerned, it is crucial to be able to enumerate lattices points in hypercubes inexpensively. Existing enumeration procedures are based on intrinsic properties of the lattices such as periodicity, orthogonality, recurrences, etc. In this paper, we present a general-purpose fast enumeration procedure based on linear programming. The procedure has proven to be very flexible and effective. We discuss its general formulation, acceleration, and relaxation techniques. Numerical experiments concerned with the enumeration of Frolov-Chebyshev lattices are also presented.

Abdellah Chkifa
Mohamed VI Polytechnic University
abdellah.chkifa@um6p.ma

MS107

Density Estimation for lognormal PDEs Using Quasi-Monte Carlo with Preintegration

Quasi-Monte Carlo (QMC) methods have shown great success in tackling difficult high-dimensional problems that often occur in uncertainty quantification. However, a key limitation is that they only approximate the expected value of the quantity of interest. One of the main reasons for this limitation is the smoothness requirements, such as requiring square-integrable mixed first derivatives. In this talk, we present a method for approximating the cumulative distribution function (cdf) and probability density function (pdf) of a quantity of interest coming from the solution of an elliptic PDE with lognormal random coefficients. The key idea is to formulate the cdf (and pdf) as an expected value, or equivalently, a high-dimensional integral, which can then be efficiently approximated by QMC. Typically QMC methods struggle to efficiently approximate the cdf because of a lack of smoothness in the integrand, which for the cdf is an indicator function. We overcome this by using an initial preintegration step to smooth out the integrand. Preintegration, also known as conditional Monte Carlo, is a method for smoothing a discontinuous function by integrating with respect to a single specially chosen variable. The result is a function in one dimension less that is now smooth (under appropriate assumptions). We will outline the QMC with preintegration method for approximating the cdf and pdf for lognormal PDEs, then present an error analysis and numerical results.

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

Frances Y. Kuo
School of Mathematics and Statistics
University of New South Wales
f.kuo@unsw.edu.au

Ian H. Sloan
University of New South Wales
School of Mathematics and Stat
i.sloan@unsw.edu.au

Abirami Srikumar
University of New South Wales
a.srikumar@student.unsw.edu.au

MS107

Density Estimation in RKHS with Application to Korobov Spaces in High Dimensions

In this talk, we will consider a kernel method for estimating a probability density function (pdf) from an i.i.d. sample drawn from such density. Our estimator is a linear combination of kernel functions, the coefficients of which are determined by a linear equation. We will present an error analysis for the mean integrated squared error in a general reproducing kernel Hilbert space setting. We will discuss how this theory can be applied to estimate pdfs belonging to weighted Korobov spaces. Under a suitable smoothness assumption, our method attains a rate arbitrarily close to the optimal rate.

Yoshihito Kazashi
École polytechnique fédérale de Lausanne
y.kazashi@uni-heidelberg.de

Fabio Nobile
EPFL, Switzerland
fabio.nobile@epfl.ch

MS107

Monte Carlo and Quasi-Monte Carlo Probability Density Estimation

Monte Carlo and quasi-Monte Carlo methods are widely used and studied for estimating the expectation of some random model X, via several realizations of this model. This data, however, can provide much more information than just the mean and a confidence interval. In fact, it can be used to estimate the entire distribution of X. In this talk, we assume that we can generate observations of X by standard Monte Carlo simulation. For that setting, we introduce novel unbiased probability density estimators based on conditional Monte Carlo and Likelihood Ratio techniques, and compare them with well known standard estimators (Kernel Density estimators). Moreover, we demonstrate when and how we can make use of randomized quasi-Monte Carlo to improve the convergence rate of the mean integrated squared error.

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

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

Art B. Owen
Stanford University

owen@stanford.edu

Amal Ben Abdellah
Université de Montréal
amalbenabdellah@gmail.com

MS108

Interplay Between Isogeometric and Stochastic Collocation for Uncertainty Quantification of Timber Beams

Timber presents numerous defects like knots and variability of grain angle, consequence of the natural grow process of the material. It is well known from the engineering literature that such imperfections deeply affect the mechanical response of structural elements. Therefore, it would be fundamental for engineers and industry to know the pointwise value of grain direction. However, in this regard, the current industrial technology is limited, since the estimation of grain direction is afforded at the surface of the board, only. It is then of primary importance the development of alternative numerical strategies that allow to include and treat the above-mentioned uncertainty in mechanic models. This presentation deals with a prismatic beam behaving in plane stress and made of a linear-elastic orthotropic material with stochastic distribution of material principal directions. To perform UQ on this model we will couple the stochastic collocation method in probability with the isogeometric collocation method in space, namely IGA will be employed to solve each instance of the mechanical deterministic problem. The performed numerical experiments show that the proposed numerical method is a promising tool for timber industry. Indeed, reliable solutions are provided at low computational cost and time scales compatible with the speed of industrial processes.

Giuseppe Balduzzi
Institute of Materials Resource Management
University of Augsburg
giuseppe.balduzzi@mrm.uni-augsburg.de

Francesca Bonizzoni
University of Augsburg
francesca.bonizzoni@math.uni-augsburg.de

Alessandro Reali
Università di Pavia
alereali@unipv.it

Lorenzo Tamellini
Istituto di Matematica Applicata e Tecnologie Informatiche
tamellini@imati.cnr.it

MS108

Isogeometric Multilevel Quadrature for Forward and Inverse Random Acoustic Scattering

We study the numerical solution of forward and inverse acoustic scattering problems by randomly shaped obstacles in three-dimensional space using a fast isogeometric boundary element method. Within the isogeometric framework, realizations of the random scatterer can efficiently be computed by simply updating the NURBS mappings which represent the scatterer. This way, we end up with a random deformation field. In particular, we show that the knowledge of the deformation field's expectation and covariance at the surface of the scatterer are already sufficient to model the surface Karhunen-Loève expansion. Leveraging on the isogeometric framework, we utilize multilevel quadrature methods for the efficient approximation of quantities of interest, such as the scattered wave's expectation and variance. Computing the wave's Cauchy data at an artificial, fixed interface enclosing the random obstacle, we can also directly infer quantities of interest in free space. Adopting the Bayesian paradigm, we finally compute the expected shape and the variance of the scatterer from noisy measurements of the scattered wave at the artificial interface. Numerical results for the forward and inverse problem are given to demonstrate the feasibility of the proposed approach.

Jürgen Dölz
University of Bonn
doelz@ins.uni-bonn.de

Helmut Harbrecht
Universitaet Basel
Department of Mathematics and Computer Science
helmut.harbrecht@unibas.ch

Carlos Jerez
Universidad Adolfo Ibáñez, Santiago de Chile
carlos.jerez@uai.cl

Michael Multerer
Università della Svizzera italiana
michael.multerer@usi.ch

MS108

Density Estimation in Uncertainty Propagation - Approximating Pushforward Measures

In many scientific areas, uncertainty propagation is employed to account for the effect of uncertain parameters in an otherwise deterministic model. Traditionally, the analysis of such problems is done through the lens of moment approximation. However, in many applications, the "full statistics" are required, i.e., we wish to approximate the probability density function (PDF). Underlying this computational problem is a fundamental question - if two "similar" functions pushforward the same measure, would the new resulting measures be close, and if so, in what sense? We will show how the PDF of the quantity of interest can be approximated, first using a spline-based method and then using spectral methods, both with theoretical guarantees. We will then present an alternative viewpoint: through optimal transport theory, a Wasserstein-distance formulation of our problem yields a much simpler and widely applicable theory.

Amir Sagiv
Columbia University
as6011@columbia.edu

Adi Ditkowski
Department of Applied Mathematics
Tel-Aviv University, Israel
adid@post.tau.ac.il

Gadi Fibich
Tel Aviv University
School of Mathematical Sciences

fibich@tau.ac.il

MS108

Shape Deformations for Maxwells Eigenproblem in an Isogeometric Setting

In many applications, splines are a beneficial choice for the spacial discretization for a geometric domain, when investigating shapes under uncertainties. In accelerator cavities, particularly the eigenvalues are sensitive to uncertainties in the shape of the geometry. Since the synchronization between the electromagnetic field and the accelerated particle is crucial, it is important to control the eigenfrequencies. Employing Isogeometric Analysis (IGA) for the spacial discretization of the cavity allows for an exact representation of the geometry and smooth fields for subsequent particle tracking. Furthermore, when considering deformations of the geometry remeshing can be avoided. When investigating the eigenvalues along shape deformations, crossings of the eigenvalues can occur. Hence, for the uncertainty quantification of eigenvalue problems, more sophisticated methods are necessary. In our work, we apply algorithms, which employ derivatives with respect to an introduced deformation parameter, to track the eigenvalues along a deformation [N. Georg et al., UQ for Maxwells eigenproblem based on IGA and mode tracking., CMAME, 350, 228-244, 2019.]. We formulate the required derivatives as shape derivatives with respect to the IGA control points. *Acknowledgement*: *This work is supported by the Graduate School CE within the Centre for Computational Engineering at Technische Universitt Darmstadt.*

Anna Ziegler
Technische Universität Darmstadt
anna.ziegler@tu-darmstadt.de

Melina Merkel
Computational Electromagnetics Group (CEM), TU Darmstadt
Centre for Computational Engineering (CCE), TU Darmstadt
melina.merkel@tu-darmstadt.de

Peter Gangl
Technische Universität Graz
gangl@math.tugraz.at

Niklas Georg
Graduate School CE, TU Darmstadt, Germany
georg@gsc.tu-darmstadt.de

Sebastian Schöps
Computational Electromagnetics Group (CEM), TU Darmstadt
Centre for Computational Engineering (CCE), TU Darmstadt
schoeps@gsc.tu-darmstadt.de

MS109

Optimal Experimental Design and Active Learning Through Objective-Based Uncertainty Quantification

Objective-based uncertainty quantification (objective-UQ) is highly useful in real-world problems that aim to achieve specific scientific or engineering objectives based on complex uncertain systems. The mean objective cost of uncertainty (MOCU) provides effective means of quantifying the impact of model uncertainty on the operational goals at hand, leading to robust optimal experimental design (OED) and active learning strategies for model improvement. In this talk, we discuss the role of objective-UQ in optimal experimental design and active learning based on uncertain scientific models and demonstrate its performance based on several examples. Furthermore, we show how machine learning schemes can be used to accelerate MOCU-based optimal experimental design, resulting in 100 1,000 fold speed-up at virtually no degradation in the OED performance.

Yoon Byung-Jun
Brookhaven National Laboratory, U.S.
bjyoon@tamu.edu

MS109

Tackling Data Fusion and Stochastic Inverse Problems with Statistical Scoring

Fusing probabilistic data represented as finite samples and solving stochastic inverse problems share a common challenge: evaluating the difference between observational datasets and distributions generated by probabilistic models in a computationally tractable way. We propose a strategy commonly used in forecast verification to quantitatively evaluate and rank data sets or different parameter combinations for a probabilistic model to address this challenge. This is typically based on scalar metrics that take as input observational data and samples for a model distribution to be evaluated. We will discuss challenges in the context of fusing high energy particle accelerators data and physics hypotheses and solving inverse problems driven by stochastic differential equations.

Emil M. Constantinescu
Argonne National Laboratory
Mathematics and Computer Science Division
emconsta@mcs.anl.gov

MS109

Smoothing with Transport Maps

Smoothing is a challenging form of data assimilation. Where filters integrate out past states to maintain constant dimensionality, smoothers solve the full inference problem. As new data points are assimilated, the inference horizon widens, appending a new set of state space dimensions with each new time step. This can quickly result in inference over extremely high-dimensional distributions, demanding some form of sparsity to render it tractable. We adopt a transport perspective on smoothing. Transport methods are a set of variational inference techniques which permit the nonlinear conversion of samples from a joint distribution into samples of any of its conditionals. This property can be molded to many inference problems, including smoothing. Formulating a dense transport map for the full data assimilation problem, we find that different established smoothing algorithms can be derived as sparse special cases of this dense map. We then compare the performance of different sparse map configurations over varying inference horizons and different ensemble sizes. We observe that in the case of limited ensemble sizes, common to high-dimensional systems with computationally expensive models, some forms of sparsity - especially those exploiting the structure of the joint distribution's graph - perform substantially better than those who do not. We conclude by exploring these findings in the context of non-

linear smoothing.

Max Ramgraber
Massachusetts Institute of Technology
mramgrab@mit.edu

Ricardo Baptista
MIT
rsb@mit.edu

Dennis McLaughlin
Civil Engineering MIT
dennism@mit.edu

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

MS109

A Generalized Deep Neural Network Model with Continuous Weights

Overparameterization is commonly used because the existing theories indicate that overparameterization will provide more abundant activation patterns and thus lead to better theoretical expressive power. However, recent work shows that the practical expressive power of a neural network is often far from the theoretical one because many redundant neurons may not be activated in all regions after training. One reason is that the standard network model is essentially a discrete architecture, and all the weight and bias components are independent. We propose to develop a continuum network architecture by converting the weight matrices of the standard network model into weight functions. We introduce a set of auxiliary variables and define the networks weights as functions of the auxiliary variables. Training the weight matrices will then become learning the weight functions in a prescribed function family. We demonstrate that the proposed model has higher expressive power than the standard model in function approximation. Additionally, we develop a new automatic differentiation strategy to compute the derivatives of the networks outputs with respect to the inputs, which can significantly improve the efficiency in using networks for PDEs.

Zezhong Zhang, Feng Bao
Florida State University
zz18c@my.fsu.edu, bao@math.fsu.edu

Guannan Zhang
Oak Ridge National Laboratory
zhangg@ornl.gov

MS110

Rates of Posterior Contraction for Direct and Inverse Problems with Besov-Space Priors

B_1^s-Besov priors have been introduced in the applied Bayesian inverse problems literature, due to their sparsity promoting and edge-preserving properties. They are defined via a wavelet basis and employ ℓ_1-type penalization of the coefficients, thus, they are expected to perform well for unknowns with blocky structure as is typical in imaging applications. In this talk, we will discuss the frequentist asymptotic performance of B_1^s-Besov priors, for direct and inverse problems, in terms of rates of posterior contraction. In particular, we will first substantiate the intuitively expected advantage of such priors over Gaussian priors for blocky unknowns, in the context of direct problems. Then, we will establish rates of contraction for nonlinear PDE inverse problems under B_1^s-Besov priors. The talk is based on published and ongoing work, done in collaboration with Masoumeh Dashti, Tapio Helin, Aimilia Savva and Sven Wang.

Sergios Agapiou
University of Warwick
agapiou.sergios@ucy.ac.cy

Masoumeh Dashti
University of Sussex
m.dashti@sussex.ac.uk

Tapio Helin
LUT University
Tapio.Helin@lut.fi

Aimilia Savva
University of Cyprus
savva.emilia@ucy.ac.cy

Sven Wang
Massachusetts Institute of
Technology
svenwang@mit.edu

MS110

Neural Networks in Infinite Dimensions

A general framework for data-driven approximation of input-output maps between infinite-dimensional spaces is developed. Motivated by the recent successes of neural networks, the proposed approach uses a combination of ideas from deep learning and model reduction. This combination results in a neural network approximation which, in principle, is defined on infinite-dimensional spaces and, in practice, is robust to the dimension of the finite-dimensional approximations of these spaces required for computation. For large classes of input-output maps, and suitably chosen probability measures on the inputs, convergence of the proposed approximation methodology is proved. Numerically, the effectiveness of the method is demonstrated on classes of parametric PDE problems with applications in reservoir modeling, the deformation of plastic materials, and the turbulent flow of fluids. Convergence and robustness of the approximation scheme with respect to the size of the discretization is established. The method is shown to be faster and more accurate than many existing algorithms in the literature.

Nikola Kovachki
California Institute of Technology
nkovachki@caltech.edu

MS110

On Polynomial-Time Computation of High-Dimensional Posterior Measures by Langevin-Type Algorithms

This talk considers the problem of generating random samples of high-dimensional posterior distributions. The main results consist of non-asymptotic computational guarantees for Langevin-type MCMC algorithms which scale polynomially in key quantities such as the dimension of the model, the desired precision level, and the number of available statistical measurements. As a direct consequence, it is shown

that posterior mean vectors as well as optimisation based maximum a posteriori (MAP) estimates are computable in polynomial time, with high probability under the distribution of the data. These results are complemented by statistical guarantees for recovery of the ground truth parameter generating the data. Our results are derived in a general high-dimensional non-linear regression setting (with Gaussian process priors) where posterior measures are not necessarily log-concave, employing a set of local 'geometric' assumptions on the parameter space, and assuming that a good initialiser of the algorithm is available. The theory is applied to a representative non-linear example from PDEs involving a steady-state Schrdinger equation.

Sven Wang
Massachusetts Institute of Technology
svenwang@mit.edu

MS111

Multi-Objective Robust Bayesian Optimization

The use of Bayesian optimization is widely spread in engineering design to reduce the number of computational expensive simulations. Input uncertainties are important to take into consideration in many real-life design optimization problems due to, e.g., manufacturing tolerances. Analyzing and propagating input uncertainty is an important step to ensure an inferred optimal design is still satisfying in reality with high probability. While this problem has been actively investigated in the single-objective Bayesian optimization setting, it is less considered for multi-objective Bayesian optimization. In this talk, we give an overview of several ways we can integrate input uncertainty in the problem formulation. We then introduce a simple yet effective robust multi-objective Bayesian optimization framework to efficiently search for robust optimal solutions. The robustness is considered by utilizing Bayes risk on the objective function, which can be efficiently inferred by a robust Gaussian Process (GP). We also elaborate on specific issues such as the prediction uncertainty of the robust GP which cannot be reduced to zero by solely sampling at a certain input due to the aggregated representation of uncertainty w.r.t input locations. This may lead to unfavorable behavior of myopic acquisition functions. This proposed framework works with flexible input uncertainty distributions, and also can be simply extended to perform parallel robust Bayesian Optimization.

Jixiang Qing
IDLab, Ghent University - imec
iGent, Technologiepark-Zwijnaarde 126, B-9052 Gent, Belgium
jixiang.qing@ugent.be

Tom Dhaene
Department of Information Technology (INTEC)
Ghent University-iMinds, Ghent, Belgium
tom.dhaene@ugent.be

Ivo Couckuyt
University of Ghent
ivo.couckuyt@ugent.be

MS111

Multi-Objective Robust Optimization Using Adaptive Kriging for Problems with Mixed Continuous-Categorical Variables

Accounting for uncertainties is crucial in the design of engineering systems. Various techniques have been developed for design optimization within a probabilistic framework. In this work, we consider simultaneously robust and multi-objective design optimization. While the former allows one to deal with uncertainties affecting the objective function, the latter allows for handling multiple conflicting objectives. Conservative quantiles are used as a single measure of robustness, trading-off the optimality and degree of robustness of the solution. The quantiles are computed using crude Monte Carlo simulation and are embedded within a classical multi-objective optimization algorithm, namely the non-dominated sorting genetic algorithm (NSGA-II). Such an approach is obviously computationally intensive. To alleviate this burden, we consider the use of surrogate models, and more specifically adaptive Gaussian process models. This approach is eventually adapted to problems with mixed continuous-categorical variables. After a validation on analytical examples, the proposed method is applied to building renovation where the goal is to find the optimal renovation strategy that minimizes both the environmental impact and the life cycle cost of a building. This is carried out in the context of life cycle analysis where there are numerous uncertainties that need to be accounted for.

Maliki Moustapha, Bruno Sudret
ETH Zurich
moustapha@ibk.baug.ethz.ch, sudret@ibk.baug.ethz.ch

MS111

A Two-Step Procedure for Time-Dependent Reliability-Based Design Optimization Involving Piece-Wise Stationary Gaussian Processes

We deal with a numerical model-based optimization problem with a simple and fast to evaluate deterministic cost function and probabilistic constraints with very high confidence levels. The main difficulty lies in the estimation of the constraints at each loop of the optimization algorithm. They are expressed as threshold exceedance probabilities of maxima and integrals of temporal random processes over long periods of time. A naive approach such as the Monte Carlo (MC) method requires too many time-consuming simulations to calculate the failure probabilities. We propose a two-step methodology which first uses limit theorems on the integrals and maxima of processes in order to reformulate the constraints in the form of time-independent and faster to evaluate expectations. To further reduce the computational cost, the second step introduce a new active kriging method "Adaptive Kriging method for Expectation Constraint Optimization" (AK-ECO). For each constraint, a metamodel is built in the augmented space which spans the design space and the space of uncertain variables. Then the AK-ECO procedure consists in carrying out cycles of optimization composed of local enrichment of the metamodels with a dedicated learning function followed by a resolution of the reformulated problem using MC with the refined metamodels. This methodology has been applied with success to an academic example, to a wind turbine industrial case and performs much better than state-of-the-art algorithms.

Alexis Cousin
IFP Energies Nouvelles
1-4 Avenue du Bois Préau, 92852 Rueil-Malmaison, France
alexis.cousin@ifpen.fr

Josselin Garnier
Ecole Polytechnique
josselin.garnier@polytechnique.edu

Martin Guiton
IFP Energies Nouvelles
1-4 Avenue du Bois Préau, 92852 Rueil-Malmaison, France
martin.guiton@ifpen.fr

Miguel Munoz Zuniga
IFP Energies Nouvelles
miguel.munoz-zuniga@ifpen.fr

MS111

Sampling Criteria for Constrained Bayesian Optimization under Uncertainty

We consider the problem of chance constrained optimization where the objective and the constraint functions are affected by uncertainties and are computationally costly. Bayesian optimization is an appropriate family of methods to address such problems. We first propose a two-step acquisition criterion defined in the joint space of optimization variables and uncertain parameters. The objective and the constraints are aggregated through a feasible improvement measure and the two steps consist in the optimization of the expectation and the one-step-ahead variance of this criterion. To ease the computational burden, an analytical approximation to the one-step-ahead variance is proposed. Additionally, we also account for the possible correlation between the constraints. This is done by considering a vector-valued "input as output" joined Gaussian process which improves the constraints modeling accuracy and consequently the optimization procedure. The correlations between the constraints are further exploited by allowing each constraint to be evaluated for different uncertain parameters and by optimally selecting a subset of constraints to be evaluated at each iteration, thus avoiding unnecessary computations. Numerical tests confirm the applicability and potential gains brought by these methods, such a faster convergence speed and better scaling with respect to the number of constraints if compared to alternative optimization methods.

Julien Pelamatti
EDF R&D, 6 quai Watier, 78 401, Chatou, France
julien.pelamatti@edf.fr

Mohamed Reda El Amri
CEA DER/SESI/LEMS, France
mohamed-reda.elamri@cea.fr

Rodolphe Le Riche
Ecole des Mines de St Etienne
leriche@emse.fr

Christophette Blanchet-Scalliet
École Centrale de Lyon, Institut Camille Jordan
christophette.blanchet@ec-lyon.fr

Céline Helbert
Institut Camille Jordan Ecole Centrale Lyon
celine.helbert@ec-lyon.fr

MS112

Hyperparameter Tuning is All You Need for LISTA

Learned Iterative Shrinkage-Thresholding Algorithm (LISTA) introduces the concept of unfolding an iterative algorithm and trains it like a neural network. It had great success on sparse recovery. In this paper, we show that adding momentum to the LISTA network achieves a better convergence rate and, in particular, the network with instance-optimal parameters is superlinearly convergent. Moreover, our new theoretical results lead to a practical approach of automatically and adaptively calculating the parameters of a LISTA network layer based on its previous layers. Perhaps most surprisingly, such an adaptive-parameter procedure reduces the training of LISTA to tuning only three hyperparameters from data: a new record set in the context of the recent advances on trimming down LISTA complexity. We call this new ultra-light weight network HyperLISTA. Compared to state-of-the-art LISTA models, HyperLISTA achieves almost the same performance on seen data distributions and performs better when tested on unseen distributions (specifically, those with different sparsity levels and nonzero magnitudes).

Jialin Liu
Alibaba DAMO academy
jialin.liu@alibaba-inc.com

MS112

Learning to Predict Nash Equilibria from Data Using Deep Equilibrium Networks

We study the problem of predicting the outcome of a contextual game, given only the context, and assuming that the player's cost functions are unknown. We use the recently introduced Deep Equilibrium Model (DEQ) framework to phrase this as a learning problem using historical data consisting of pairs of context and game outcomes. Using several "tricks" (e.g. Davis-Yin operator splitting, constraint decoupling) we improve the efficiency of this scheme to the extent that it can be readily applied to large games with complicated constraint sets. Finally, we demonstrate the efficacy of this approach on a collection of real-world traffic routing problems.

Daniel McKenzie
University of California, Los Angeles, U.S.
mckenzie@math.ucla.edu

Howard Heaton, Samy Wu Fung, Qiuwei Li
University of California, Los Angeles
hheaton@ucla.edu, swufung@mines.edu, liqiuweiss@gmail.com

Stanley J. Osher
University of California
Department of Mathematics
sjo@math.ucla.edu

Wotao Yin
Alibaba Group (US)
Damo Academy

wotaoyin@math.ucla.edu

MS112

A Symbolic Approach for Learning to Optimize

Recent studies on Learning to Optimize (L2O) suggest a promising path to automating and accelerating the optimization procedure for complicated tasks. Existing L2O models parameterize optimization rules by neural networks, and learn those numerical rules via meta-training. However, they face two common pitfalls: (1) scalability: the numerical rules represented by neural networks create extra memory overhead for applying L2O models, and limits their applicability to optimizing larger tasks; (2) interpretability: it is unclear what each L2O model has learned in its black-box optimization rule, nor is it straightforward to compare different L2O models in an explainable way. To avoid both pitfalls, I will discuss our recent work showing that we can "kill two birds by one stone", by introducing the powerful tool of symbolic regression to L2O. We establish a holistic symbolic representation and analysis framework for L2O, which yields a series of insights for learnable optimizers. Leveraging our findings, we further propose a lightweight L2O model that can be meta-trained on large-scale problems and outperformed human-designed and tuned optimizers.

Atlas Wang
University of Texas, Austin
atlaswang@utexas.edu

MS112

Efficient Training of Infinite-Depth Neural Networks via Jacobian-Free Backpropagation

A promising trend in deep learning replaces fixed depth models by approximations of the limit as network depth approaches infinity. This approach uses a portion of network weights to prescribe behavior by defining a limit condition. This makes network depth implicit, varying based on the provided data and an error tolerance. Moreover, existing implicit models can be implemented and trained with fixed memory costs in exchange for additional computational costs. In particular, backpropagation through implicit depth models requires solving a Jacobian-based equation arising from the implicit function theorem. We propose a new Jacobian-free backpropagation (JFB) scheme that circumvents the need to solve Jacobian-based equations while maintaining fixed memory costs. This makes implicit depth models much cheaper to train and easy to implement. Numerical experiments on classification, CT reconstructions, and predicting traffic models are provided.

Samy Wu Fung
University of California, Los Angeles
swufung@mines.edu

MS113

Proportional Marginal Effects for Sensitivity Analysis with Correlated Inputs

Global sensitivity analysis of a numerical model aims at quantifying importance measures for the model inputs on a statistical quantity of interest related to a model output. In case of independent inputs, the Sobol' indices associate to each input a percentage of the output's variance. For the dependent inputs case, the Shapley value concept (coming from cooperative game theory in order to distribute gains between players), has been associated to Sobol indices, leading to the so-called Shapley effects. However, a first drawback (the so-called Shapley's joke) is that an input not included in the model can be associated to a strictly positive index if it is correlated to an input present in the model. More generally, a lack of discrimination between the influence of the inputs is observed due to the underlying equitable redistribution principle. In this work, we use another game theory allocation rule called proportional value. This result is not originally defined for cost functions with null values, but an extension to these cases is proposed. Based on these results, novel sensitivity indices are proposed: the proportional marginal effects (PME). The PME do not fall under the Shapley's joke and have a much larger power of discrimination between inputs than Shapley effects. These behaviors are proved in the general case and are studied through toy-cases. An estimation strategy is also presented, with application to an industrial use-case.

Margot Hérin
Sorbonne University
margot.herin@polytechnique.edu

Marouane Il Idrissi, Vincent Chabridon, Bertrand Iooss
EDF R&D
marouane.il-idrissi@edf.fr, vincent.chabridon@edf.fr, bertrand.iooss@edf.fr

MS114

Nonstationary Seasonal Model for Daily Mean Temperature Distribution Bridging Bulk and Tails

In traditional extreme value analysis, the bulk of the data is ignored, and only the tails of the distribution are used for inference. Extreme observations are specified as values that exceed a threshold or as maximum values over distinct blocks of time, and subsequent estimation procedures are motivated by asymptotic theory for extremes of random processes. For environmental data, nonstationary behavior in the bulk of the distribution, such as seasonality or climate change, will also be observed in the tails. To accurately model such nonstationarity, it seems natural to use the entire dataset rather than just the most extreme values. It is also common to observe different types of nonstationarity in each tail of a distribution. Most work on extremes only focuses on one tail of a distribution, but for temperature, both tails are of interest. This paper builds on a recently proposed parametric model for the entire probability distribution that has flexible behavior in both tails. We apply an extension of this model to historical records of daily mean temperature at several locations across the United States with different climates and local conditions. We highlight the ability of the method to quantify changes in the bulk and tails across the year over the past decades and under different geographic and climatic conditions. The proposed model shows good performance when compared to several benchmark models that are typically used in extreme value analysis of temperature.

Mitchell Krock
Rutgers University
mk1867@stat.rutgers.edu

Julie Bessac
Argonne National Laboratory
jbessac@anl.gov

Michael Stein

Rutgers University
ms2870@stat.rutgers.edu

Adam Monahan
University of Victoria
School of Earth and Ocean Science
monahana@uvic.ca

MS114

Divide-and-Conquer Methods for Extreme Value Analysis of Large Spatial Datasets

Extreme weather events frequently exhibit spatial and temporal dependence that is computationally prohibitive to model for as few as a dozen observations, with supposed computationally efficient approaches like the composite likelihood remaining computationally burdensome with a few hundred observations. In this paper, we propose a partitioning approach based on local modelling of subsets of the spatial domain that delivers a computationally and statistically efficient procedure. Marginal and dependence parameters are estimated locally on subsets of observations using censored pairwise composite likelihood, and combined using a modified Generalized Method of Moments procedure. We demonstrate consistency and asymptotic Normality of estimators, and show empirically that this approach leads to a surprising reduction in bias of parameter estimates over a full data approach.

Brian J. Reich, Emily Hector
North Carolina State University
brian_reich@ncsu.edu, ehector@ncsu.edu

MS114

Estimating Ground-Level Ozone Exceedance Probabilities Using Non-Stationary Extreme Value Modeling Procedures

Ground-level ozone is a harmful air pollutant whose negative effects are intensified when it is at its most extreme levels. The United States Environmental Protection Agency monitors ozone levels throughout the country, and Federal legislation sets air quality standards that localities must meet in order to stay in compliance. These standards are written in terms of the three-year average of the fourth-highest daily eight-hour ozone value, and this maximum allowable value has been changed multiple times through recent years. In this work, we develop methods to estimate exceedance probabilities for surface-level ozone extremes. Our methods are built within the framework provided by extreme value theory; to allow for the possibility of changing ozone extremes, our modeling procedure incorporates non-stationarity. We apply our approach to analyze surface-level ozone data from several US locations over the last 25 years, and use these data to estimate exceedance probabilities in the context of the US ozone standards. We find that ozone extremes have been decreasing over this time period at many locations, and these decreases seem to be related to decreases in levels of ozone precursors.

Brook Russell, Jax Li, Whitney K. Huang
Clemson University
brookr@clemson.edu, jaxl@g.clemson.edu, wkhuang@clemson.edu

MS114

Removing the Influence of Daily Weather Systems on Detected Changes in Precipitation Extremes

The detection of changes over time in the distribution of precipitation extremes is significantly complicated by noise at the spatial scale of daily weather systems. Traditional approaches for quantifying observed trends in extreme precipitation are generally based on single-station analyses which fail to account for the spatial coherence of individual storms and hence yield unrealistic and potentially misleading estimates of the true underlying changes in extremes. In this paper, we demonstrate how the use of a flexible statistical method that robustly accounts for the so-called storm dependence in measurements of daily precipitation removes a challenging source of noise and results in improved estimates of trends in precipitation extremes. Applying the methodology to long-term in situ records of daily precipitation from the central United States, we find that properly accounting for storm dependence furthermore leads to increased detection of statistically significant trends relative to existing approaches. Additionally, the approach allows us to calculate changes in the risk of concurrent extreme precipitation, a quantity that single-station analyses cannot provide.

Likun Zhang, Mark Risser
Lawrence Berkeley National Laboratory
likunz@lbl.gov, mdrisser@lbl.gov

MS115

PINNs for Solving Forward and Inverse Problems Governed by Stochastic-Fractional PDEs

We consider solutions of forward and inverse problems governed by stochastic-fractional PDEs (SFPDEs), and in particular we target long-time integration. To this end, we develop a Physics-Informed Neural Network (PINN), and propose a new extension for SFPDEs, which incorporates the bi-orthogonal constraints of stochastic modeling into the loss function with an implicit form. This approach can overcome some of the drawbacks of the original bi-orthogonal methods for time dependent SDEs. We will demonstrate the efficiency of the new network via several numerical examples, demonstrating how intrigate is the interplay between stochasticity and fractional order and designing new PINN features to deal with the multiscale solutions.

Rongxin Li
Shanghai Normal Univeristy
1000478817@smail.shnu.edu.cn

Ling Guo
Shanghai Normal University
lguo@shnu.edu.cn

George E. Karniadakis
Brown University
Division of Applied Mathematics
george_karniadakis@brown.edu

MS115

Data-Based Model Reduction and Mori-Zwanzig Formalism for Random Dynamical Systems

The Mori-Zwanzig (MZ) projection operator formalism provides a general framework for constructing reduced models for dissipative chaotic and stochastic dynamical systems, particularly in situations without sharp scale separation and hence significant memory effects. In this talk,

I will report on an approach to data-based model reduction for stochastic and random dynamical systems using the NARMAX (Nonlinear Auto-Regressive Moving Average with eXogenous inputs) representation of stochastic processes, widely used in time series analysis and data-driven modeling. I will explain how the NARMAX approach may be formally derived from the original dynamical model using a discrete-time version of the MZ formalism. These ideas are illustrated on a stochastically-forced PDE. Time permitting, I will also discuss our recent efforts to improve the efficiency and scalability of our data-based modeling procedure.

Kevin K. Lin
Department of Mathematics
University of Arizona
klin@math.arizona.edu

Fei Lu
John Hopkins University
flu15@jhu.edu

MS115

Markov Chain Generative-Adversarial Neural Networks for Solving Bayesian Inverse Problems

State and parameter estimation based on noisy measurements is known to be computationally expensive. This is especially the case when one is not only interested in the estimate but also the uncertainty of said estimate. In such cases the aim is to compute a posterior distribution over the full state and corresponding parameters conditioned on available observations. Due to the complexity of such posteriors, it is common practice to resort to sampling methods such as Markov Chain Monte Carlo (MCMC) methods. While MCMC methods are known to be efficient, it is still typically required to sample 50000 or more times before the chain has converged. This is especially the case when dealing with high-dimensional problems. Since every sample requires solving the forward PDE, it is, in general, infeasible to sample enough times if one wants the posterior approximation in real-time. We propose a method to speed up this procedure by replacing the forward model and prior with a generative adversarial network. The state and parameter space are being replaced with a latent space that is pushed forward using the GAN. This results in a methodology where one samples from a low-dimensional latent space and learn the forward map in an offline stage. We showcase the methodology on problems from fluid mechanics.

Nikolaj T. Mücke
Centrum Wiskunde & Informatica (CWI) and Utrecht University
nikolaj.mucke@cwi.nl

Cornelis W. Oosterlee
CWI, Centrum Wiskunde & Informatica, Amsterdam
c.w.oosterlee@cwi.nl

Sander Bohte
CWI, Centrum Wiskudne & Informatica
s.m.bohte@cwi.nl

Benjamin Sanderse
Centrum Wiskunde & Informatica (CWI)
Amsterdam, the Netherlands
b.sanderse@cwi.nl

MS115

Probabilistic Machine Learning and Likelihood for Learning Stochastic Parameterizations for Climate Modelling

Machine Learning (ML) revolutionised generative models, such as in the natural language domain. Climate models are ultimately generative models we use them to generate potential climatic trajectories. How can the insights from the ML literature in generative modelling be combined with those from the climate parameterization field to improve climate models? Parameterization means modelling processes which are too small-scale to be explicitly resolved in climate models. Our work tackles this by using generative models and probabilistic ML to invent hidden states to better model the climate. The developed models offer improvements over existing baselines in the Lorenz 96 simulator, permitting more realistic simulations. By using a probabilistic framework and likelihood (a standard procedure from probability modelling), many state-of-the-art ML tools can be easily deployed for stochastic climate modelling. Likelihood also allows for a standardized probabilistic evaluation procedure for both ML-trained and human-designed parameterizations. But it is known that models with good likelihoods can produce poor samples. By examining cases from our work, we consider how likelihood should be used to evaluate such dynamical systems, and when a good likelihood will correspond to physically consistent samples.

Raghul Parthipan, Damon Wischik
University of Cambridge
rp542@cam.ac.uk, damon.wischik@cl.cam.ac.uk

Scott Hosking
British Antarctic Survey, Cambridge
jask@bas.ac.uk

Hannah M. Christensen
University of Oxford, United Kingdom
Hannah.Christensen@physics.ox.ac.uk

MS116

Operator Networks with Predictive Uncertainty for Partial Differential Equations with Inhomogeneous Boundary Conditions

We present a neural network training procedure for solving partial differential equations with inhomogeneous boundary conditions. Using a light-weight extension of the DeepONet operator network architecture, the trained networks are designed to provide rapid predictions along with simultaneous uncertainty estimates to help identify potential inaccuracies in the network predictions. In particular, the predictive uncertainty of the network is calibrated to anticipate network errors by implementing a loss function which interprets the network prediction as a probability distribution as opposed to a single point-estimate. The proposed technique is also capable of solving problems on irregular, non-rectangular domains, and a series of experiments are presented to evaluate the network accuracy as well as the quality of the predictive uncertainty estimates.

Guang Lin, Nick Winovich
Purdue University
guanglin@purdue.edu, nwinovi@sandia.gov

Lu Lu
University of Pennsylvania
lulu1@seas.upenn.edu

MS116

Learning Deep Nonlocal Operator for Heterogeneous Material Modeling

Constitutive modeling based on the continuum mechanics theory has been a classical approach for modeling the mechanical responses of materials. However, when constitutive laws are unknown or when defects and/or high degrees of heterogeneity present, these classical models may become inaccurate. In this work, we propose to use data-driven modeling which directly utilizes high-fidelity simulation and/or experimental measurements on displacement fields, to predict a material's response without the necessity of using conventional constitutive models. Specifically, the material response is modeled by learning maps between loading conditions and its resultant displacement fields, so that the network is a surrogate for a solution operator. To model the complex material responses, we develop a novel deep neural operator architecture based on the Fourier Neural Operator (FNO) method, which we coin DeepFNO. In DeepFNO, we model the increment between layers as an integral operator, to capture long-range dependencies in the feature space and allow for accelerated learning techniques for deep networks. We demonstrate the performance of our method for a number of examples, including hyperelastic, anisotropic and brittle fracture materials. As an application, we employ the proposed approach to learn material models directly from digital image correlation (DIC) tracking measurements, and show that the learnt solution operators substantially outperform conventional constitutive models.

Yue Yu
Department of Mathematics, Lehigh University
yuy214@lehigh.edu

Huaiqian You, Quinn Zhang
Lehigh University
huy316@lehigh.edu, quz222@lehigh.edu

Colton Ross, Chung-Hao Lee
The University of Oklahoma
cjross@ou.edu, ch.lee@ou.edu

MS116

Convergence Rates of DeepOnets for Operator Regression

As an operator regression technique, DeepONet has been successfully applied in simple and complex systems. Yet, it is unclear why deepONet works well. In this talk, we present error estimates of DeepOnets for generic Hölder continuous operators and solution operators from both linear and nonlinear advection-diffusion equations. We find that the convergence rates depend on the architecture of branch networks and the smoothness of inputs and outputs of solution operators.

Zhongqiang Zhang
Worcester Polytechnic Institute
zzhang7@wpi.edu

MS117

Higher-Order Statistics Using Machine Learning Second-Order Closures for Turbulent Dynamical Systems

This work analyzes the statistics of turbulent systems via a data-informed and reduced-order modelling scheme. To achieve this, we first derive the dynamical equations for the mean and covariance of such systems. These equations are then complemented and evolved together with a spatio-temporally nonlocal, neural-network-based closure for the higher-order statistics. During the training of these neural-networks appropriate physical constraints are imposed. In more detail, energy preservation constraints ensure that energy transfers between stochastic modes are properly modelled, resulting in the numerical stability of the system. These constraints are shown to be essential for the model to correctly capture the statistical equilibria of the full reference system. The validity of this formulation is showcased in a multitude of different systems used to study turbulence in the ocean and atmosphere. Comparisons with direct numerical simulations are carried out both for the mean and energy spectrum as well as the probability of occurrence of intermittent extreme events. These numerical investigations are carried out both for flows with Gaussian and strongly non-Gaussian statistics.

Alexis-Tzianni Charalampopoulos
Massachusetts Institute of Technology
alexchar@mit.edu

Themistoklis Sapsis
Massachusetts Institute of Techonology
sapsis@mit.edu

MS117

Quantifying and Reducing Uncertainty from HED Experiments Using Bayesian Optimal Experimental Design

High Energy Density (HED) science is the study of the behaviour of material under extreme conditions of temperature and pressure. Understanding the growth and properties of hydrodynamic instabilities and the transition into turbulence is important in many HED processes and also yields insights into other areas where hydrodynamic instabilities occur. Unlike classical fluids experiments, examining hydrodynamic instabilities in an HED regime is much more difficult. HED experiments are expensive and often performed at oversubscribed facilities. Additionally, there are many limitations for the available diagnostics. Further, these problems contain complex and interaction physics: hydrodynamics flows, radiation, conduction, and magnetic fields, only to name a few. Such complexity means that modelling can become prohibitively expensive. Improving the models from limited experimental and high-fidelity simulation data is therefore of great importance.A big question that exists within the context of improving simulations is: how do we create a general model that can both accurately capture the underlying physics and be applicable to more than one problem? In an attempt to answer this question, we incorporate experimental design and uncertainty quantification techniques to connect experimental data with simulations and take a statistical approach in order to efficiently plan experiments so that the data obtained can be analysed to yield valid and objective conclusions.

Codie Kawaguchi, Xun Huan
University of Michigan

cfkawa@umich.edu, xhuan@umich.edu

MS117

Gradient-Free Optimization with Bayesian Echo State Networks: Applications in Acoustics

We develop a versatile optimization method, which finds the design parameters that minimize time-averaged acoustic cost functionals. The method is gradient-free, model-informed, and data-driven with reservoir computing based on echo state networks. First, we analyse the predictive capabilities of echo state networks both in the short- and long-time prediction of the dynamics. We find that both fully data-driven and model-informed architectures learn the chaotic acoustic dynamics, both time-accurately and statistically. Informing the training with a physical reduced-order model with one acoustic mode markedly improves the accuracy and robustness of the echo state networks, whilst keeping the computational cost low. Second, we couple echo state networks with a Bayesian technique to explore the design thermoacoustic parameter space. The computational method is minimally intrusive. Third, we find the set of flame parameters that minimize the time-averaged acoustic energy of chaotic oscillations, which are caused by the positive feedback with a heat source, such as a flame in gas turbines or rocket motors. The optimal set of flame parameters is found with the same accuracy as brute-force grid search, but with an order of magnitude faster convergence rate. This work opens up new possibilities for non-intrusive optimization of chaotic systems, in which the cost of generating data, for example from high-fidelity simulations and experiments, is high.

Francisco Huhn
University of Cambridge
fh360@eng.cam.ac.uk

Luca Magri
Imperial College London
l.magri@imperial.ac.uk

MS117

Estimating Floating-Point Errors Using Automatic Differentiation

Floating-point errors are a testament to the finite nature of computing and if left uncontrolled they can have catastrophic results. As such, for high-precision computing applications, quantifying these uncertainties becomes imperative. There have been significant efforts to mitigate such errors by either extending the underlying floating-point precision, using alternate compensation algorithms or estimating them using a variety of statistical and non-statistical methods. A prominent method of dynamic floating-point error estimation is using Automatic Differentiation (AD). However, most state-of-the-art AD-based estimation software requires manually adapting or annotating the source code by some amount. Moreover, operator overloading AD based error estimation tools call for multiple gradient recomputations to report errors over a large variety of inputs and suffer from all the shortcomings of the underlying operator overloading strategy such as reduced efficiency. In this work, we propose a customizable way to use AD to synthesize source code for estimating uncertainties arising from floating-point arithmetic in C/C++ applications. Our work presents an automatic error annotation framework that can be used in conjunction with custom user-defined error models. We also present our progress with error estimation on GPU applications.

Vassil Vassilev
Princeton University, CERN
vasil.georgiev.vasilev@cern.ch

Garima Singh
Manipal Institute of Technology
garimasingh0028@gmail.com

MS118

Self-Assembly of Hydrocarbons

Molecular dynamics simulations allow us to get an experimental insight into self-assembly of hydrocarbons. These simulations showed the generation of over 7 thousand chemical species and over 13 thousand of unique chemical reactions. Unfortunately, these simulations a very expensive computationally as chemical reactions occur rarely on the timescale of molecular motion. We propose an approach relying on graph data analysis and Markov chains. Combined with the experimental data and give a comprehensive coarse-grained description of processes taking place in this chemical system.

Maria K. Cameron
University of Maryland
mariakc@umd.edu

Vincent Decieux, Evan Reed
Stanford University
vdufourd@stanford.edu, evanreed@stanford.edu

Christopher Moakler
University of Maryland
cmoakler@umd.edu

MS118

Methods for Extreme Quantile Regression in High Dimensions

Quantile regression relies on minimizing the conditional quantile loss, which is based on the quantile check function. This has been extended to flexible regression functions such as the gradient forest (Athey et al., 2019). These methods break down if the quantile of interest lies outside of the range of the data. Extreme value theory provides the mathematical foundation for estimation of such extreme quantiles. A common approach is to approximate the exceedances over a high threshold by the generalized Pareto distribution. For conditional extreme quantiles, one may model the parameters of this distribution as functions of the predictors. Up to now, the existing methods are either not flexible enough or do not generalize well in higher dimensions. We develop two new approaches for extreme quantile regression that estimate the parameters of the generalized Pareto distribution in a flexible way even in higher dimensions. The first approach is based on gradient boosting and the second one on random forests. These estimators outperform classical quantile regression methods and methods from extreme value theory in simulations studies. We illustrate the methodology at the example of U.S. wage data.

Sebastian Engelke
University of Geneva
sebastian.engelke@unige.ch

Jasper Velthoen

Delft University of Technology
j.j.velthoen@tudelft.nl

Clement Dombry
Universite Bourgogne Franche-Comte
clement.dombry@univ-fcomte.fr

Juan-Juan Cai
Vrije Universiteit Amsterdam
j.cai@vu.nl

Edossa Merga Terefe, Nicola Gnecco
University of Geneva
edossamerga@gmail.com, nicola.gnecco@unige.ch

MS118

Exploiting Self-Similarity in Distribution Tails for Minimizing Extreme Risks with Limited Data

Formulating optimization objectives or constraints in terms of tail risk measures such as VaR/CVaR often incurs a steep-price in sample complexity, primarily due to (i) the rarity with which relevant risky samples are witnessed, and (ii) the resulting amplification of estimation errors while performing optimization. To tackle this difficulty, we introduce a novel tail modeling framework built on the ubiquitous, yet less-exploited phenomenon, namely, self-similarity in tail distributions. Interestingly, this similarity in tails translates to similarity in optimal decisions at different tail levels, and can be utilized to re-formulate the problem as an importance-weighted optimization formulation with low variance. While estimation of the nuisance importance weight function accurately is arguably a data-intensive exercise, we show that errors in its estimation can be canceled by exploiting the structure in tail models. This debiasing exercise leads to consistent decisions with exponentially fewer data samples than required by empirical risk minimization.

Anand Deo
Singapore University of Technology and Design
deo_avinash@sutd.edu.sg

Karthyek Murthy
Singapore University of Technology and Design (SUTD)
karthyek_murthy@sutd.edu.sg

MS118

Data-Driven Rare Event Simulation for Stochastic Dynamical Systems: A Koopman Operator Approach

We present a data-driven approach for designing multilevel splitting schemes for rare event simulation in nonlinear stochastic dynamical systems. Multilevel splitting methods are known to be more efficient than standard Monte Carlo while being more robust than importance sampling for rare event simulation. The method is more efficient at producing trajectories that reach rare events by splitting promising trajectories that move towards the desired regions. The approach requires determining the splitting locations, which are defined by level sets of a so-called importance function. The optimal importance function is related to the solution of a Hamilton-Jacobi-Bellman (HJB) PDE, which typically requires knowledge of the model. We exploit the relationship between the stochastic Koopman operator (sKO) and the HJB to approximate the optimal importance function for splitting. We rely on data-driven methods for computing eigenfunctions of the sKO, such as dynamic mode decomposition and diffusion maps, so that the resulting algorithm is completely black box. The approach is demonstrated on a variety of canonical dynamical systems.

Benjamin J. Zhang
Massachusetts Institute of Technology
bjz@mit.edu

Quan Long
United Technologies Research Center
longq@utrc.utc.com

Joshua White
Massachusetts Institute of Technology
jkwhite@mit.edu

Tuhin Sahai
Raytheon Technologies Research Center
tuhin.sahai@rtx.com

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

MS119

Surrogate Modeling for Rare Event Simulation of Material Fatigue

Rare events are expected to occur infrequently (for example, order of 10^{-3} or less) according to a probability model. In uncertainty quantification, the rare events often correspond to failure of systems designed for high reliability and their probability is mathematically represented as multivariate integral over a high dimensional space of uncertain parameters. Standard Monte Carlo simulation for computing such rare-event probabilities is computationally inefficient as it requires sampling from the (usually expensive) system response a prohibitive number of times. Alternative methods include importance sampling, subset simulation and splitting. In this work, we present a hybrid approach by introducing surrogate models for the system response and sampling both the surrogate (cheap) model in a "large: portion of the probability space and the original (expensive) system in a "small" portion. We explore a variety of linear and non-linear surrogate models. We apply our model-agnostic technique for calculating probabilities of rare events that involve crack nucleation of titanium alloy materials found in modern aircraft engines.

Dimitris Konomis, Youssef M. Marzouk
Massachusetts Institute of Technology
dkonomis@mit.edu, ymarz@mit.edu

MS119

Non-Intrusive Parametric Reduced-Order Modeling via Operator Inference

We formulate a new approach to reduced modeling of parameterized, time-dependent partial differential equations (PDEs). The method employs Operator Inference, a scientific machine learning framework combining data-driven learning and physics-based modeling. The parametric structure of the governing equations is embedded directly into the reduced-order model, and parameterized reduced-order operators are learned via a data-driven linear regression problem. The result is a reduced-order model that

can be solved rapidly to map parameter values to approximate PDE solutions. Such parameterized reduced-order models may be used as physics-based surrogates for uncertainty quantification and inverse problems that require many forward solves of parametric PDEs. Numerical issues such as well-posedness and the need for appropriate regularization in the learning problem are considered, and an algorithm for hyperparameter selection is presented. The method is illustrated for a parametric heat equation and demonstrated for the FitzHugh-Nagumo neuron model.

Shane A. McQuarrie
University of Texas at Austin, U.S.
shanemcq@utexas.edu

Parisa Khodabakhshi, Karen E. Willcox
UT Austin
parisa@austin.utexas.edu, kwillcox@oden.utexas.edu

MS119

Versatile Gaussian Process and Bayesian Optimization with Computational Materials Science Applications

Gaussian process (GP) has been one of the cornerstones in non-parametric Bayesian machine learning methods. Using GP as an underlying surrogate model, Bayesian optimization (BO) aims to balance exploration and exploitation and refine the GP model as the optimization progresses. With a solid mathematical foundation, these two methods have been widely adopted across multiple disciplines. While they are among one of the most popular data-driven approaches, there are many limitations in the classical GP and BO that do not naturally fit in the practical settings. In this talk, we discuss a wide range of extensions, including multi-objective, multi-fidelity, mixed-integer, parallel, scalable, and high-dimensional from theoretical and computational perspectives. We conclude the talk with real-world engineering applications for materials and manufacturing.

Anh Tran
Sandia National Laboratories
anhtran@sandia.gov

MS119

Design and Analysis of Finite Element Simulations

Finite Element Analysis (FEA) is a powerful tool in engineering applications, which solves partial differential equations by discretizing the space into a set of finite elements. The numerical accuracy of FEA depends on the number of elements used in the discretization, which can be adjusted. The larger the number of elements the more accurate the results are. However, the computational cost increases with it. In current practice, the experimenter chooses the number of mesh elements that is expected to produce a reasonably accurate result and for which the simulation can be completed in a reasonable amount of time. In this article, we propose an experimental design method that enables the experimenter to select a range of element numbers and complete the full set of simulations within the same time constraints. We also explain how the simulations performed given different numbers of mesh elements can be integrated to produce a predictive model that is much more accurate than a model obtained with same element numbers. We illustrate our approach using an analytical function and a cantilever beam finite element simulation.

Henry Yuchi, Roshan V. Joseph, C. F. Jeff Wu
Georgia Institute of Technology
shaowu.yuchi@gatech.edu, roshan@gatech.edu, jeff.wu@isye.gatech.edu

MS121

Multifidelity Uncertainty Quantification for Nonlocal Models

A wide class of nonlocal models accounts for distant interactions through the use of integral formulations, as opposed to partial differential equations (PDEs). Even though compared to conventional PDE models, nonlocal models are better suited for representing some physical phenomena; their increased computational costs inhibit their use in practical applications. This drawback becomes even more significant for outer-loop applications where numerous model evaluations are required. Multifidelity methods aim at reducing the computational cost of an outer-loop application by splitting the budget between high-fidelity model evaluations (used for unbiasedness and fidelity) and a set of low-fidelity model evaluations (used for speedup). In this study, we use a multifidelity approach for a nonlocal Cahn-Hilliard model to be used in an uncertainty quantification setting. The Cahn-Hilliard model is generally used to describe the process of phase separation for binary alloys, and its nonlocal version can represent sharp interfaces between the two phases (as opposed to the diffuse interface in conventional PDE models).

Parisa Khodabakhshi
UT Austin
parisa@austin.utexas.edu

Olena Burkovska
Technical University of Munich
burkovskao@ornl.gov

Karen E. Willcox
UT Austin
kwillcox@oden.utexas.edu

Max Gunzburger
Florida State University
Department of Scientific Computing
mgunzburger@fsu.edu

MS121

Comparing Multi-Index Stochastic Collocation and Multi-Fidelity Stochastic Radial Basis Functions for Forward Uncertainty Quantification of Ship Resistance

In this talk we present a comparison of two methods for the forward Uncertainty Quantification (UQ) analysis of a passengers ferry advancing in calm water and subject to two operational uncertainties, namely the ship speed and payload. Specifically, the performance of Multi-Index Stochastic Collocation (MISC) and multi-fidelity Stochastic Radial Basis Functions (SRBF) surrogates is assessed. The estimation of the expected value of the (model-scale) resistance to advancement, as well as of its higher order moments and probability density function, are presented and discussed. Both methods need to repeatedly solve the free-surface Navier-Stokes equations for different configurations of the operational parameters. The required CFD simulations are obtained by a multi-grid Reynolds Aver-

aged Navier–Stokes (RANS) equations solver. Both MISC and SRBF use as fidelity levels the intermediate grids employed by the RANS solver. A relevant aspect for the comparison of the two methods is that the CFD simulations are affected by numerical noise, which is due to the iterative algorithm on which the solver is based. In particular, we discuss the impact of the noise on the forward UQ analysis and investigate some strategies to improve the performance of the two methods with respect to this issue.

Chiara Piazzola
Consiglio Nazionale delle Ricerche - Istituto di Matematica Applicata e Tecnologie Informatiche (CNR - IMATI)
chiara.piazzola@imati.cnr.it

Lorenzo Tamellini
Istituto di Matematica Applicata e Tecnologie Informatiche
tamellini@imati.cnr.it

Riccardo Pellegrini, Riccardo Broglia, Andrea Serani
Consiglio Nazionale delle Ricerche
Istituto di Ingegneria del Mare (CNR-INM)
riccardo.pellegrini@inm.cnr.it, riccardo.broglia@cnr.it, andrea.serani@cnr.it

Matteo Diez
CNR-INM
National Research Council-Institute of Marine Engineering
matteo.diez@cnr.it

MS121

Uncertainty Quantification in Computational Modeling of Plasma-Surface Interactions

We perform a global sensitivity analysis with two coupled codes that model the interaction between the plasma boundary and the material surface in fusion energy applications. The coupled codes depend on 19 parameters, including uncertain operating conditions, uncertain material properties and uncertain numerical parameters. We construct a set of generalized polynomial chaos surrogate models that allow us to extract the desired Sobol' sensitivity indices as a post-processing step. We illustrate how the predicted sensitivity indices change when considering both codes in isolation, or when considering the coupled setting. Finally, we investigate if multifidelity polynomial chaos methods can be used to alleviate the computational burden in the coupled setting.

Pieterjan Robbe, Tiernan Casey, Khachik Sargsyan
Sandia National Laboratories
pmrobbe@sandia.gov, tcasey@sandia.gov, ksargsy@sandia.gov

Habib N. Najm
Sandia National Laboratories
Livermore, CA, USA
hnnajm@sandia.gov

MS121

Accelerating Monte Carlo Methods for Random Heterogeneous Media

Uncertainty quantification for fine scale models of random heterogeneous materials is computationally challenging, because, in principle, one needs to resolve the small scale variations for every realization. For some local quantities of interest, however, a good approximation for each sample can be obtained by resolving the microstructure only in some parts of the computational domain and using an upscaled model elsewhere. In this talk, we show an error estimator-driven procedure that exploits this fact to construct a sequence of surrogate models for a given local quantity of interest. These models are then combined in a multilevel framework to accelerate Monte Carlo sampling. Numerical experiments for steady-state heat conduction and linear elasticity on a microstructure generated via a hierarchical procedure show the effectiveness of the proposed algorithm.

Laura Scarabosio
Seminar for Applied Mathematics
ETH Zuerich
scarabosio@science.ru.nl

MS122

Mixtures of Gaussian Process Experts by Sequential Monte Carlo Methods

Gaussian processes exhibit cubic computational complexity due to the need of inverting a full covariance matrix. To circumvent this, mixtures of Gaussian process experts have been considered where data points are assigned to independent experts, reducing the complexity by allowing likelihood estimation to use smaller covariance matrices. Previous approaches have included parallelizable importance sampling based methods. We extend the existing methodology by utilizing nested particle filters to estimate hyperparameters of the individual experts while simultaneously sampling partitions of the data.

Teemu Härkönen
LUT University
teemu.harkonen@lut.fi

MS122

Bayesian Inversion with Hierarchical Random Field Priors: Computational Strategies

Gaussian random fields are popular models for spatially varying uncertainties, arising, e.g., in geotechnical engineering, hydrology, or image processing. A Gaussian random field is fully characterised by its mean and covariance operator. In more complex models these can also be partially unknown. In this case we need to handle a family of Gaussian random fields indexed with hyperparameters. Sampling for a fixed configuration of hyperparameters is already very expensive due to the nonlocal nature of many classical covariance operators. Sampling from multiple configurations increases the total computational cost severely. In this talk we employ parameterised Karhunen-Love expansions and adaptive cross approximations for sampling. To reduce the cost we construct a reduced basis surrogate built from snapshots of Karhunen-Love eigenvectors in the first case. In the second case, we propose a parameterised version of the adaptive cross scheme. In numerical experiments we consider Matrn-type covariance operators with unknown correlation length and standard deviation. Here, we study the approximation accuracy of reduced basis and cross approximation. As an application we consider Bayesian inversion with an elliptic partial differential equation where the logarithm of the diffusion coefficient is a parameterised Gaussian random field. Indeed, we employ Markov chain Monte Carlo on

the reduced space to generate samples from the posterior measure.

Jonas Latz
Heriot-Watt University
j.latz@hw.ac.uk

MS122

Efficient Bayesian Inversion Using Gaussian Priors with Fractional Laplacian

We consider an infinite dimensional Bayesian inverse problem involving Gaussian priors, in which the covariance operator involves the inverse of a fractional Laplacian. We use an integral representation of the fractional power of an elliptic operator, discrete in space using finite elements and the integral using sinc quadrature. We show how to efficiently apply the prior covariance, its square root, and their inverses, by using Krylov solvers for shifted linear systems. We then derive efficient methods to obtain the MAP estimate, and linearized uncertainty estimates such as conditional realizations, and posterior variance. We will also discuss various options for a hierarchical Bayesian formulation and demonstrate the performance of our approach and solvers on test problems from deblurring and X-ray tomography.

Arvind Saibaba
North Carolina State University
saibab@ncsu.edu

Harbir Antil
George Mason University
Fairfax, VA
hantil@gmu.edu

MS122

Uncertainty Quantification by Comparing Point Cloud Densities

In the context of Bayesian statistics it is possible to find a vast number of models for which it is difficult if not impossible to estimate the parameters using classical methods. In the last decade, approaches have emerged that do not use data sets directly but are based on informative statistics obtained from them to allow parametric estimation in more complex cases. In these approaches, however, it must be borne in mind that the quality of the result will be greatly influenced by how much information is lost in mapping the data into the selected statistic. Finding sufficient statistics can be a real problem, especially in the case of complex models. Certainly, there are ways to facilitate the extraction of information from the data set, such as combining different statistics in the creation of likelihood, where each of them is specialized in extracting particular features from the latter. It is therefore interesting to look for a general type of statistics that can extract most of the information, regardless of the particular underlying multidimensional distribution from which the data can be assumed to be sampled. The approach we present could represent a right way in that direction. Our novelty consists in directly comparing the multidimensional distribution of the tested data with those of the reference data set. The obtained statistic is a value that indicates the pointwise density variability between the tested data set and the reference one.

Sebastian Springer
University of Oulu
sebastian.springer@oulu.fi

Heikki Haario
Lappeenranta University of Technology
Department of Mathematics and Physics
heikki.haario@lut.fi

Alessandro Laio
International School for Advanced Studies
Statistical and Biological Physics sector
laio@sissa.it

MS123

Fast Surrogate of 3-D Patient-Specific Computational Fluid Dynamics Using Statistical Shape Modeling and Deep Learning

Optimization and uncertainty quantification (UQ) have been receiving an increasing amount of attention in computational hemodynamics. However, existing methods based on principled modeling and classic numerical techniques have faced significant challenges, particularly when it comes to complex 3D patient-specific shapes in the real world. First, it is notoriously difficult to parameterize the input space of arbitrarily complex 3-D geometries. Second, the process often involves massive forward simulations, which are extremely computationally demanding or even infeasible. We propose a novel deep learning solution to address these challenges and enable scalable geometric UQ and optimization. Specifically, a statistical generative model for 3-D patient-specific shapes will be constructed based on a handful of available baseline patient geometries. An unsupervised shape correspondence solution is used to enable geometric morphing and a compact geometric design space can then be constructed by the statistical generative shape model. In order to build a fast forward map between geometric input space to the solution space of functional information, we propose a supervised deep learning solution. With the fast surrogate model, we ran test cases to demonstrate its application in shape optimization and UQ analysis in a massively scalable manner.

Pan Du, Xiaozhi Zhu, Jian-Xun Wang
University of Notre Dame
pdu@nd.edu, xiaozhi.zhu.60@nd.edu ¡xiaozhi.zhu.60@nd.edu¿, jwang33@nd.edu

MS123

Stochastic Deep Ritz Method for Uncertainty Quantification

Randomness is ubiquitous in natural sciences and modern engineering. The uncertainty is often modeled as a random coefficient in the partial differential equations (PDEs) that describe the physics. Due to the curse of dimensionality, traditional methods for solving random coefficient PDEs, e.g., polynomial chaos and stochastic collocation, are limited to problems of low dimension. For high dimensional problems, deep learning provides a promising alternative. In this work, we recast the random coefficient PDE into a variational problem whose associated energy functional can be used naturally as the error function in the deep learning framework. The resulted stochastic deep Ritz method does not require any data from the solution space and can be solved by using stochastic gradient descent. We validate our approach by solving a family of random elliptic equations in high dimension.

Ting Wang

CISD
US Army Research Laboratory
tingw@udel.edu

Kenneth Leiter, Jaroslaw Knap
U.S. Army Research Laboratory
kenneth.w.leiter2.civ@mail.mil,
jaroslaw.knap.civ@mail.mil

MS123

Learning Based Computational Solid Mechanics

The recent decades have seen various attempts at accelerating the process of developing materials targeted towards specific applications. The performance required for a particular application leads to the choice of a particular material system whose properties are optimized by manipulating its underlying microstructure through processing. The specific configuration of the structure is then designed by characterizing the material in detail, and using this characterization along with physical principles in system level simulations and optimization. These have been advanced by multiscale modelling of materials, high-throughput experimentations, materials databases, topology optimization and other ideas. Still, developing materials for extreme applications involving large deformation, high strain rates and high temperatures remains a challenge. This talk reviews a number of learning based methods that advance the goal of designing materials targeted by specific applications.

Burigede Liu
University of Cambridge
BL377@cam.ac.uk

MS123

Neural Networks with Physics-Informed Architectures and Constraints: Toward Robust and Data-Efficient Learning of Dynamical Systems

Effective inclusion of physics-based knowledge into deep neural network models of dynamical systems can greatly improve data efficiency and model robustness. Such a priori knowledge might arise from physical principles (e.g., conservation laws) or from the system's design (e.g., the Jacobian matrix of a robot), even if large portions of the system dynamics remain unknown. We develop a framework to learn dynamics models from trajectory data while incorporating a priori system knowledge as inductive bias. More specifically, the proposed framework uses physics-based side information to inform the structure of the neural network itself and to place constraints on the values of the outputs and the internal states of the model. It represents the system's vector field as a composition of known and unknown functions, the latter of which are parametrized by neural networks. The physics-informed constraints are enforced via the augmented Lagrangian method during the model's training. We experimentally demonstrate the benefits of the proposed approach on a variety of dynamical systems – including a suite of robotics environments featuring large state spaces, nonlinear dynamics, external forces, contact forces, and control inputs. By exploiting a priori system knowledge, the proposed approach learns to predict the system dynamics two orders of magnitude more accurately than a baseline approach that does not include prior knowledge, while also enforcing physics-based constraints.

Franck Djeumou
University of Texas, U.S.
fdjeumou@utexas.edu

Cyrus Neary
The University of Texas at Austin
cneary@utexas.edu

Ufuk Topcu
University of Texas
utopcu@utexas.edu

MS124

Autodeuq: Automated Deep Ensemble with Uncertainty Quantification

TBD..

Prasanna Balaprakash
Argonne National Laboratory
pbalapra@anl.gov

MS124

Gaussian Processes Meet Neural ODEs: A Bayesian Framework for Learning the Dynamics of Partially Observed Systems from Scarce and Noisy Data

We present a machine learning framework for Bayesian systems identification from partial, noisy, sparse and irregular observations of nonlinear dynamical systems. The proposed method takes advantage of recent developments in differentiable programming to propagate gradient information through ordinary differential equation solvers and perform Bayesian inference with respect to unknown model parameters using Markov Chain Monte Carlo sampling and Gaussian Processes. This allows us to efficiently infer posterior distributions over plausible models with quantified uncertainty, while the use of sparsity-promoting priors such as the Finnish Horseshoe distribution enables the discovery of interpretable and parsimonious representations for the underlying latent dynamics. A series of numerical studies is presented to demonstrate the effectiveness of the proposed methods including predator-prey systems, systems biology, and a 50-dimensional human motion dynamical system. Taken all together, our findings put forth a novel, flexible and robust workflow for data-driven model discovery under uncertainty.

Mohamed Aziz Bhouri
University of Pennsylvania
bhouri@seas.upenn.edu

MS124

Machine Learning, Systematics, and Statistics: A Cosmological Perspective

Over the last few decades, cosmology has transitioned from being data-poor to data-rich; it remains, however, an observational science without the possibility of conducting isolatable, well-designed experiments. At the same time, there is sufficient regularity in the data and connection to fundamental theory that various modeling approaches can be fruitfully applied. Consequently, as a field, cosmology presents many opportunities to apply and develop methods to solve statistical inverse problems. In this talk I will discuss the flavor of several of these problems with a general emphasis on the technical difficulties and approaches

to resolving them.

Salman Habib
Argonne National Laboratory
habib@anl.gov

MS124

A Possible Quantum Advantage for Uncertainty Quantification

A method of uncertainty quantification on a quantum circuit is to place parametrized samples of a linear model on a single block diagonal matrix. This approach leverages the logarithmic scaling of the number of qubits with respect to matrix size. This application area of uncertainty quantification to linear systems can experience a quantum advantage using the method presented in principle. Extensive empirical testing, presented here for the first time, has pinpointed the factor determining accuracy of quantum circuit calculations. Specifically, the accuracy of a quantum circuit simulator has been benchmarked against the classical solution for random matrices which are constricted by various criteria. These criteria include orthogonality, condition number, orthonormality, angle between vectors within the linear system, diagonality, and the angle between the solution vector and the right-hand side vector. The result of these rigorous tests show that the single most powerful factor for accuracy is the angle between the solution vector and the right- hand side vector. In further pursuit of a quantum advantage, a preconditioner for achieving ideal accuracy may be developed for any linear system which will align the solution vector and right- hand side vector. Although expensive to obtain once, the same preconditioner can be applied across all block-diagonal samples which are uncertainty perturbations about the first preconditioned sample.

Eric Walker
University of Buffalo
ericwalk@buffalo.edu

MS125

Hierarchical Methods for High-Dimensional Nonlinear Filtering

Hierarchical extensions of the ensemble Kalman filter (EnKF) has been shown to achieve better efficiency than standard EnKF through variance reduction combined with sampling few ensemble members on the fine-resolution levels. In this talk, I will present two hierarchical filtering methods for high-dimensional filtering problems, namely the multilevel EnKF and the multi-grid EnKF. The focus will be on known convergence and efficiency results, challenges around improving these, and the potential for combining hierarchical methods with localization techniques in large-scale problems.

Hakon Hoel
RWTH Aachen Univeristy
haakonah1@gmail.com

MS125

A Multi-Grid Ensemble Kalman Filter (MGEnKF) Estimator for Sequential Assimilation of Unsteady Flows

We present here an advanced constrained estimator strategy, which combines an EnKF approach with the multigrid method. For this reason, the algorithm is referred to as Multigrid Ensemble Kalman Filter (MGEnKF). Multigrid methods are a family of tools which employ multi-level techniques to obtain the time advancement of the flow. In particular, the geometric multigrid uses different levels of the resolution in the computational grid to obtain the final state. In the MGEnKF, the EnKF error covariance matrix reconstruction is performed using information from several ensemble members which are generated over a coarse level mesh of a multigrid approach. The state estimation obtained at the coarse level and the associated ensemble statistics are used to obtain a single solution calculated on a high-resolution mesh grid. This procedure allows to i) reduce the computational costs of the EnKF and ii) ensure the conservativity and smoothness of the final solution. In addition, owing to the algorithmic structure of the problem, all the simulations on the fine and coarse level can be run simultaneously in parallel calculations, providing a tool able to perform in-streaming DA for unsteady flow problems. The efficiency of the strategy will be illustrated via the analysis of one-dimensional and two-dimensional test cases, using different dynamical equations.

Gabriel Moldovan
Institut Pprime, CNRS -ISAE-ENSMA -Université de Poitiers
gabriel-ionut.moldovan@ensma.fr

Guillaume Lehnasch
Institute PPRIME, CNRS
Université de Poitiers - ENSMA
guillaume.lehnasch@isae-ensma.fr

Laurent Cordier
Institute PPRIME, CNRS
laurent.cordier@univ-poitiers.fr

Marcello Meldi
Institute PPRIME, CNRS
Université de Poitiers - ENSMA
marcello.meldi@ensma.fr

MS125

Accounting for Multilevel Modeling Error in Multilevel Data Assimilation

In ensemble-based Data Assimilation (DA), utilization of lower-fidelity reservoir simulations reduces the computational cost per ensemble member, thereby rendering the possibility of increasing the ensemble size without increasing the total computational cost. Increasing the ensemble size will reduce Monte-Carlo errors and therefore benefit DA results. Use of lower-fidelity reservoir models will, however, introduce modeling errors in addition to those already present in conventional-fidelity simulation results. Multilevel simulations utilize a selection of models for the same entity that constitute hierarchies both in fidelities and computational costs. In this talk, we present techniques for estimating and approximately accounting for the Multilevel Modeling Error (MLME), that is, the part of the total modeling error that is caused by using a multilevel model hierarchy instead of a single conventional model to calculate model forecasts. To this end, four computationally inexpensive approximate MLME correction schemes are presented, and their abilities to correct the multilevel model forecasts for reservoir models with different types of MLME are assessed. Additionally, we assess the performances of the different MLME-corrected model forecasts

in assimilation of acoustic impedance data.

Mohammad Nezhadali
NORCE Norwegian Research Center, University of Bergen (UiB)
NORCE Norwegian Research Center
mone@norceresearch.no

Tuhin Bhakta
NORCE Norwegian Research Centre
tubh@norceresearch.no

Kristian Fossum
NORCE
kristian.fossum@norceresearch.no

Trond Mannseth
NORCE Norwegian Research Centre
trma@norceresearch.no

MS125

Model Forest Extensions to the Multifidelity Ensemble Kalman Filter

This talk will extend the idea of multilevel and multifidelity inference from a model hierarchy to the concept of model forests—collections of models and their respective surrogates. These ideas will be applied to the multifidelity ensemble Kalman filter (MFEnKF) and tested on a quasi-geostrophic model with a standard POD surrogate, and a right-invertible autoencoder-based surrogate.

Andrey A. Popov
Virginia Polytechnic Institute and State University
apopov@vt.edu

Adrian Sandu
Virginia Polytechnic Institute and State University
Computational Science Laboratory
sandu@cs.vt.edu

MS126

Data-Driven Stochastic Modeling of Organized Cloud Fields and Dynamical Insights

The emergence of organised multiscale patterns resulting from convection is ubiquitous, observed throughout different cloud types around the world. The reproduction of such patterns by cloud-resolving models or by large eddy simulation models remains a grand challenge. The core of this modelling challenge lies in the multiscale nature of cloud physics, characterised by the presence of a wide spectrum of variables evolving across a broad range of spatial and temporal scales. With the new advances in data-driven modelling techniques, the discovery of dynamical equations governing time-evolving observations issued from complex systems has gained a lot of attention. This talk will expose how the data-driven model discovery of basic stochastic equations can be solved efficiently for simulating organised, mesoscale continental shallow cumulus, from high-resolution satellite datasets. It will be shown that these equations provide a dynamical generator for the temporal evolution of the multiscale cloud patterns and thereby dynamical insights about the latter. The approach allows not only for reproducing the cloud field's key organisational features, such as cloud streets and convective inertia gravity waves, but also allow for simulating a wide range of out-of-sample days, presenting their own variability while respecting the patterns' physical characteristics.

Mickael Chekroun
Weizmann Institute of Science
michael-david.chekroun@weizmann.ac.il

Tom Dror
Department of Earth and Planetary Sciences
Weizmann Institute
tom.drorschwartz@weizmann.ac.il

Orit Altaratz, Ilan Koren
Department of Earth and Planetary Sciences
Weizmann Institute of Science
orit.altaratz@weizmann.ac.il, ilan.koren@weizmann.ac.il

MS126

Randomized Multilevel Monte Carlo Methods for Bayesian Inference in Science and Engineering Applications

Recently, a growing number of methods have emerged which allow one to leverage cheap low-fidelity models in order to precondition algorithms for performing inference with more expensive models and make Bayesian inference tractable in the context of high-dimensional and expensive models. Some notable examples are multilevel Monte Carlo (MLMC), multi-index Monte Carlo (MIMC), and their randomized counterparts (rML/MIMC), which are able to provably achieve a dimension-independent (including infinite-dimension) canonical complexity rate with respect to mean squared error (MSE) of 1/MSE. Some parallelizability is typically lost in an inference context, but recently this has been largely recovered via novel double randomization approaches. Such an approach delivers i.i.d. samples of quantities of interest which are unbiased with respect to the infinite resolution target distribution. This talk will describe the general approach with a focus on a Markov chain Monte Carlo method. Time permitting, some sequential Monte Carlo methods will be discussed. Over the coming decade, this family of algorithms has the potential to transform data centric science and engineering, as well as classical machine learning applications such as deep learning, by scaling up and scaling out fully Bayesian inference.

Kody Law
University of Manchester
kodylaw@gmail.com

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

MS126

Stochastic Closure Model Via Parametric Inference

Closure models aim to account for the effects of the unresolved scales. We construct closure models in a statistical learning framework, with additional stochastic forces quantifying the uncertainty. A fundamental idea is the approximation of the discrete-time flow map for the dynamics of the resolved variables. The flow map, which encodes the effects of the unresolved scales, is a functional of the resolved scales, thus its inference faces the curse of dimensionality. We investigate a semi-parametric approach that derives parametric models from the structure of the full model. We show that this approach leads to effective reduced models for deterministic and stochastic PDEs, such

as the Kuramoto-Sivashisky equation to viscous stochastic Burgers equations. In particular, we highlight the shift from the classical nonlinear Galerkin method to statistical inference.

Fei Lu
John Hopkins University
flu15@jhu.edu

Kevin K. Lin
Department of Mathematics
University of Arizona
klin@math.arizona.edu

MS126

Data-Driven Variational Multiscale Reduced Order Models for Conditional Gaussian Nonlinear Systems

We propose a new data-driven nonlinear stochastic reduced order models (ROM) for conditional Gaussian nonlinear systems, which uses the variational multiscale (VMS) methodology and data-driven approach. In particular, the new ROM in this framework remains in a relatively simple form and it uses available data to model the closure term that accounts for the nonlinear interactions between resolved scales and unresolved scales. Moreover, this new ROM framework uses VMS to model the interactions between large resolved scales and small resolved scales. The special mathematical structure used in this new ROM allows the development of an efficient algorithm to learn the crucial parameterizations in the ROM when only partial observations are available, which is the typical situation in practice and moreover, it facilitates an efficient and accurate nonlinear data assimilation scheme which provides an extremely useful tool for the uncertainty quantification and forecast. We test the new ROM in the numerical simulation of one dimensional stochastic Burgers equation.

Changhong Mou
Department of Mathematics
Virginia Tech
cmou3@wisc.edu

Nan Chen
University of Wisconsin-Madison
chennan@math.wisc.edu

Traian Iliescu
Department of Mathematics
Virginia Tech
iliescu@vt.edu

MS126

Data-Driven Homogenization of Multiscale Langevin Dynamics

Learning parameters in stochastic models from data is an important problem that arises in many applications. Quite often the stochastic dynamics is characterized by several widely separated length scales and one is then interested in obtaining a reduced, coarse-grained description of the dynamics that is valid at macroscopic scales. Inferring parameters in coarse-grained models using data from the multiscale dynamics is a challenging problem since data are compatible with the surrogate model only at the macroscopic scale. In this talk we consider the framework of the overdamped multiscale Langevin equation, for which a single-scale coarse-grained model exists due to the theory of homogenization, and we study the problem of fitting effective dynamics from observations of the multiscale system. In particular, we propose two novel techniques based on filtered data to estimate the drift coefficient of the homogenized equation. One is a modification of the maximum likelihood estimator and the other relies on the eigenvalues and eigenfunctions of the generator of the homogenized dynamics. We prove theoretically that our estimators are asymptotically unbiased in the limit of infinite data and when the multiscale parameter vanishes. A series of numerical experiments illustrates the advantages of our approach and demonstrates that our methodology can be employed to robustly infer effective models from complex phenomena.

Andrea Zanoni, Assyr Abdulle, Giacomo Garegnani
EPFL
andrea.zanoni@epfl.ch, assyr.abdulle@epfl.ch,
giacomo.garegnani@epfl.ch

Grigorios Pavliotis
Imperial College London
Department of Mathematics
g.pavliotis@imperial.ac.uk

Andrew Stuart
Computing + Mathematical Sciences
California Institute of Technology
astuart@caltech.edu

MS127

Model-Constrained Deep Learning Methods for Forward and Inverse Problems

Deep Learning (DL), in particular deep neural networks (DNN), by design is purely data-driven and in general does not require physics. This is the strength of DL but also one of its key limitations when applied to science and engineering problems in which underlying physical properties - such as stability, conservation, and positivity - and desired accuracy need to be achieved. DL methods in their original forms are not capable of respecting the underlying mathematical models or achieving desired accuracy even in big-data regimes. On the other hand, many data-driven science and engineering problems, such as inverse problems, typically have limited experimental or observational data, and DL would overfit the data in this case. Leveraging information encoded in the underlying mathematical models, we argue, not only compensates missing information in low data regimes but also provides opportunities to equip DL methods with the underlying physics and hence obtaining higher accuracy. This short communication introduces several model-constrained DL approaches - including both feed-forward DNN and autoencoders - that are capable of learning not only information hidden in the training data but also in the underlying mathematical models to solve inverse problems. We present and provide intuitions for our formulations for general nonlinear problems.

Hai Nguyen
The University of Texas at Austin
hainguyen@utexas.edu

Tan Bui-Thanh
University of Texas at Austin

tanbui@oden.utexas.edu

MS127

Deep Neural Network Expression of Posteriors in Bayesian PDE Inversion

Deep Neural Network Expression of Posteriors in Bayesian PDE Inversion. We present results on bounds for DNN expression rates and on constructive DNN expression for data-to-QoI maps and posterior densities in Bayesian PDE inversion. Forward PDE models are assumed countably-parametric. Data is assumed subject to additive, centered gaussian noise, in finite dimension. DNN expression rates are free from the Curse of Dimensionality.

Christoph Schwab
ETH-Zurich
schwab@sam.math.ethz.ch

Lukas Herrmann
RICAM, Linz, Austria
lukas.herrmann@sam.math.ethz.ch

Joost Opschoor
SAM. ETH Zurich
joost.opschoor@sam.math.ethz.ch

MS127

Regularized Random Feature Methods for Reduced Order Modeling

We present a data-driven approximation method for reduced order modeling. The surrogate model uses over-parameterized random feature maps to approximate the behavior of the data projected onto low-dimensional subspaces. We show that the approximations remain accurate even in the data scarce regime. Examples include shape modeling and modeling spatio-temporal dynamics.

Hayden Schaeffer
Carnegie Mellon University
schaeffer@cmu.edu

MS127

Normalizing Field Flow: Solving Forward and Inverse Stochastic Differential Equations Using Physics-Informed Flow Model

We introduce normalizing field flows (NFF) for learning random fields from scattered measurements. More precisely, we construct a bijective transformation (a normalizing flow characterizing by neural networks) between a reference random field (say, a Gaussian random field with the Karhunen-Loève (KL) expansion structure) and the target stochastic field, where the KL expansion coefficients and the invertible networks are trained by maximizing the sum of the log-likelihood on scattered measurements. This NFF model can be used to solve data-driven forward, inverse, and mixed forward/inverse stochastic partial differential equations in a unified framework. We demonstrate the capability of the proposed NFF model for learning Non-Gaussian processes, mixed Gaussian processes, and forward & inverse stochastic partial differential equations.

Ling Guo
Shanghai Normal University
lguo@shnu.edu.cn

Hao Wu
Tongji University
School of Mathematical Sciences
hwu@tongji.edu.cn

Tao Zhou
Institute of Computational Mathematics
Chinese Academy of Sciences
tzhou@lsec.cc.ac.cn

MS128

A Deep-Learning-Based Inverse Design Framework Using Compressed Simulation Data for Self-Oscillating Gels

We consider data-driven design of soft robots using a novel locomotion mechanism based on a reaction-driven motion of an oscillating sheet. While a forward model for the simulation of this sheet is available as a PDE solver by Alben et. al (2019), design applications require the inversion of sheet motion into geometric and reaction parameters of the sheet. While PDE constrained approaches for this inversion are possible, they may be time consuming and computationally expensive. Instead, we develop a deep-learning based model mapping sheet motion to sheet parameters. This framework is based on a stacked MLP design, whose hyperparameters we carefully tune to reconstruct the geometric and reaction parameters with high accuracy. However, this approach to train an inverse design framework for self-oscillating gels requires a large amount of data, which both makes training demanding and slows down the training procedure. To tackle these types of challenges in more general inverse-design settings, we investigate compact representations of the simulation data through low-rank matrix and tensor decomposition methodologies on different network structures. We compare the compressed-data inverse design framework with our original framework in terms of estimation accuracy of the continuous and discrete gel-sheet parameters, training speed, and memory requirements.

Doruk Aksoy
Department of Aerospace Engineering
University of MIchigan
doruk@umich.edu

Alex Gorodetsky
University of Michigan
goroda@umich.edu

Silas Alben
University of Michigan, USA
alben@umich.edu

Robert Deegan
University of Michigan
Department of Physics
rddeegan@umich.edu

Shravan Veerapaneni
Department of Mathematics
University of Michigan
shravan@umich.edu

MS128

Reduced Basis Approximation for Variational Data Assimilation over Hyper-Parameterized Model Un-

certainties

Mathematical models, such as partial differential equations (PDEs), are widely used to predict the behavior of a physical system. However, any model can only provide an approximation to the underlying physics, and can be subject to a variety of model errors, such as uncertainties in the loading or initial condition. Variational Data Assimilation can be used to improve models through the incorporation of measurement data. Here, the inverse solution requires many evaluations of the full-order forward model, already leading to large computational costs. When the inverse solution needs to be computed for a variety of hyper-parameters describing flexible system configurations, it becomes prohibitive. In this talk we present reduced basis (RB) approximations to the 3D-VAR and 4D-VAR variational data assimilation methods posed over hyper-parameterized PDEs with linear model uncertainties. After a preparatory offline phase, the RB-3D-VAR and RB-4D-VAR methods can be evaluated at significantly reduced cost, thereby enabling many-query approximations of the inverse solutions for different hyper-parameters and data, while the approximation error can be monitored through rigorous and certified a posteriori error bounds. We discuss implementational challenges and solutions for parabolic problems for an exemplary continuous-Galerkin space-time finite element discretization. We demonstrate our results on a contaminant-dispersion problem over a Taylor-Green vortex velocity field.

Nicole Aretz
RWTH Aachen University
aretz@aices.rwth-aachen.de

Francesco Silva
TU/e Eindhoven University of Technology
f.a.b.silva@tue.nl

Martin Grepl
RWTH Aachen University
grepl@igpm.rwth-aachen.de

Karen Veroy-Grepl
TU/e Eindhoven University of Technology
k.p.veroy@tue.nl

MS128

Adaptive Quantile Surrogate Based Methodology for Design Optimization under Uncertainties

In the processes for designing structural systems, it is essential to incorporate the effects of various uncertainties arising from the lack of data, modeling approximation, or inherent randomness in the systems and their environment. To this end, reliability-based design optimization (RBDO) has been extensively studied and applied to make the best decision based on the structural capabilities to predict the risk caused by uncertainties. However, their practical applications are often hampered by the huge computational costs required for the repeated reliability analysis. To address the challenge, this paper presents the two quantile-surrogate-based RBDO methods recently developed by the authors [Kim, J., & Song, J. Quantile surrogates and sensitivity by adaptive Gaussian process for efficient reliability-based design optimization. Mechanical Systems and Signal Processing, 2021, 161: 107962]. The methods estimate the quantile surrogates of the performance functions to identify the admissible design domain. Quantile surrogates are adaptively trained through an exploration-exploitation trade-off by an active learning framework. In addition, the approaches derive the parameter sensitivity formulations of quantile surrogates to facilitate applications to high-dimensional RBDO problems. The numerical examples demonstrate that the proposed methods facilitate convergence to reliable optimum design with fewer computational simulation data compared to existing RBDO approaches.

Jungho Kim
Seoul National University
cong020@snu.ac.kr

Junho Song
Seul National University
junhosong@snu.ac.kr

MS128

Stochastic Learning Approach to Binary Optimization for Optimal Design of Experiments

We present a novel stochastic approach to binary optimization for optimal experimental design (OED) for Bayesian inverse problems governed by mathematical models such as partial differential equations. The OED utility function, namely, the regularized optimality criterion, is cast into a stochastic objective function in the form of an expectation over a multivariate Bernoulli distribution that models the probability of a measurement at a given location. The probabilistic objective is then optimized using a stochastic-gradient optimization approach to find an optimal observational policy. The proposed approach is analyzed from an optimization perspective and also from a machine-learning perspective corresponding to policy gradient reinforcement learning. The approach is evaluated numerically using a sensor placement problem for parameter identification for an idealized two-dimensional Bayesian linear inverse problem.

Sven Leyffer
Argonne National Laboratory
leyffer@mcs.anl.gov

Ahmed Attia
Argonne National Laboratory
attia@anl.gov

Todd Munson
Argonne National Laboratory
Mathematics and Computer Science Division
tmunson@mcs.anl.gov

MS129

Reduced Model Bayesian Multiscale Method and Uncertainty Quantification

In many applications, in inverse problems, the underlying data generating mechanisms are determined by a series of differential equations which make the likelihood computation or forward solution computationally burdensome. Bayesian methods give a natural framework for inverse problem solution by specifying prior and computing the high probability solution given the observed data, and quantifying the uncertainty. Reduced model Bayesian method plays an important role by approximating the likelihood. Similarly, posterior approximation techniques can be helpful for fast scalable computation of the posterior distribution. This talk focuses on such reduced model ap-

proximate solutions and their properties.

Nilabja Guha
UMASS Lowell
nguha1@umbc.edu

MS129

Bayesian Learning of Neural Networks for Small Or Imbalanced Data Sets

Data-based predictive models such as neural networks are showing great potential to be used in various scientific and engineering fields. They can be used in conjunction with physics-based models to account for missing or hard-to-model physics, or as surrogates to replace high-fidelity, overly costly physics-based simulations. However, in many engineering fields data scarcity and / or data imbalance is a challenge. Bayesian training of neural networks allows for a comprehensive account of both aleatory and epistemic uncertainties that arise from the inherent stochasticity of the physical processes being modeled and the data scarcity. This talk will discuss two engineering applications of Bayesian neural networks: surrogate materials modeling, and ambulance travel time prediction. In the first case, data is scarce as it is obtained from expensive high-fidelity materials simulations. Bayesian methods based on variational inference and model averaging are utilized for training. In the second case, real data from ambulance trips in NYC, obtained as part of a collaboration with the New York City Fire Department, is used to train a neural network for travel time prediction. This data is highly imbalanced, affected by current policies on ambulances deployment. Using Bayesian methods allows us to integrate knowledge from a road network analysis to build a meaningful prior, and thoroughly account for uncertainties.

Audrey Olivier
University of Southern California
audreyol@usc.edu

Michael D. Shields, Lori Graham-Brady
Johns Hopkins University
michael.shields@jhu.edu, lori@jhu.edu

Andrew Smyth
Department of Civil Engr. & Mech.
Columbia University
smyth@civil.columbia.edu

Sevin Mohammadi
Columbia University
sm4894@columbia.edu

Matthew Adams, Kat Thomson
New York City Fire Department
matthew.adams@fdny.nyc.gov, kat.thomson@fdny.nyc.gov

Henry Lam, Enrique Lelo de Larrea, Elioth Sanabria
Columbia University
henry.lam@columbia.edu, enrique.lelodelarrea@columbia.edu, m.elioth@columbia.edu

MS129

Bayesian Emulation of Complex Computer Models with Structured Partial Discontinuities

Complex mathematical computer models are used across many scientific disciplines and industry to improve the understanding of the behaviour of physical systems and provide decision support. Long evaluation times combined with large numbers of inputs and outputs renders full uncertainty quantification calculations prohibitively expensive. Instead, Bayesian emulators are often employed to provide fast predictions for as yet unevaluated parameter settings along with a statement of the uncertainty. The speed of emulator evaluation then facilitates a full uncertainty quantification. Emulator construction typically assumes that the computer model is continuous and potentially smooth; hence they struggle to handle various types of structured discontinuities. We address the emulation of functions that possess a finite set of discontinuities of possibly complex form, where the endpoint locations of each discontinuity may lie within the input space. This involves embedding the input space on a hypersurface within a higher-dimensional space; torn along the locations of the existing structured discontinuities. Emulation proceeds over the higher-dimensional space with full flexibility of emulator forms. The embedding induces a warped emulator when projected onto the original inputs which we counteract via a carefully constructed non-stationary covariance function. We demonstrate this for a petroleum well placement problem with discontinuities induced by partial geological fault boundaries

Jonathan Owen
Durham University, UK
jonathan.owen@durham.ac.uk

MS129

Efficient Emulation of Epidemiological Spatio-Temporal Patch Models

Modern computational science allows complex scientific processes to be described by mathematical models implemented in computer codes, or simulators. When these simulators are computationally expensive, it is common to approximate them using statistical emulators constructed from computer experiments. Often, the simulator output represents system behaviour across a large spatial and/or temporal domain, which can make efficient emulation computationally challenging. Epidemic progression is an important application area utilising temporal (and often spatial) simulation. Our motivating application simulator is a multi-patch compartmental model of flu, with each patch representing a region of Botswana. We have extended the methodologies of Outer Product Emulators (OPEs) and Parallel Partial Emulators (PPEs) to emulate the epidemiological model for the spread of flu through a spatially dispersed population. Specifically, the model produces time series of the number of infected people in different patches of the space, making the output spatio-temporal. We compare and analyse the benefits and drawbacks of each emulation approach in application to the flu model.

Daria Semochkina
University of Southampton
d.semochkina@soton.ac.uk

MS130

Variance Reduction for Estimation of Shapley Effects and Adaptation to Unknown Input Distribution

The Shapley effects are global sensitivity indices: they quantify the impact of each input variable on the output variable in a model. In this work, we suggest new estimators of these sensitivity indices. When the input dis-

tribution is known, we investigate the already existing estimator defined in [E. Song, B. L. Nelson, and J. Staum, SIAM/ASA J. Uncertain. Quantif., 4 (2016), pp. 1060–1083] and suggest a new one with a lower variance. Then, when the distribution of the inputs is unknown, we extend these estimators. We provide asymptotic properties of the estimators studied in this article. We also apply one of these estimators to a real data set.

Francois Bachoc
Institut de Mathematiques de Toulouse
francois.bachoc@math.univ-toulouse.fr

MS130

Global Sensitivity Analysis of High Dimensional Neuroscience Models

The complexity and size of state-of-the-art cell models have significantly increased in part due to the requirement that these models possess complex cellular functions which are thought–but not necessarily proven–to be important. Modern cell models often involve hundreds of parameters; the values of these parameters come, more often than not, from animal experiments whose relationship to the human physiology is weak with very little information on the errors in these measurements. The concomitant uncertainties in parameter values result in uncertainties in the model outputs or Quantities of Interest (QoIs). New Global Sensitivity Analysis (GSA) approaches are required to deal with increased model size and complexity; a three stage methodology consisting of screening (dimension reduction), surrogate modeling, and computing Sobol' indices, is presented. The methodology is used to analyze a physiologically validated numerical model of neurovascular coupling which possess hundreds of uncertain parameters.

Pierre Gremaud
Department of Mathematics
North Carolina State University
gremaud@ncsu.edu

MS130

On Quantile Oriented Sensitivity Analysis

We propose to study quantile oriented sensitivity indices (QOSA indices) and show some of their theoretical properties. These have a number of shortcomings, when dealing both with independent and dependent inputs, which lead us to define new generic indices based on the Shapley values named Goal-Oriented Shapley Effects (GOSE). In particular, we focus on Quantile-Oriented Shapley effects (QOSE) and subsequently perform several calculations of QOSA indices and QOSE in order to better understand the behaviour and the respective interest of each.

Veronique Maume-Deschamps
University of Lyon / AMIES
veronique.maume@univ-lyon1.fr

MS130

Efficient Estimation of the Anova Mean Dimension, with An Application to Neural Net Classification

The mean dimension of a black box function of d variables is a convenient way to summarize the extent to which it is dominated by high or low order interactions. It is expressed in terms of 2**d-1 variance components but it can be written as the sum of d Sobol' indices that can be estimated by leave one out methods. We compare the variance of these leave one out methods: a Gibbs sampler called winding stairs, a radial sampler that changes each variable one at a time from a baseline, and a naive sampler that never reuses function evaluations and so costs about double the other methods. For an additive function the radial and winding stairs methods are most efficient. For a multiplicative function the naive method can easily be most efficient if the factors have high kurtosis. As an illustration we consider the mean dimension of a neural network classifier of digits from the MNIST data set. The classifier is a function of 784 pixels. For that problem, winding stairs is the best algorithm. We find that inputs to the final softmax layer have mean dimensions ranging from 1.35 to 2.0.

Art B. Owen, Christopher Hoyt
Stanford University
owen@stanford.edu, crhoyt@stanford.edu

MS131

Near-optimal compression in near-linear time

We introduce Kernel Thinning-Compress++ (KT-Compress++), an algorithm based on two new procedures for compressing a distribution P nearly optimally and much more effectively than i.i.d. sampling/standard thinning. Given a suitable reproducing kernel k and $O(nlog^3n)$ time, KT-Compress++ compresses an n-point approximation to P into an sqrt(n)-point approximation with better than Monte Carlo integration error rates for functions in the associated reproducing kernel Hilbert space (RKHS). First, we show that for any fixed-function KT-Compress++ provides dimension-free guarantees for any kernel, any distribution, and any fixed-function in the RKHS. Second, we show that with high probability, the maximum discrepancy in integration error is $O_d(n^{-1/2})$ up to logarithmic factors for a broad class of P on R^d. In contrast, an equal-sized i.i.d. sample from P suffers at least $n^{-1/4}$ integration error. Our guarantees nearly match the known lower bounds for several settings, and while resembling the quasi-Monte Carlo error rates for uniform P on $[0,1]^d$ they apply to general distributions on R^d and several universal kernels. En route, we introduce a new simple meta-procedure Compress++, that can significantly speed up the runtime of a generic thinning algorithm while suffering at most a factor of 4 error. Finally, we present several vignettes illustrating the practical benefits of KT-Compress++ in dimensions $d = 2$ to 100.

Raaz Dwivedi
Harvard University
raaz@mit.edu

Lester Mackey
Microsoft Research New England
lmackey@microsoft.com

Abhishek Shetty
UC Berkeley

MS131

Bayesian Multilevel Monte Carlo

Multilevel Monte Carlo is a common tool for approximating integrals of computationally expensive models. Its efficiency stems from the use of approximations of the integrand of interest at several levels of accuracy which can

be evaluated at a reduced computational cost, and naturally lead to an estimator with lower variance than classical Monte Carlo methods. In this paper, we propose a Bayesian multilevel Monte Carlo method. We show that our approach leads to significant further improvements in accuracy.

Kaiyu Li, Daniel Giles
University College London
kaiyu.li.19@ucl.ac.uk, d.giles@ucl.ac.uk

Toni Karvonen
Aalto University
Department of Electrical Engineering and Automation
toni.karvonen@helsinki.fi

Serge Guillas, Francois-Xavier Briol
University College London
s.guillas@ucl.ac.uk, f.briol@ucl.ac.uk

MS131

Monte Carlo Variance Reduction Using Stein Operators

This talk will focus on two new methods for estimating posterior expectations when the derivatives of the log posterior are available. The proposed methods are in a class of estimators that use Stein operators to generate control variates or control functionals. The first method applies regularisation to improve the performance of popular Stein-based control variates for high-dimensional Monte Carlo integration. The second method, referred to as semi-exact control functionals (SECF), is based on control functionals and Sard's approach to numerical integration. The use of Sard's approach ensures that our control functionals are exact on all polynomials up to a fixed degree in the Bernstein-von-Mises limit. Several Bayesian inference examples will be used to illustrate the potential for reduction in mean square error. If time permits, I will also briefly describe some benefits and challenges of Stein-based control variates in the unbiased Markov chain Monte Carlo setting.

Leah F. South
Queensland University of Technology
l1.south@qut.edu.au

Toni Karvonen
Aalto University
Department of Electrical Engineering and Automation
toni.karvonen@helsinki.fi

Chris Drovandi
Queensland University of Technology
Australia
c.drovandi@qut.edu.au

Antonietta Mira
Universita della Svizzera italiana, Switzerland
University of Insubria, Italy
antonietta.mira@usi.ch

Christopher Nemeth
Lancaster University
c.nemeth@lancaster.ac.uk

Mark Girolami
University of Cambridge
mag92@cam.ac.uk

Chris Oates
Newcastle University
Chris.Oates@newcastle.ac.uk

MS131

Vector-Valued Control Variates

Control variates are post-processing tools for Monte Carlo estimators which can lead to significant variance reduction. This approach usually requires a large number of samples, which can be prohibitive for applications where sampling from a posterior or evaluating the integrand is computationally expensive. Furthermore, there are many scenarios where we need to compute multiple related integrals simultaneously or sequentially, which can further exacerbate computational costs. In this paper, we propose *vector-valued control variates*, an extension of control variates which can be used to reduce the variance of multiple integrals *jointly*. This allows the transfer of information across integration tasks, and hence reduces the overall requirement for a large number of samples. We focus on control variates based on kernel interpolants and our novel construction is obtained through a generalised Stein identity and the development of novel matrix-valued Stein reproducing kernels. We demonstrate our methodology on a range of problems including multifidelity modelling and model evidence computation through thermodynamic integration.

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

Alessandro Barp
University of Cambridge
ab2286@cam.ac.uk

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

MS132

Generalized Feature Fitting Approach for Remote Imaging Spectroscopy

Imaging spectroscopy is a remote-sensing technique that allows the mapping of surface features by recording reflective light in narrow and continuously spaced spectral bands. Many materials on Earths surface have unique spectral signitures, which are detectable via these measurements. A generalized framework that can account for various sources of uncertainty is needed to calculate and accurately map the abundance of materials with known signatures. In this work, we developed a new data model that accounts for observation and modeling uncertainties so these quantities can be leveraged for improved modeling performance and propagated for the production of mineral maps with error bars. We use generalized least squares and chi2 hypothesis testing to provide the capability to decline a proposed model and to compare between competing models on the same pixel. We tested this algorithm on real-world data from Cuprite Hills, NV., a relict hydrothermal alteration mineral site, measured using NASAs Airborne Visible InfraRed Imaging Spectrometer - Next Generation (AVIRIS-NG). Our results show improved retrieval performance compared with the traditional approach, as the algorithm can leverage uncertainties that vary spatially between pixels. This methodology was developed for the study case of mapping surface minerals but is readily appli-

cable to other applications that map materials with known spectral signatures, e.g., vegetation mapping, soil property mapping, and target detection.

Nimrod Carmon
Jet Propulsion Laboratory
nimrod.carmon@jpl.nasa.gov

David R. Thompson
Jet Propulsion Laboratory
California Institute of Technology
david.r.thompson@jpl.nasa.gov

Michael Turmon
JPL
michael.j.turmon@jpl.nasa.gov

Hai Nguyen
Jet Propulsion Laboratory
California Institute of Technology
hai.nguyen@jpl.nasa.gov

MS132

Bayesian Spatial Statistical Modeling and Computation for Global Imaging Spectroscopy

The Bayesian approach to inverse problems arising in imaging spectroscopy can quantify uncertainty in retrievals and help elucidate the value of different information sources, but it can be computationally intractable in practice. In many Bayesian inverse problems, however, there exists a low-dimensional likelihood-informed subspace that describes both optimal projections of the data and directions in parameter space that are most informed by the data. We utilize this subspace in inverse problems for fitting surface and atmospheric models to imaging spectrometer data, with the goal of developing a Markov chain Monte Carlo (MCMC) retrieval algorithm sufficiently fast for operations. As the current approach focuses on pixel-by-pixel retrievals, we explore extensions of this approach that exploits correlations in a spatial field.

Kelvin Leung, Youssef M. Marzouk
Massachusetts Institute of Technology
kmleung@mit.edu, ymarz@mit.edu

MS132

Radiative Transfer Emulation for Hyperspectral Imaging Retrievals with Advanced Kernel Flows-Based Gaussian Process Emulation

Changes in geophysical properties on the surface of the Earth can be monitored with remote imaging spectroscopy. These properties are derived from surface reflectance, which needs to be obtained in a first-stage retrieval from measured radiance spectra. Performing this retrieval requires solving a non-linear inverse problem and the quality of the second-stage retrievals of the biophysical properties depends on how well the first-stage retrieval is performed. A computational bottleneck for these retrievals is radiative transfer: for each retrieval a radiative transfer model (RTM) needs to be queried several times, and these models are in general computationally costly. A way to resolve this issue is to construct a lookup table representation of the RTM, which is then interpolated. While kriging offers a simple solution, its accuracy and the underlying uncertainties are highly sensitive to the prior selection of an underlying kernel. We solve this problem by using the Kernel Flows (KF) algorithm to learn a data-dependent kernel, with which we construct a Gaussian process emulator for the RTM. This emulator can then be used in retrievals. In this presentation we discuss the KF algorithm implementation and the emulator design for the reflectance retrieval problem. We describe the computational details, evaluate the performance of the emulator, and describe what effects keeping track of uncertainties in approximating radiative transfer has on retrieved surface reflectances.

Jouni Susiluoto
Jet Propulsion Laboratory
jouni.i.susiluoto@jpl.nasa.gov

Amy Braverman
Jet Propulsion Laboratory
California Institute of Technology
Amy.Braverman@jpl.nasa.gov

Philip Brodrick
Jet Propulsion Laboratory
philip.g.brodrick@jpl.nasa.gov

Boumediene Hamzi
AlFaisal University
Department of Mathematics
boumediene.hamzi@gmail.com

Margaret Johnson, Otto Lamminpää
Jet Propulsion Laboratory
maggie.johnson@jpl.nasa.gov,
otto.m.lamminpaa@jpl.nasa.gov

Houman Owhadi
Applied Mathematics
Caltech
owhadi@caltech.edu

Clint Scovel
Caltech
clintscovel@gmail.com

Joaquim Teixeira
Jet Propulsion Laboratory
joaquim.p.teixeira@jpl.nasa.gov

Michael Turmon
JPL
michael.j.turmon@jpl.nasa.gov

MS133

Causal Inference for Extreme Values

Extreme value statistics concerns the maxima of random variables and relations between the tails of distributions rather than averages and correlations. The vast majority of models are centered around max-stable distributions, the Gaussian analogs for extremes. However, max-stable multivariate have an intractable likelihood, severely limiting statistical learning and inference. Directed graphical models for extreme values (aka max-linear Bayesian networks) have only appeared in 2018, but have seen many applications in finance, hydrology and extreme risks modelling. This talk (1) highlights how they differ from usual Bayesian networks, (2) discusses their connections to tropical convex geometry, (3) shows performances of current learning algorithms on various hydrology datasets, and (4)

outlines major theoretical and practical challenges.

Ngoc Tran
University of Texas at Austin
Department of Mathematics
ntran@math.utexas.edu

Johannes Buck
Munich University of Technology
graphmod71739

Claudia Kluppelberg
Munich University of Technology
Germany
cklu@ma.tum.de

MS133

Spatial Modeling of Heavy Precipitation by Coupling Weather Station Recordings and Ensemble Forecasts with Max-Stable Processes

Due to complex physical phenomena, the distribution of heavy rainfall events is difficult to model spatially. Physics-based numerical models can provide physically coherent spatial patterns, but may miss important precipitation features like heavy rainfall intensities. Measurements at ground-based weather stations supply adequate rainfall intensities, but most national weather recording networks are spatially too sparse to capture rainfall patterns. To fully leverage these two information sources, climatologists and hydrologists seek models that can merge different types of rainfall data. One inherent difficulty is capturing the appropriate multivariate dependence structure among rainfall maxima. Multivariate extreme value theory (EVT) suggests the use of a max-stable process, which can be represented by a max-linear combination of independent copies of a hidden stochastic process weighted by a Poisson point process. In practice, the choice of this hidden process is non-trivial, especially if anisotropy, non-stationarity and nugget effects are present in the spatial data. By coupling forecast ensemble data from the French national weather service with local observations, we construct and compare different types of data-driven max-stable processes that are parsimonious in parameters, easy to simulate and capable of reproducing nugget effects and spatial non-stationarities. We compare our new method with classical approaches from spatial EVT such as Brown-Resnick processes.

Marco Oesting
University of Stuttgart
marco.oesting@mathematik.uni-stuttgart.de

Philippe Naveau
Laboratoire des Sciences du Climat et de lEnvironnement
philippe.naveau@lsce.ipsl.fr

MS133

Neural Extreme Value Copulas

We propose a neural-network-based method for calibrating and sampling multivariate extreme value distributions (EVDs). EVDs arise from Extreme Value Theory as the necessary class of models when extrapolating a distributional fit over large spatial and temporal scales based on data observed in intermediate scales. EVDs are governed by the so-called Pickands dependence function, which is known to possess certain structural properties. We demonstrate that these structural properties, which we carefully maintain in our neural-network architecture, enable learning EVDs using relatively small amounts of data. Moreover, we present new methods for sampling high-dimensional EVDs using a generative model. We demonstrate our methodology with a number of experiments, in data ranging from environmental sciences to financial mathematics, as well as on synthetic data. The empirical results demonstrate the considerable promise of our proposed methods.

Vahid Tarokh, Ali Hasan, Khalil Elkhalil
Duke University
vahid.tarokh@duke.edu, ali.hasan@duke.edu,
khalil.elkhalil@duke.edu

Joao M. Pereira
University of Texas, Austin
jmoraiscp@gmail.com

Sina Farsiu
Duke University
sina.farsiu@duke.edu

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

MS133

Distributionally Robust Non-Parametric Bayesian Conditional Mean Estimation

We study a distributionally robust optimization formulation (i.e. a minmax game) for the problem of non-parametric Bayesian conditional mean estimation. We choose the best mean-squared error predictor on a Hilbert space against an adversary which chooses the worst-case model in a Wasserstein ball around an infinite-dimensional Gaussian model. The Wasserstein cost function is chosen to control features such as the amount of roughness or smoothness on the paths that the adversary is allowed to inject. We show that the game has a well defined value (i.e. strong duality holds in the sense that max-min = min-max) and existence of a Nash equilibrium which can be computed by a sequence of finite-dimensional approximations. We showcase the versatility of our modeling framework through a set of numerical experiments.

Xuhui Zhang, Jose H. Blanchet
Stanford University
xuhui.zhang@stanford.edu, jose.blanchet@stanford.edu

MS134

Apik: Active Physics-Informed Kriging Model with Partial Differential Equations

Kriging (or Gaussian process regression) becomes a popular machine learning method for its flexibility and closed-form prediction expressions. However, one of the key challenges in applying kriging to engineering systems is that the available measurement data is scarce due to the measurement limitations or high sensing costs. On the other hand, physical knowledge of the engineering system is often available and represented in the form of partial differential equations (PDEs). We present in this paper a PDE Informed Kriging model (PIK), which introduces PDE information via a set of PDE points and conducts posterior prediction similar to the standard kriging method. The proposed PIK model can incorporate physical knowledge from both linear and nonlinear PDEs. To further improve learning performance, we propose an Active PIK framework (APIK)

that designs PDE points to leverage the PDE information based on the PIK model and measurement data. The selected PDE points not only explore the whole input space but also exploit the locations where the PDE information is critical in reducing predictive uncertainty. Finally, an expectation-maximization algorithm is developed for parameter estimation. We demonstrate the effectiveness of APIK in two synthetic examples, a shock wave case study, and a laser heating case study.

Jialei Chen
University of Georgia
jialei.chen@uga.edu

Zhehui Chen, Chuck Zhang
Georgia Tech
zhchen@gatech.edu, chuck.zhang@gatech.edu

Jeff Wu
Georgia Institute of Technology
jeffwu@isye.gatech.edu

MS134

Bayesian Learning of Stochastic Dynamical Models for PDE Quantities of Interest: Avoiding Full-Field Reduced Modeling

We present a formulation for the identification of dynamical systems that is derived from probabilistic principles and accounts for three types of uncertainty present in system ID: parameter, model, and measurement. Most existing system ID methods will only consider uncertainty in the data/output of the system. In contrast, we additionally account for the uncertainty present in our dynamics/parameters by modeling the system as a stochastic process. This modeling choice can be shown to produce a regularizing term in the posterior that can counteract multimodality and make optimization/sampling easier. Furthermore, this regularization is derived directly from the model of the system rather than from heuristic techniques such as penalizing the parameter norm. Recent algorithmic developments include parallel tempering MCMC sampling techniques that can help overcome the difficulties of sampling complicated posteriors. We present the benefits of this method for the identification of nonlinear dynamics, including chaos and PDEs. When considering high-dimensional systems, we demonstrate our methods applicability to identifying the dynamics of a quantity of interest, which is oftentimes desirable to avoid the computational expense of full-field simulation.

Nicholas Galioto, Alex Gorodetsky
University of Michigan
ngalioto@umich.edu, goroda@umich.edu

MS134

Characterizing Uncertainty in Deep Learning Models: An Application To Image Data from Combustion Experiments

In this talk, we discuss the uncertainty quantification process for a U-net used to predict the fuel regression rate for a hybrid rocket experiment. The U-net is a convolutional neural network (CNN) architecture used for image segmentation: the task of classifying pixels and identifying objects and their location in an image. Propulsion research is an area where the U-net may be gainfully employed due to the ample availability of experimental image data and the need for deriving combustion-related quantities of interest from such data. Uncertainty quantification of such predictive CNN models is essential: experimental images are noisy and these models include inherent variation in their predictions. We focus here on the fuel regression rate, i.e. the rate that the fuel recedes over the course of a burn. The uncertainty quantification process includes two parts: model uncertainty (How confident is the model for each pixel prediction in the segmented fuel image?) and input data uncertainty (What is inherent uncertainty in the training data and how does it impact the predictions?). To quantify model uncertainty, we introduce dropout layers to the U-net, which allows to produce an uncertainty map for each prediction. For the input data, we analyzed how uncertainty is introduced through each process in preparation of the training images: camera calibration, experimental noise, image resizing, and human error.

Georgios Georgalis, Abani Patra
Tufts University
georgios.georgalis@tufts.edu, abani.patra@tufts.edu

MS134

Efficient Large-Scale Bayesian Inference for Predictive Modeling in Precision Health Balance Training

Deep neural networks (DNNs) are a powerful tool for automating expert assessment tasks in the fields of medicine and physical therapy, and can substantially reduce the workload of trained technicians. In this work, we use convolutional neural networks in a novel application of assessing balance potential for older adults and adults with inner ear disorders. However, DNNs often fail to sufficiently (or at all) convey appropriate uncertainty information, which is critical to both medical professionals and patients alike. We quantify uncertainty in DNN models by performing Bayesian inference for their weight parameters, i.e., creating Bayesian neural networks. With parameter dimension often reaching hundreds of thousands for DNNs, we conduct approximate inference using Stein variational gradient descent (SVGD). Parameter dimensionality is further reduced by projected parameters onto Hessian-informed active subspaces, via projected SVGD. Prior distributions for the DNN weights also tend to be very abstract, and at times arbitrary. We illustrate that priors on DNN parameters can in fact be used to create intuitive prior predictions while also inducing prediction-enhancing regularization. Furthermore, a hierarchical Bayesian framework is also demonstrated in which the model can learn the quality of data and carefully weigh the impact of the data on modifying prior.

Jeremiah Hauth, Jamie Ferris, Safa Jabri, Steven Teguhlaksana, Kathleen Sienko, Xun Huan
University of Michigan
hauthj@umich.edu, jcferris@umich.edu, safajb@umich.edu, steventl@umich.edu, sienko@umich.edu, xhuan@umich.edu

MS135

Transient Effects of Fire on Landscapes and Implications for Modeling Water-Related Hazards

Fire temporarily alters soil and vegetation properties, promoting increases in runoff and erosion that can dramatically increase the likelihood of destructive flash floods and debris flows downstream of burned areas. Debris flows, or fast-moving landslides that consist of a mixture of water, mud, and rock, initiate after fires when surface water runoff rapidly erodes sediment on steep slopes. Due to the

complex interactions between runoff generation, soil erosion, and post-fire debris-flow initiation/growth, the study of post-fire debris-flow hazards necessitates an approach that couples these processes within a common modeling framework. Models used to simulate these processes, however, often contain a number of poorly constrained parameters, particularly in post-fire settings where there is limited time to collect data and where parameters related to soil and vegetation properties will change over time as the landscape recovers. Here, we describe the immediate impact of fire on the landscape, as well as physics-based models designed to simulate runoff, erosion, and debris flow processes in burned areas, and attempts to parameterize temporal changes in key model input parameters as the landscape recovers. We highlight existing gaps in our ability to assess post-fire debris-flow hazards and motivate the need to expand our ability to use numerical modeling tools and quantify uncertainties in model outputs in order to support post-fire decision support systems.

Luke McGuire
University of Arizona
lmcguire@arizona.edu

MS135

Building Computer Models for Rain Induced Flows in Burnt Areas

We will present the careful development of models for post wild-fire debris flows in the framework of the popularly used TITAN2D code. These are informed by sparse data from recent events. The data are restricted to observations of flow footprints. Inputs to the model include difficult to characterize rainfall rates, rheology and entrainment parameters and topography. Our models must represent these well so further analysis of the uncertainty and hazards may carried out.

Palak Patel, Abani Patra
Tufts University
palak_kalpeshbhai.patel@tufts.edu,
abani.patra@tufts.edu

MS135

Characterizing Parallel Partial Emulation

Gaussian Process emulators (GPs) are a non-parametric regression technique that provides the best, linear, unbiased estimator of a scalar output, given data. In the classical approach of emulation, predicting the output at, say, every point of a spatial grid in a simulation of PDEs requires each space point to be its own dimension, increasing the required computational work to make such a prediction impractical. Recent work by Gu, Berger, and colleagues has proposed an extension of GP emulation to predict the functional output of simulations, with feasible computational work, by a clever sharing of certain parameters computed in the GP construction. In order to use this extended GP emulator in uncertainty analysis it is helpful to examine properties of the predicted solutions How well do the predicted solutions satisfy the underlying differential equation? Are properties of the differential equation such as conservation satisfied by the predicted solutions? We examine these questions through simple analysis and computational studies.

E Bruce Pitman
University at Buffalo
pitman@buffalo.edu

MS135

Emulators of Post-Fire Debris-Flows

Numerical models of post-fire debris-flow bulking and runout are computationally expensive. These models depend on poorly constrained and difficult to measure parameters related to fire-altered soil and vegetation, some of which change in time. Further, the development of debris flows (as opposed to clear flows) also depends on the rainfall intensity of potential storms. To date, modeling based hazard analysis has focused on if" a debris-flow will might triggered on a given fire scared hillside, and not on the extent or footprint of potential debris-flow runouts. We employ Gaussian process emulators to high-dimensional debris-flow model output to quantify uncertainties and aid in model-based hazard assessments of post-fire debris-flow inundation hazard assessments.

Elaine Spiller
Marquette University
elaine.spiller@marquette.edu

MS136

Risk Averse Engineering Design Using MLMC Estimators

Accurate and efficient estimation of risk measures and their sensitivities with respect to design parameters is important to the robust design of civil engineering structures. Gradient-based algorithms can be used to solve these optimal design problems, for which knowledge of these sensitivities are essential. Monte Carlo methods have been used for the statistical estimation of these risk measures and their sensitivities since the input uncertainties are often high-dimensional. Multi-Level Monte Carlo (MLMC) methods improve upon the performance of Monte Carlo methods by using a sequence of approximations of the underlying problem to reduce the cost of solution. We propose novel MLMC estimators and error estimators for risk measures such as the Conditional-Value-at-Risk (CVaR), as well as their sensitivities. We also present a Continuation-MLMC (C-MLMC) algorithm to adaptively calibrate the parameters of the MLMC estimator. C-MLMC algorithms rely on error estimators and adaptivity to achieve a prescribed tolerance with minimal computational cost. We will also show how MLMC estimators can be used within a gradient-based algorithm for robust design. We demonstrate a simple example of optimal design, as well as an example that is practically relevant to the engineering community.

Sundar Ganesh
EPFL
Switzerland
sundar.ganesh@epfl.ch

MS136

Improving Digital Twins by Learning from a Fleet of Assets

A digital twin is an evolving virtual model of a system or physical asset that assimilates data during the system's lifecycle so that it becomes an asset-specific model that underpins intelligent automation and drives key decisions. The digital twin concept was devised under the observation that information based on fleet (class) statistics is often not useful for assessing the health of an individual asset, since

the condition of assets varies across a fleet due to variability in manufacturing and operating conditions. While it is true that the average health of the asset class cannot be used to reliability certify the health of a single asset, data from multiple assets can be used in combination with local asset data to improve a digital twin. In this talk we leverage the digital twin concept to improve inferences about a single asset by utilizing observations of similar assets within an asset class. With this goal, we build networks of digital twins that encode the conditional dependence between the states of each asset in a class. We present numerical examples that highlight the generality of our approach through its application to model problems.

Daniel T. Seidl
Sandia National Laboratories
dtseidl@sandia.gov

John D. Jakeman
Sandia National Labs
jdjakem@sandia.gov

MS136

Multi-Index Sequential Monte Carlo Ratio Estimators for Bayesian Inverse Problems

In this talk, we consider the problem of estimating expectations of some quantity of interests with respect to a target distribution with an unknown normalising constant. We work under the assumption that the unnormalised target is intractable and needs to be approximated at finite resolution. Under such an assumption, our ratio estimators are constructed by applying a multi-index Sequential Monte Carlo method which combines the multi-index Monte Carlo for complexity improvement and Sequential Monte Carlo for efficient inference. With theoretical results, the multi-index Sequential Monte Carlo ratio estimators can achieve canonical complexity of MSE^{-1} for many problems where the existing methods require $\text{MSE}^{-\xi}$. Numerical results are shown on examples of Bayesian inverse problems.

Kody Law
University of Manchester
kodylaw@gmail.com

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

Neil Walton
School of Mathematics
The University of Manchester
neil.walton@manchester.ac.uk

Shangda Yang
University of Manchester
shangda.yang@manchester.ac.uk

MS137

Application of AdaAnn with Normalizing Flows to Approximate Probability Density Functions

Normalizing flows are invertible mappings used to transform simpler probability densities into ones that are more complex. Through optimizing parameters associated with these mappings, normalizing flows are used in statistics and machine learning for density estimation and variational inference. In the framework of variational inference, to improve the efficiency and accuracy of the approximation via normalizing flows especially when a target distribution is multimodal, annealing of the target distribution with a constant annealing schedule in the optimization is often employed. On the other hand, a constant annealing schedule often applies slowly-changing temperatures to the target distribution and can be inefficient computationally. In this talk, we will introduce a more efficient adaptive annealing scheduler called AdaAnn that automatically adjusts the incremental step in the annealing schedule to the KL-divergence between two adjacent tempered distributions. We will demonstrate the computational efficiency of normalizing flows with AdaAnn in approximating multimodal distributions and obtaining Bayesian inferences of the parameters for dynamical systems.

Emma Cobian
University of Notre Dame
ecobian@nd.edu

Jonathan Hauenstein
University of Notre Dame
Dept. of App. Comp. Math. & Stats.
hauenstein@nd.edu

Fang Liu, Daniele E. Schiavazzi
University of Notre Dame
fang.liu.131@nd.edu, dschiavazzi@nd.edu

MS137

Probabilistic Approaches to Transfer Learningfor Sparse and Noisy Data Environments

Machine learning (ML) models have thus far been applied to tasks and domains that, while impactful, have sufficient volume of data. For predictive tasks of national security relevance, ML models of great capacity are often needed to capture the complex underlying physics. Such models normally require an abundance of training data to exhibit sufficient predictive accuracy, which might not be available due to (1) excessive expense of computer simulations, (2) prohibitive experimental data acquisition cost, or (3) limited access to classified/sensitive data. To alleviate such difficulties, transfer learning (TL) may be used in which similar data from existing datasets or domains is used. We present a novel probabilistic TL framework to enhance the trust in ML models within noisy and sparse data settings. The framework will assess when it is worth applying TL, which ML model to use in TL, and how much knowledge is to be transferred. We rely on extensions of concepts and techniques from the fields of Bayesian inversion, sequential data assimilation, uncertainty quantification, and information theory. We provide insights through an application to polynomial-based surrogate model construction. We investigate that extent to which TL alleviates sparsity in training data that may jeopardize the reliability of such surrogates.

Mohammad Khalil, Wyatt H. Bridgman
Sandia National Laboratories
mkhalil@sandia.gov, whbridg@sandia.gov

Reese Jones
Co-author
rjones@sandia.gov

Ahmad A. Rushdi
Stanford University

rushdi@stanford.edu

MS137

Convergence of Deep Neural Networks: a Big Gap Between Theory and Practice

Mesh refinement provides a cornerstone of verification and validation for FEM; following from the consistency and stability of Galerkin methods, demonstrating convergence of a result on a sequence of refined grids is establishes that a solution is accurately arriving at a discrete solution to the continuous problem. For neural networks a number of approximation theory results establish best approximation results, namely, existence of neural networks which converge in Sobolev norms with respect to network width and depth. In practice however such optimal weights and biases are not obtained when training a network with gradient descent, leading to O(1) errors. Consequently, when using neural networks for regression and PDE solution, one cannot simply increase the size of the network to demonstrate convergence, leading to a new source of uncertainty. In this talk we provide an overview of this approximation theory in the literature and survey our recent work remedying this in practice.

Nathaniel Trask
Sandia National Laboratories
natrask@sandia.gov

MS138

Friends Dont Let Friends Deploy Black-Box Models: The Importance of Intelligibility in Machine Learning - Part I

In machine learning often tradeoffs must be made between accuracy, privacy and intelligibility: the most accurate models usually are not very intelligible or private, and the most intelligible models usually are less accurate. This can limit the accuracy of models that can safely be deployed in mission-critical applications such as healthcare where being able to understand, validate, edit, and trust models is important. EBMs (Explainable Boosting Machines) are a recent learning method based on generalized additive models (GAMs) that are as accurate as full complexity models, more intelligible than linear models, and which can be made differentially private with little loss in accuracy. EBMs make it easy to understand what a model has learned and to edit the model when it learns inappropriate things. In the talk Ill present case studies where EBMs discover surprising patterns in data that would have made deploying black-box models risky. Ill also show how were using these models to uncover and mitigate bias in models where fairness and transparency are important.

Rich Caruana
Microsoft Research
rcaruana@microsoft.com

MS138

Quantifying Reliability of Machine Learning Predictions via Physics-Informed Metrics

Machine learning (ML) based models are adopted to analyze direct image data from experimental science, including cold atom studies. This type of data suffers information losses from the imperfections of the techniques we used to prepare and measure them and we sometimes fail to leverage our simulation techniques to interpret them. In this talk, I will discuss how we adopt an ML model to detect multiple solitons in time-of-flight absorption images in order to automate related experiments. Furthermore, we quantify the reliability of such a model by combing it with a physics-informed module with our comprehension of the physical phenomena. In addition to the quantity of detected excitation, this module can classify each solitonic feature into one of three categories: kink soliton, solitonic vortex, or "partial" soliton. At last, we provide an open-source python package for solitonic excitation detection from both the pre-trained model with our data and any user-defined cold atom absorption image data.

Shangjie Guo
University of Maryland
shangjie.guo@gmail.com

MS138

Friends Dont Let Friends Deploy Black-Box Models: The Importance of Intelligibility in Machine Learning - Part II

In machine learning often tradeoffs must be made between accuracy, privacy and intelligibility: the most accurate models usually are not very intelligible or private, and the most intelligible models usually are less accurate. This can limit the accuracy of models that can safely be deployed in mission-critical applications such as healthcare where being able to understand, validate, edit, and trust models is important. EBMs (Explainable Boosting Machines) are a recent learning method based on generalized additive models (GAMs) that are as accurate as full complexity models, more intelligible than linear models, and which can be made differentially private with little loss in accuracy. EBMs make it easy to understand what a model has learned and to edit the model when it learns inappropriate things. In the talk Ill present case studies where EBMs discover surprising patterns in data that would have made deploying black-box models risky. Ill also show how were using these models to uncover and mitigate bias in models where fairness and transparency are important.

Harsha Nori
Microsoft Research, U.S.
hanori@microsoft.com

MS138

Building a Noise-Tolerant Framework for Quantum Dot Auto Tuning

Gate-defined quantum dots (QDs) are a promising quantum computing platform, yet initialization of these devices is non-trivial. While there has been some progress toward autonomous tuning, most of the proposed approaches lack an assessment of data or tuner reliability. This leads to unexpected failures when imperfect data is processed by an autonomous system. We demonstrate a framework for robust QD autotuning by pairing a machine learning (ML) state classifier with a data quality control module. The latter subsystem verifies image quality before processing by the state classifier, avoiding potential high uncertainty classification of imperfect data. To train both ML systems, we enhance the QD simulation by incorporating synthetic noise typical of QD experiments. Thus, we prevent high uncertainty predictions by leveraging outside knowledge about the possible types of noise in our data. We confirm that the augmentation of our data with synthetic noise significantly improves the performance of our state classifier, resulting in an accuracy of 95.1(7) % when tested on ex-

perimental data. We then validate the functionality of the data quality control module by showing the state classifier performance deteriorates with decreasing data quality, as expected. Our results establish a robust and flexible ML framework for autonomous tuning of noisy QD devices.

Joshua Ziegler
National Institute of Standards and Technology
joshua.ziegler@nist.gov

MS140

Learning Graphs, Density Estimation, and the Information Gap

Recent algorithms to learn the undirected graphical model (UGM) of a dataset rely on measure transport to represent the corresponding distribution. This initial density estimation step is of course expensive in high dimensions, and must happen before the real algorithm begins—computing the Hessian of the log density which reveals conditional independence (and thus the UGM). But consider that the UGM is a coarse summary of the density: a given density respects the conditional independence properties of a single graph, but any single graph is consistent with infinitely many densities. We call this the *information gap*, and claim that learning the graph should therefore be easier than learning a proper density estimate. In this work, we prove that, for a particular class of distributions, this is entirely possible—we can still, to known error, learn the correct graph with very cheap (incorrect) density approximations. In this talk, I'll review some of the theory, various numerical examples, and possible extensions of the information gap.

Rebecca Morrison
CU Boulder
rebeccam@colorado.edu

Ricardo Baptista
MIT
rsb@mit.edu

Estelle Basor
AIM
ebasor@aimath.org

MS140

Localized Transport Maps for Non-Gaussian Random Fields with Applications in Sea Ice

Gaussian processes and Gaussian Markov random fields are a common way of defining prior distributions for Bayesian inference problems. While ubiquitous, there are many applications where Gaussian priors cannot capture the structure observed in reality, even when transformations of the Gaussian distribution are employed (e.g., lognormal processes). We illustrate this point with a sea ice inverse problem where historic remote sensing data can be used for prior specification. We then discuss the use of transport maps for constructing non-Gaussian Markov random fields from historic data. Possible approaches for introducing additional structure like isotropy and stationarity will be introduced and demonstrated on a data fusion example for remote characterization of sea ice roughness.

Matthew Parno
Dartmouth College
matthew.d.parno@dartmouth.edu

Casey Dowdle
Cold Regions Research and Engineering Laboratory
caseyldowdle@gmail.com

Paul-Baptiste Rubio
MIT
rubiop@mit.edu

MS140

Optimal Transport Normalizing Flows

A data driven procedure is developed to compute the optimal map between two conditional probabilities $\varrho(x|z_1,\dots,z_L)$ and $\mu(x|z_1,\dots,z_L)$ depending on a set of covariates z_i. The procedure is tested on synthetic data from the 2017 ACIC Data Analysis Challenge for the estimation of treatment effect. Exactly solvable examples and simulations are performed to highlight the differences with ordinary optimal transport.

Giulio Trigila
Baruch college - CUNY
giulio.trigila@baruch.cuny.edu

Esteban G. Tabak
Courant Institute
New York University
tabak@cims.nyu.edu

Wenjun Zhao
Courant Institute of Mathematical Sciences
New York University
wenjun.zhao@nyu.edu

MS140

A Triangular Approach to Generative Modeling

The triangular map is a popular tool for transforming any source probability density to any target density. Based on triangular maps, we propose a general framework for high-dimensional density estimation, by specifying one-dimensional transformations (equivalently conditional densities) and appropriate conditioner networks. This framework (a) reveals the commonalities and differences of existing autoregressive and flow based methods, (b) allows a unified understanding of the limitations and representation power of these recent approaches and, (c) motivates us to uncover a new Sum-of-Squares (SOS) flow that is interpretable, universal, and easy to train. We further show that the density quantile functions of the source and target density provide a precise characterization of the slope of transformation required to capture tail behaviours, and propose tail-adaptive flows that adjust the source density and the triangular map simultaneously. Lastly, we argue that the triangular map also leads naturally to a multivariate extension of the univariate quantile function and we illustrate its relevance in novelty detection applications.

Yaoliang Yu, Priyank Jaini
University of Waterloo
Vector Institute
yaoliang.yu@uwaterloo.ca, pjaini@uwaterloo.ca

Jingjing Wang
University of Waterloo
jingjing.wang@uwaterloo.ca

Sun Sun

National Research Council Canada
sun.sun@uwaterloo.ca

MS141

Bayesian Estimation and Forward Propagation of Data-Driven Potential Uncertainty in Lennard-Jones Cluster Dynamics

A common way of representing the long-time dynamics of materials is in terms of a Markov chain over metastable states. This chain can either be used to generate trajectories using kinetic Monte Carlo, or analyzed directly, e.g., in terms of first passage times between distant states. While a number of approaches have been proposed to infer such a representation from direct molecular dynamics (MD) simulations, challenges remain. For example, chains inferred from a finite amount of MD will in general be incomplete, leading to uncertainties in the inferred quantities of interest (QoI). We quantify the local completeness of the chain in terms of Bayesian estimators of the yet unobserved rate, and its global completeness in terms of the residence time of trajectories within the explored subspace. This allows ones to characterize the uncertainty on the measured QoI, as well as to develop optimal sampling strategies to further improve the accuracy of the target QoI at the smallest computational cost. We illustrate our approach with examples from materials science, including the computation of breakup times of defect clusters, as well as of transport coefficients.

Dallas Foster
Massachusetts Institute of Technology
fostdall@mit.edu

MS141

Uncertainty Quantification in First-Principle Materials Modelling: Overview and Challenges

First-principle material modelling involves a coupling of multiple computational layers. For example the determination of macroscopic observables requires long-running molecular dynamics (MD) simulations. These are driven using interatomic potentials, which are learned based on data, which in turn is generated using computations, e.g. density-functional theory (DFT). Overall a sizable number of parameters are needed to tune physical approximations and the numerics to solve them. The ability to address uncertainties across these scales is thus crucial to make reliable parameter choices. However, systematic approaches for uncertainty quantification (UQ) are underdeveloped. One challenge is that workflows are too computationally demanding to allow for black-box sampling techniques. To make progress it is thus key to gain insight on reduced problems. Moreover an optimal exploitation of tools such as algorithmic differentation are needed e.g. to unlock invasive techniques. This contribution will present an overview of standard materials modelling and associated sources of uncertainty. It will discuss recent efforts within the multidisciplinary CESMIX project to develop composable software for materials simulation written in Julia. Main goal is to support both established software (such as LAMMPS) as well as emerging Julia codes (such as DFTK.jl). This facilitates the development of UQ on reduced problems, but still allows a future integration into standard packages of the field.

Michael F. Herbst
RWTH Aachen University
Germany
herbst@acom.rwth-aachen.de

MS142

Assessing the Risk of Data-Driven Optimal Designs Using Scenario Theory

This talks outlines the foundations of scenario optimization theory and exemplifies its usage by means of a system identification example. The scenario optimization approach is a technique for obtaining solutions to robust optimization and chance-constrained optimization problems based on a sample of the constraints. The approach not only yields an optimal solution but also a formally-verifiable, non-asymptotic, distribution-free certificate of its correctness. Such a certificate is of paramount importance in assessing the risk of solutions for safety-critical applications. For constraints that are convex a deep theoretical analysis has been established which shows that the probability of a new constraint being violated follows a distribution that is dominated by a Beta distribution. This result is exact for a whole class of convex problems. Extensions to non-convex optimization have been recently developed. This framework will be illustrated by building a computational model of a dynamic system according to multivariable input-output data. An Interval Predictor Model prescribes the parameters of a computational model as a set thereby making each predicted output an interval-valued function of the input. The formulation proposed seeks the parameter set leading to the tightest enclosure of the data. Scenario theory is used to compute an upper bound on the probability of future data falling outside the predicted interval.

Luis G. Crespo
NASA Langley
luis.g.crespo@nasa.gov

MS142

Data-Driven Predictions in Safety-Critical Systems: Challenges and Opportunities

This talk outlines challenges and opportunities that arise in safety-critical systems when data-driven predictions are introduced, and some reflections from practical UQ applications. Safety-critical systems are all around us, from commercial airplanes to nuclear power plants. Historically, the safety of such systems has been addressed by analysis of potential scenarios to which the system can be subjected, and cautious engineering of the capacity of these systems. As the complexity of our engineering systems increases, and are more interconnected and controlled by computers, our human minds have become hard-pressed to cope with this enormous and dynamic complexity. Data-driven modeling, such as ML, has revolutionized the way we approach such problems and may soon become integral to many safety-critical systems. However, it also introduces additional uncertainties that must be properly handled. This is essential if we want to be confident that future engineering systems remain sufficiently safe. There are different approaches to increase confidence. This could be from statistical guarantees to ensure robustness, or by uncertainty reduction, for instance through optimal experimental design. There are also methods that are not directly linked to data uncertainty. For instance, leveraging the knowledge of the relevant physical system through physics-based constraints, or explainable model diagnostics. From a more general safety perspective, this may be just as important.

Carla J. Ferreira

DNV AS
carla.ferreira@dnv.com

Christian Agrell
DNV
christian.agrell@dnv.com

Simen Eldevik
DNV AS
simen.eldevik@dnv.com

MS142

A Machine Learning Approach to Safer Airplane Landings: Predicting Runway Friction Using XGBoost and Explainable AI

The presence of snow and ice on runway surfaces reduces the available tire-pavement friction needed for retardation and directional control and causes potential economic and safety threats for the aviation industry. To activate appropriate safety procedures, pilots need accurate and timely information on the actual runway surface conditions. This talk shows how machine learning (ML) can be used to predict complex physical phenomena such as surface friction, when the underlying physical model is difficult to retrieve. A problem when working with safety-critical systems is that we do not have sufficient experience data from dangerous situations, which in our case means very slippery conditions. We approach this problem by combining prediction models trained with different focus areas on subgroups of the data. The ML models are compared to several state-of-the-art runway assessment methods and outperform them all. However, the ML algorithms provide highly complex models, as scores from hundreds of decision trees are combined. This makes it difficult to interpret and trust how the predictions are made. To decrease the uncertainty involved in the use of ML algorithms, the prediction models are combined with SHAP approximations, which presents arguments for the predictions. This provides a comprehensible decision support system for airport operators and pilots, which can contribute to safer and more economic operations of airport runways.

Alise D. Midtfjord
University of Oslo
alisedm@math.uio.no

MS143

Learning Non-Intrusive Data-Driven Reduced Models for Design Optimization of Rotating Detonation Engines

The design optimization of rotating detonation engines is a formidably challenging task for a number of reasons. First, a single high-fidelity simulation can require up to millions of core hours even on massively parallel supercomputers, thus restricting the total number of possible simulations to only a handful. In addition, the optimization procedure usually depends on a large number of design parameters which makes it computationally prohibitive. To address these challenges, we (i) split the computational domain into separate components (injectors, combustion chamber and nozzle), (ii) construct a reduced model for the critical component, i.e., the combustion chamber and (iii) use simplified models for the remaining two components. To construct reduced models for the combustion chamber, we employ scientific machine learning and enhance scientific machine learning-based procedures with a simple yet powerful preprocessing step that leads to more accurate and predictive reduced models. We test the proposed approach in a complex and realistic rotating detonation engine scenario with more than 22 million degrees of freedom and show that a low cardinality training set is sufficient to construct a predictive reduced-order model.

Ionut-Gabriel Farcas
Technical University of Munich
Chair of Scientific Computing in Computer Science
ionut.farcas@austin.utexas.edu

Ramakanth Munipalli
Edwards Air Force Base
Air Force Research Laboratory
ramakanth.munipalli@us.af.mil

Karen E. Willcox
UT Austin
kwillcox@oden.utexas.edu

MS143

Adversarial Attacks on the Uncertainty Neural Network Interatomic Potentials for Active Learning

Neural networks (NNs) have proven extremely effective interpolators in the physical sciences. In particular, NN interatomic potentials can be trained to replicate the mapping of atomistic structure to energy (the so-called Potential Energy Surface) of expensive quantum mechanical methods. While they reduce the computational cost by orders of magnitude, NN potentials are brittle and struggle to generalize to points outside the training data (rare events) that maybe be visited when they are deployed in simulations. Here, we will describe how uncertainty quantification methods based on NN ensembles and on customized loss functions, including mean-variance estimation or evidential learning, enable gradient-based active learning to systematically improve NN potentials. Because the uncertainty metrics are differentiable end-to-end with respect to model inputs, it is possible to distort input atomic positions towards regions of high thermodynamic likelihood (low energy) and high uncertainty, so they can be labeled with new quantum chemical simulations. These adversarial attacks on uncertainty allow exploring the "phase space" of uncertainty much more efficiently than traditional approaches and the training of accurate and generalizable surrogate functions.

Aik Rui Tan, Daniel Scwhalbe-Koda
MIT DMSE
atan14@mit.edu, dskoda@mit.edu

Rafael Gomez-Bombarelli
Assistant Professor
Massachusetts Institute of Technology
rafagb@mit.edu

MS143

Particle-Based Methods for Inference and Real-Time Control of Dynamical Material Systems

Autonomous materials research platforms offer the opportunity to accelerate the discovery and design of new materials through optimal, closed-loop decision-making that strategically explores a priori unknown spaces of materials synthesis and processing recipes. In addition, newer platforms perform in-situ characterization and dynamical tuning of growth conditions, allowing an autonomous agent

to query the material system more actively, utilize and learn governing dynamical laws, and drive the evolution towards target states. Therefore, such platforms pose interesting problems in physics-based, real-time control and inference, above more traditional perspectives that view the system as a black box with fixed control parameters. This talk presents particle-based, Bayesian methods and episodic reinforcement learning policies to drive material systems in real-time under uncertainty while inferring effective dynamics and hidden state variables. We will present these methods applied to two problems: driving chemical reactions, and block copolymer self-assembly.

Kristofer Reyes
Buffalo University
kreyes3@buffalo.edu

MS144

Rank Based Estimator for Sobol Indices

In this talk, we will discuss our recent results on the Chatterjee estimator of the first order Sobol' index. This estimator is based on the pick-freeze method but uses only one sample. We will explain how it is possible to show a CLT for this estimator.

Fabrice Gamboa
Institut de Mathématiques de Toulouse. AOC Project
University of Toulouse
fabrice.gamboa@math.univ-toulouse.fr

Thierry Klein
ENAC (Ecole nationale de l'aviation civile) and IMT
University of Toulouse
thierry.klein@math.univ-toulouse.fr

Agnès Lagnoux
Institut de Mathématiques de Toulouse
Université Toulouse 2
lagnoux@univ-tlse2.fr

MS145

Parameter Subset Selection for a Mathematical Model of Antibody Therapies for Neurological Diseases

A significant challenge in the development of drugs to treat central nervous system (CNS) disorders is to attain sufficient delivery of antibodies across blood-brain barriers (BBB). Since not all antibodies can pass through BBB, it is crucial to understand antibody exposure in the CNS quantitatively to construct drug characteristics and identify proper dosing regimens. We focus on a minimal physiologically-based pharmacokinetic (mPBPK) model of the brain for antibody therapeutics, which was developed by Bloomingdale, Bakshi, Maass, et al. (2021). This model is the reduced form of an existing multi-species platform brain PBPK model. The original PBPK model consists of 100 differential equations while the mPBPK model contains 16 differential equations which improves the speed of simulations. The model includes thirty one parameters and their values are obtained from the original brain PBPK model. In this presentation, we will discuss the use of a sensitivity-based parameter subset selection algorithm to determine those parameters which are identifiable in the sense that they can be uniquely determined by data. We illustrate this for ascending human doses. Issues to be discussed include the computation of sensitivities using sensitivity equations and complex-step approximations and local versus quasi-global analysis. We will also discuss qualitative verification techniques as well as quantitative techniques based on energy statistics.

Kamala Dadashova, Ralph Smith
North Carolina State University
kdadash@ncsu.edu, rsmith@ncsu.edu

MS145

Sensitivity Analysis of a Model for Anti-PCSK9 tTherapy to Lower LDL Cholesterol

PCKS9 plays an important role in the bodys low density lipoprotein cholesterol (LDLc) regulation, which greatly impacts risk of hyperlipidemia and cardiovascular disease. A quantitative systems pharmacology (QSP) model of production and removal of cholesterol in an individual has suggested that an anti-PCSK9 therapy may be an effective treatment of these diseases (Gadkar et al. 2014). Like many QSP models, this model includes a number of parameters that are biological in nature. In practice, many of those are calibrated and held constant in practice. Yet, a full sensitivity analysis suggests the model is quite sensitive to some of these parameters. We consider an active subspace technique that allows us to better control parameter range selection and consider additional effects of combined changes in parameter values. We will also show results from a partial rank correlation coefficient analysis and describe how these methods can be used to guide Gaussian process-based surrogate construction. Further, we will provide a comparison of parameter sensitivities under a variety of potential dosing regimens and discuss the suggested efficacy of these regimens at reducing the steady-state levels of LDLc.

Eli Horner
Marquette University
eli.horner@marquette.edu

Patrick Hanafin
University of North Carolina
patrickh@email.unc.edu

Jaimit Parikh
GSK
jaimit.parikh@gmail.com

Elaine Spiller
Marquette University
elaine.spiller@marquette.edu

MS145

Identifying Therapeutic Targets for Atopic Dermatitis (AD) Using Sensitivity Analysis (SA)

AD is an inflammatory skin disease with a worldwide prevalence of 525% (Deckers IA et al. PLoS One 2012;7:e39803). AD pathogenesis involves epidermal barrier abnormalities, loss of topical, symbiotic microflora, and immunological dysregulations characterized by immune activation and inflammatory cytokine production. Several therapies have been developed to alleviate AD flare-ups. Our objective was to perform SA on an existing AD mathematical systems model to identify influential parameters of AD disease pathogenesis, and thus druggable targets. The previously-developed model described AD pathogenesis with 14 equations and 51 parameters, incorporating skin barrier integrity, infiltrated pathogens, cytokines, and T cells (Miyano T et al. Allergy 2021;00:1-13). The

clinical endpoint, percent change in Eczema Area and Severity Index score at 24 weeks (PIE24), is derived from model-predicted skin permeability and infiltrated pathogen scores. We performed SA on the AD model using a "one-at-a-time" local SA and global SA (partial rank correlation coefficients (PRCC) and the Morris method) in MATLAB (v2021a) to ascertain parameters most influential of PIE24. T cell elimination rate ($|PRCC|$:0.92), pathogen elimination via skin turnover (0.89), IL-13 elimination (0.89), and naïve T cell differentiation into Th2 cells (0.85) were most influential of PIE24. Future therapeutic development should target these pathways to maximize impact on AD disease pathogenesis and reduce AD flare-ups.

Patrick Hanafin
University of North Carolina
patrickh@email.unc.edu

Jaimit Parikh
GSK
jaimit.parikh@gmail.com

Gauri Rao
University of North Carolina
gaurirao@live.unc.edu

Ralph C. Smith
North Carolina State Univ
Dept of Mathematics, CRSC
rsmith@ncsu.edu

Helen Moore
Applied BioMath
helen.moore@medicine.ufl.edu

MS145

Applications of Sensitivity Analysis on a QSP Model of Bone Mineral Density Loss

Osteoporosis, caused by decreased bone mineral density (BMD), results in increasing bone fractures, associated mortality rates, and health-care costs. The screening and development of osteoporosis treatments has relied heavily on finding predictive biomarkers of BMD loss. Mathematical models incorporating multiple bone turnover markers have been developed to predict BMD loss and related physiological phenomena. However, due to high complexity and variability, the key factors that cause BMD loss have not yet been determined. Here, we performed both local and global sensitivity analyses on a QSP bone model [Hasegawa and Duffull, CPT:PSP, 2018] to identify the most significant physiological factors that influence BMD changes under denosumab treatment. For local sensitivity analysis, we introduced a perturbation of 10% and 20% on each parameter in the model and compared the changes of model outcomes. We used partial rank correlation coefficient global sensitivity analysis to determine relationships between the parameters and system outputs [Gallaher et al., JTB, 2018]. Our analysis indicated that the osteoblast effect on BMD production and osteoclast effects on BMD degradation were most significantly associated with BMD changes. To summarize, we explored a QSP bone model and evaluated model parameter influence on BMD. Our work indicates potential targets for future osteoporosis therapies.

Jiawei Zhou, Patrick Hanafin
University of North Carolina
zhoujw@unc.edu, patrickh@email.unc.edu

Jaimit Parikh
GSK
jaimit.parikh@gmail.com

Ralph Smith
North Carolina State University
rsmith@ncsu.edu

Helen Moore
Applied BioMath
helen.moore@medicine.ufl.edu

MS146

Accrue: Accurate and Reliable Uncertainty Estimate in Deterministic Models

We focus on the problem of assigning uncertainties to single-point predictions generated by a deterministic model that outputs a continuous variable. This problem applies to any state-of-the-art physics or engineering models that have a computational cost that does not readily allow running ensembles and estimating the uncertainty associated to single-point predictions. Essentially, we devise a method to easily transform a deterministic prediction into a probabilistic one. We show that for doing so, one has to compromise between the accuracy and the reliability (calibration) of such a probabilistic model. Hence, we introduce a cost function that encodes their trade-off, and we call this new method ACCRUE (ACCurate and Reliable Uncertainty Estimate). We use the continuous rank probability score to measure accuracy and we derive an analytic formula for the reliability, in the case of forecasts of continuous scalar variables expressed in terms of Gaussian distributions. The new ACCRUE cost function is then used to estimate the input-dependent variance, given a black-box oracle mean function, by solving a two-objective optimization problem. We show several examples both with synthetic data, where the underlying hidden noise can accurately be recovered, and with large real-world datasets.

Enrico Camporeale
University of Colorado at Boulder
enrico.camporeale@noaa.gov

MS146

Learning Ensembles of Probabilistic Predictions for Multiple-Days Ahead Geomagnetic Indices Forecast

We present a new model for the probability that the Disturbance storm time (Dst) index exceeds a certain threshold, with a lead time between 1 and 3 days. Dst provides essential information about the strength of the ring current around the Earth caused by the protons and electrons from the solar wind, and it is routinely used as a proxy for geomagnetic storms. The model is developed using an ensemble of Convolutional Neural Networks (CNNs) that are trained using SoHO images (MDI, EIT and LASCO). The relationship between the SoHO images and the solar wind has been investigated by many researchers, but these studies have not explicitly considered using SoHO images to predict the Dst index. This work presents a novel methodology to train the individual models and to learn the optimal ensemble weights iteratively, by using a customized class-balanced mean square error (CB-MSE) loss function tied to a least-squares (LS) based ensemble. The proposed model can predict the probability that Dst ¡ -100nT 24 hours ahead with a True Skill Statistic (TSS) of 0.62 and Matthews Correlation Coefficient (MCC) of 0.37.

The weighted TSS and MCC from Guastavino et al. (2021) is 0.68 and 0.47, respectively. An additional validation during non-Earth-directed CME periods is also conducted which yields a good performance. Finally, tests performed during different solar cycle periods are presented.

Andong Hu
Multiscale Dynamic Group,
Centrum Wiskunde & Informatica (CWI)
andong.hu@colorado.edu

Carl Shneider, Ajay Tiwari
Centrum Wiskunde & Informatica (CWI)
shneider.carl@gmail.com, ajaynld13@gmail.com

Enrico Camporeale
University of Colorado at Boulder
enrico.camporeale@noaa.gov

MS146

Dynamic Data Driven Thermospheric Mass Density Forecasting with Quantified Uncertainties

The United States Space Surveillance Network is currently tracking more than 29,000 resident space objects (RSOs). 57% of these RSOs lie in the low Earth Orbit (LEO) regime. These LEO RSOs are strongly perturbed by the Earths upper atmosphere. Determining and predicting the orbit of these RSOs require accurate estimates of the local thermospheric mass density. The thermospheric mass density is highly dynamic and strongly driven by solar and geomagnetic activity. Traditionally, the thermospheric mass density is estimated using empirical models or physics-based models. Empirical models can provide fast predictions; however, they have limited accuracy. Meanwhile, physics-based models have good forecasting capabilities but are computationally expensive and require parallel resources for real-time evaluation. In this work, a data-driven dynamic reduced order thermospheric density model is developed to improve the forecasting capability while keeping computational complexity low. The dynamic nature of the model allows the current mass density field to evolve in time based on the evolution of space weather indices as well as assimilating external measurement sources such as GPS data, satellite range, and two-line element set (TLE). A data-driven reduced-order model is used to approximate the high dimensional thermospheric mass density field while preserving the essential behavior and dominant effects of the thermospheric mass density field to keep the computational requirement low.

Peng Mun Siew, Richard Linares
Department of Aeronautics and Astronautics
Massachusetts Institute of Technology
siewpm@mit.edu, linaresr@mit.edu

MS147

Solution Verification of a Controlled Flow

We present error estimation for a feedback flow control problem, where the control is applied at the boundary. In particular, we predict the transport of the errors contained in the boundary conditions throughout the computational domain as well as the discretization error associated with the finite element approximation used for the simulation.

Jeff Borggaard
Virginia Tech
jborggaard@vt.edu

MS147

Verification of Anomalous Wave Solutions in Viscous and Inviscid Non-Ideal Gases

We verify one- and two-dimensional anomalous wave solutions in inviscid and viscous van der Waals gases. The anomalous waves considered include rarefaction shock waves and isentropic compression fans. These unusual waves will be shown to satisfy the second law of thermodynamics. An exact solution for an inviscid shock tube with a van der Waals equation of state will be presented and used for comparison with numerical solutions. Error convergence for numerical simulations of inviscid and viscous shock tube problems will be shown. Inviscid simulations are performed with a third order Runge-Kutta method in time and a fifth order Mapped Weighted Essentially Non-Oscillatory (WENO5M) discretization with global Lax-Friedrichs flux splitting in space. Viscous simulations are performed with a first order forward difference in time and a second order central difference in space. This simple method is sufficient to capture thin anomalous waves on a fine grid. The numerical solution is verified by comparison to an exact viscous solution.

Katherine Pielemeier, Alexander Davies
University of Notre Dame
kpieleme@nd.edu, adavies2@nd.edu

Joseph M. Powers
University of Notre Dame
Aero/Mech Engineering
powers@nd.edu

MS148

An Information Field Theory Interpretation of Physics-Informed Neural Networks

The objective of this work is to automate the discovery of a systems Lagrangian from data. We represent the Lagrangian density, a functional of the systems state variables and its spatial/temporal derivatives of all orders, using a deep neural network (DNN). The systems state variables, all vector/tensor functions of space and time, are also represented by DNNs. The true physical state is connected to the Lagrangian via the principle of least action, i.e., the true state is a stationary point of the spatio-temporal integral of the Lagrangian density with fixed initial and final conditions. Inspired by Information Field Theory, we recast the principle of least action in a relaxed probabilistic fashion and use a likelihood function to connect the systems state to the observed data. We characterize the posterior over the parameters of the Lagrangian DNN using variational Bayes. This posterior captures the epistemic uncertainty about the Lagrangian (model-form uncertainty) induced by the limited/noisy data.

Alexander Alberts, Ilias Bilionis
Purdue University
albert31@purdue.edu, ibilion@purdue.edu

MS148

Calibration of Physics Informed Computer Models with Functional Inputs

Bayesian calibration of a functional input to a time-consuming simulator based on a Gaussian process (GP)

emulator involves two challenges that distinguish it from other parameter calibration problems. First, one needs to specify a flexible stochastic process prior for the input, and represent it with a tractable number of random variables. Second, a sequential experiment design criterion that reduces the effect of emulator prediction uncertainty on calibration results is needed and the criterion should be scalable for high-dimensional input and output. In this research, we will introduce a new method to address these two issues. For the first issue, we employ a GP with a prior density for the correlation parameter as prior for the functional input, and the Karhunen-Love expansion of this non-Gaussian stochastic process to reduce its dimension. We show that this prior gives substantially more robust inference results than a GP with a fixed correlation parameter. For the second issue, we propose the weighted prediction variance criterion (with posterior density of the functional input as weight) and prove the consistency of the sequence of emulator-based likelihoods obtained with the criterion. The proposed approach is illustrated with examples on estimation of hydraulic transmissivity in groundwater flow.

Zhaohui Li
School of Data Science
City University of Hong Kong
zhaohui.li@gatech.edu

Matthias H. Tan
City University of Hong Kong
matthtan@cityu.edu.hk

MS148

Parametric Machine Learning Surrogates for Gravitational Wave Signals

Gravitational wave astronomy is the field concerned with learning features of the universe through observation and analysis of gravitational waves. Observatories capable of detecting these extremely weak signals have only existed for a few years, and there is currently a great need for accurate, fast, and versatile models in order to analyze the observed signals. Numerical solvers for the Einstein Field Equations are extremely computationally expensive and difficult to develop, so a broad array of surrogate models, based mostly on physical approximations, has been developed to make analysis of the signals tractable. In this work, we build on recent developments in using neural networks to learn the morphology of a signal generated by more expensive gravitational wave models. We extend these models to be accurate over a desired prior distribution of source parameters, so that they can be used in the context of solving a Bayesian inverse problem, which is the key objective of gravitational wave analysis.

Bassel Saleh
University of Texas at Austin
bassel@utexas.edu

Thomas O'Leary-Roseberry
The University of Texas at Austin
Oden Institute for Computational and Engineering Sciences
tom@oden.utexas.edu

Brendan Keith
Brown University
brendan_keith@brown.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS148

Model-Based Techniques for Radiation Source Localization in an Urban Environment

This presentation will focus on techniques to localize a radiation source in an urban environment. This problem is made challenging by the highly variable nature of buildings, the fact that vehicles may block detector paths, the presence of background radiation due to granite structures, and time-dependent changes in background radiation due to weather. Moreover, high-fidelity simulation codes are typically too computationally complex to permit real-time source localization. To address these issues, we will discuss the construction of surrogate models, which incorporate fundamental physics but run in fractions of a second. We will also discuss the use of Bayesian statistical algorithms to infer uncertainties associated with source locations and intensities. A final component of the presentation will focus on strategies to employ moving detectors, such as drones, to improve search capabilities.

Ralph Smith, Paul Miles, Chris Edwards, Jared A. Cook
North Carolina State University
rsmith@ncsu.edu, prmiles.dev@gmail.com, cjedwar3@ncsu.edu, jacook8@ncsu.edu

MS149

The UQTk C++/Python Toolkit for Uncertainty Quantification: Overview and Applications

The UQ Toolkit (UQTk) is a collection of libraries, tools and apps for the quantification of uncertainty in numerical model predictions. As one of the software tools offered by the DOE SciDAC FASTMath Institute, UQTk offers intrusive and non-intrusive methods for forward uncertainty propagation, tools for sensitivity analysis, sparse surrogate construction, low-rank-tensor approximations, Bayesian inference via various flavors of MCMC, model error assessment, as well as several other capabilities. The core libraries are implemented in C++ but a Python interface is available for easy prototyping and incorporation in UQ workflows. The talk will give an overview of UQTk capabilities and illustrate its application to representative scientific workflows.

Luke Boll
Sandia National Labs
lboll@sandia.gov

Katherine M. Johnston
Sandia National Laboratories
University of Washington
kjohnst@sandia.gov

Khachik Sargsyan, Cosmin Safta
Sandia National Laboratories
ksargsy@sandia.gov, csafta@sandia.gov

Bert J. Debusschere
Chemistry, Combustion and Materials Center
Sandia National Laboratories, Livermore CA

bjdebus@sandia.gov

MS149

OpenTURNS and Persalys: Open-Source Software for Advanced Uncertainty Quantification

Available both as a C++ and a Python library, OpenTURNS provides state-of-the-art tools for uncertainty propagation relying on probabilistic methods. Developed by a partnership of five industrial companies and institutions (EDF, Airbus, ONERA, Phimeca and IMACS), it evolves based on feedback from both engineers and research engineers. UQ algorithms related to central dispersion, probability of exceedance, reliability analysis, sensitivity analysis, surrogate models, functional modeling, calibration are efficiently implemented. Connecting a new simulator to OpenTURNS is easy thanks to its wrapper services. For users who do not want to deal with programming interfaces, Persalys, developed by EDF and Phimeca, provides a graphical interface for many OpenTURNS services. In this talk, we present the main Persalys features: central dispersion analysis, global sensitivity analysis, threshold probability estimate, calibration. We also present advanced Paraview-based graphical features, including the plot matrix view and the parallel coordinate plot. Features are available for data extracted from CSV files or callable physical models (some of them can even be used on physical models with time series output). Finally, we show how Persalys into HPC contexts with limited input from the user.

Joseph Muré, Michael Baudin, Anne Dutfoy, Anthony Geay, Ovidiu Mirescu
EDF R&D
joseph.mure@edf.fr, michael.baudin@edf.fr,
Anne.Dutfoy@edf.fr, anthony.geay@edf.fr,
ovidiu.mirescu@edf.fr

MS149

Cossan Software: a Pathway Towards a Modular and Portable Library Collection

Cossan software is a collaborative development aim at offering advanced and recent algorithms for performing risk, reliability, and uncertainty analysis of complex engineering systems. The software includes a user-friendly frontend known as COSSAN-X with associated tutorials and wizards designed to guide users through the different steps of the analysis making this tool ideal for industry and for training as well. The computational engine OpenCossan is released under the LGPL license and represents an ideal environment for research and academics to access the state-of-the-art algorithms for dealing with uncertainty. Although the Cossan software has been programmed in Matlab and Java using a modular approach it still a quite big and complex software, for unfamiliar users. Dedicated tools and libraries applications has been recently released and available on cossan.co.uk. Examples are: SMARTool for performing uncertainty quantification adopting Robust Neural Networks coupled with Adaptive Bayesian Model Selection, Bayesian and Credal Networks toolbox, Interval Predictor Model and Virtual Human expert classificator. Current work involves porting OpenCossan in Julia, implementing a probabilistic programming language as a tool to automate statistical inference problems and Probability bounds analysis allowing models to be described and composed using a simplified syntax and provides a representation of sets of distributions in structures called probability boxes

Temitope Ohiani
University of Strathclyde
temitope.ohiani@strath.ac.uk

Ander Gray
University of Liverpool
ander.gray@liverpool.ac.uk

Edoardo Patelli
University of Strathclyde
edoardo.patelli@strath.ac.uk

MS150

Interval and Fuzzy Physics-Informed Neural Networks for Uncertain Fields

Temporally and spatially dependent uncertain parameters are regularly encountered in engineering applications. Commonly these uncertainties are accounted for using random fields and processes which require knowledge about the appearing probability distributions functions which is not readily available. In these cases non-probabilistic approaches such as interval analysis and fuzzy set theory are helpful uncertainty measures. Partial differential equations involving fuzzy and interval fields are traditionally solved using the finite element method where the input fields are sampled using some basis function expansion methods. This approach however is reliant on knowledge about the spatial correlation fields. In this work we utilize physics-informed neural networks (PINNs) to solve interval and fuzzy partial differential equations. The resulting network structures termed interval physics-informed neural networks (iPINNs) and fuzzy physics-informed neural networks (fPINNs) show promising results for obtaining bounded solutions of equations involving spatially and/or temporally uncertain parameter fields. In contrast to finite element approaches no correlation length specification of the input fields as well as no averaging via Monte-Carlo simulations are necessary. In fact, information about the input interval fields is obtained directly as a byproduct of the presented solution scheme. Furthermore all major advantages of PINNs are retained and ease of inverse problem set-up.

Nikolaos Bouklas
Cornell University
nb589@cornell.edu

MS150

Bayesian Nonlocal Operator Regression (BNOR): Towards the Characterization of Uncertainty in Heterogeneous Materials

We consider the problem of modeling heterogeneous materials where small-scale dynamics and interactions affect the global behavior; these situations are ubiquitous in engineering and scientific applications. The material's microstructure, properties, interfacial conditions, and operating environments cause variability in the material's response; hence it is often non-practical, if not impossible, to provide quantitative characterization for each sample. The goal of this work is to develop a Bayesian framework to characterize the uncertainty of material response when using a nonlocal model to describe wave propagation through heterogeneous, disordered materials. Our approach is based on the nonlocal operator regression (NOR) technique, and Bayesian inference. Specifically, we use a

MCMC method to predict the probability distribution of the nonlocal constitutive law that embeds the material's properties. As an application, we consider the wave propagation problem in a heterogeneous bar with randomly generated microstructure layers. In particular, we apply the proposed approach to model the stress wave propagation and provide a characterization of the uncertainty in the material microstructure. With several numerical tests, we illustrate the effectiveness of our approach in predicting the posterior distribution of the optimal nonlocal model.

Yiming Fan
Lehigh University
yif319@lehigh.edu

Yue Yu
Department of Mathematics, Lehigh University
yuy214@lehigh.edu

Marta D'Elia, Stewart Silling
Sandia National Laboratories
mdelia@sandia.gov, sasilli@sandia.gov

Habib N. Najm
Sandia National Laboratories
Livermore, CA, USA
hnnajm@sandia.gov

MS150

An Efficient Multiscale Surrogate for Brittle Fracture Analysis

A critical component to the overall safety of structures is the behavior of fractures in material. Existing practices for fracture simulations employ deterministic approaches; however as a result of the obligation to quantify risk in regulatory decision-making, probabilistic approaches have become ubiquitous. A prime example of a random input is initial crack size, which is seldom accurately known and has a strong influence on lifetime. Crack initiation involves complex mechanisms of interatomic cohesion, mobility and interactions, which is driven by material parameters and often needs to be considered uncertain. Therefore, we need a direct representation of uncertainties using efficient approximating models and distributed inputs. Considering the high complexity of simulation techniques, efficiency is a difficult issue to address when parametric uncertainty is accounted. The aim of this work is to develop a DeepONet based surrogate model for mapping the fracture toughness, applied stress, flaw size, and component geometry to the failure path using probabilistic approaches. The DeepONet based surrogate model will couple atomisticcontinuum multiscale materials to analyze the crack propagation. The region near the crack tip is modeled using peridynamics (PD), and the region away from the damaged zone will be approximated as a continuum region.

Somdatta Goswami, Khemraj Shukla
Brown University
somdatta_goswami@brown.edu,
khemraj_shukla@brown.edu

Yue Yu
Department of Mathematics, Lehigh University
yuy214@lehigh.edu

George E. Karniadakis
Brown University
george karniadakis@brown.edu

MS150

An Asymptotically Compatible Probabilistic Collocation Method for Randomly Heterogeneous Nonlocal Problems

We consider nonlocal elliptic type problems in heterogeneous media with coefficients depending on finitely many parameters. The parameters are realizations of random variables which could come from a truncated Karhunen-Loeve decomposition of a random field. Following the work of Cohen, DeVore and Schwab (2011), we first show the analytic regularity of the dependence of the solution on the parameters in the coefficients. This leads to an algebraic or sub-exponential convergence rate in the parameter space by using the Smolyak-type sparse grid probabilistic collocation method. The spatial discretization of the nonlocal problem is done with an asymptotically compatible meshfree method which are robust under the change of a nonlocal horizon parameter charactering the length of nonlocality. The efficiency of the method is demonstrated for nonlocal diffusion and nonlocal mechanics problems.

Xiaochuan Tian
UC San Diego
xctian@ucsd.edu

MS151

Computer Model Calibration with Time Series Data Using Deep Learning and Quantile Regression

Computer models play a key role in many scientific and engineering problems. One major source of uncertainty in computer model experiment is input parameter uncertainty. Computer model calibration is a formal statistical procedure to infer input parameters by combining information from model runs and observational data. The existing standard calibration framework suffers from inferential issues when the model output and observational data are high-dimensional dependent data such as large time series due to the difficulty in building an emulator and the non-identifiability between effects from input parameters and data-model discrepancy. To overcome these challenges we propose a new calibration framework based on a deep neural network (DNN) with long-short term memory layers that directly emulates the inverse relationship between the model output and input parameters. Adopting the 'learning with noise' idea we train our DNN model to filter out the effects from data model discrepancy on input parameter inference. We also formulate a new way to construct interval predictions for DNN using quantile regression to quantify the uncertainty in input parameter estimates. Through a simulation study and real data application with WRF-hydro model we show our approach can yield accurate point estimates and well calibrated interval estimates for input parameters.

Won Chang
University of Cincinnati
changwn@ucmail.uc.edu

MS151

Reconstructing Turbulence with Deep Learning: Uncertainty Quantification and Outlook

Gaining awareness of the state of fluid flow is important in

controlling and analyzing turbulence. Achieving situational awareness of turbulent phenomena is challenging due to the rich nonlinear dynamics. We consider deep learning-based turbulent flow field reconstruction from sparse sensors. We first introduce the use of super-resolution analysis for reconstructing turbulent flows. This analysis reconstructs a high-resolution flow field from its low-resolution counterpart. To reconstruct flows with an arbitrary number of moving sensors, the Voronoi-tessellation projection for input measurements is also considered. This approach supports machine learning models to overcome a major difficulty with conventional methods in handling sparse sensors that are in motion without constant model retraining. We also assess uncertainties in data and model associated with machine learning-based flow reconstruction. For data uncertainty, we capitalize on probabilistic neural network to effectively quantify the uncertainty of unseen data using mixture Gaussian distributions. In contrast, Gaussian stochastic weight averaging is leveraged for model-form uncertainty. These techniques successfully provide physically interpretable confidence intervals while distinguishing the source of uncertainty of deep-learning flow reconstruction. We finally discuss the importance of data preparation with physics-inspired data scaling for deep learning-based turbulent flow analysis.

Kai Fukami
University of California, Los Angeles
kfukami1@g.ucla.edu

Romit Maulik, Nesar Ramachandra
Argonne National Laboratory
rmaulik@anl.gov, nramachandra@anl.gov

Masaki Morimoto
Keio University, Tokyo
spitz.lemon@keio.jp

Ricardo Vinuesa
KTH Royal Institute of Technology
rvinuesa@mech.kth.se

Koji Fukagata
Keio University, Japan
fukagata@mech.keio.ac.jp

Kunihiko Taira
University of California, Los Angeles
Mechanical and Aerospace Engineering
ktaira@seas.ucla.edu

MS151

High-Speed Simulation of Heart Valve Function Using a Neural Network PDE Approach

Due to the natural variations in structures, the mechanical behaviors of myocardium can vary dramatically within the heart. Thus, to obtain the responses of the myocardium with different realizations of structures, the resulting hyperelastic problem needs to be solved with spatially varying parameters and in certain cases different boundary conditions. To alleviate the associated computational costs at the time of simulation, we have developed a neural network-based direct PDE solution method. The resulting neural network was then trained in a physics-informed approach by searching for θ that minimizes the potential energy of the hyperelastic problem on the training dataset generated by sampling over the physiological range. The present method is intended for the low data problem; it does not require generating a large, labelled training datasets, which are also computationally intractable. The neural network model was trained with satisfactory convergence, it can be used to give fast predictions of complex 3D deformations in full kinematic space with population-based fiber structures by forward passes in the neural network. Due to their transfer learnability characteristics, the neural network on subsequent specimens more quickly. Scaled up for complete organ-level cardiac models to provide efficient and robust computational models for to improve patient outcomes in clinically relevant timeframes.

Michael Sacks
Oden Institute
UT Austin
msacks@oden.utexas.edu

MS152

Data-Driven Isogeometric Analysis for Quantifying High-Dimensional Uncertainty of Nonlinear System

Quantifying the Multivariate/high-dimensional input and output uncertainty that are omnipresent and have significant influence in particularly nonlinear system is quite challenging and little to be investigated, as where traditional model-based computational simulators are generally restricted. Therefore, this paper proposes a data-driven full-field Emulator to quantify high-dimensional material and load uncertainties in linear and nonlinear probabilistic (mechanics) system. Firstly, we generate raw training data, including the high-dimensional (material and load) uncertainty input and their (high-dimensional) full-field displacement output, by Monte Carlo simulator using isogemetric analysis. Secondly, we freshly transfer the map from raw uncertain parameters to full-field displacement, to the map from the reduced basis coefficients of input to the ones of output. Thirdly, a machine learning Emulator, from the viewpoint of Bayesian statistics, is built by the order-reduced data (reduced basis coefficients/features). Consequently, we can directly and fast obtain the system full-field solution to the newly given input using the built Emulator. Several engineering examples, involving both material and load uncertainties in both linear and nonlinear system, illustrated the significant performance in accuracy and efficiency.

Chensen Ding
University of Exeter
c.ding@exeter.ac.uk

MS152

An Efficient High-Order and Matrix-Free Isogeometric Galerkin Method for KarhunenLove Expansion

In recent work, the interpolation based quadrature of the weak form of the eigenvalue problem corresponding to the KarhunenLove expansion has been shown to perform remarkably well for the smooth Gaussian covariance kernel. In particular, the matrix-free isogeometric Galerkin method with interpolation based quadrature outperforms the matrix-free isogeometric collocation method in terms of the computational cost and the accuracy of the solution. The computational efficiency is achieved by leveraging tensor product splines, tensor contraction and a kernel interpolation scheme tailored specifically to the weak form of the eigenvalue problem. The computational cost of a single iteration of the eigenvalue solver scales with N^2, where N

denotes the number of degrees of freedom in the interpolation space. The cost is independent of the polynomial order of the solution and the interpolation space and thus enables high-order approximations. In order to solidify recent findings, we explore the possibilities and the limitations of such high-order approaches and extend previous benchmarks to additional kernels. The study is accompanied by a novel open-source computational framework, which in its core design supports development of fast isogeometric methods on Cartesian grids and tensor-product spaces and implements techniques such as sum-factorization based formation and assembly.

Michal L. Mika, René Hiemstra, Dominik Schillinger
Technische Universität Darmstadt, Germany
michal.mika@tu-darmstadt.de, hiemstra@mechanik.tu-darmstadt.de, dominik.schillinger@tu-darmstadt.de

MS152

Isogeometric Analysis of Diffusion Problems on Random Surfaces

We consider the numerical solution of diffusion equations on random surfaces within the isogeometric framework. Complex computational geometries, given only by surface triangulations, are recast into the isogeometric context by transforming them into quadrangulations and a subsequent interpolation procedure. Moreover, we describe in detail, how diffusion problems on random surfaces can be modeled and how quantities of interest may be derived. In particular, we propose a low rank approximation algorithm for the high-dimensional space-time correlation of the random solution. Numerical studies are provided to quantify and validate the approach.

Michael Multerer, Wei Huang
Università della Svizzera italiana
michael.multerer@usi.ch, wei.huang@usi.ch

MS153

Deep Learning of Unknown Systems with Noisy Data

Recently, a general framework has been developed for learning the evolution operator or flow map of an unknown time-dependent system from its trajectory data using deep neural networks. This allows for the creation of a predictive model for the system. Most works in this direction have strictly used noiseless data in training, which is rare in real applications. Therefore, in this talk we discuss the modifications required to account for noisy training and testing data within the existing framework. A variety of numerical examples are considered to support our findings.

Victor Churchill, Zhongshu Xu
The Ohio State University
churchill.77@osu.edu, xu.4202@osu.edu

Dongbin Xiu
Ohio State University
xiu.16@osu.edu

MS153

Improving Efficiency of Deep Learning Approaches for Scientific Machine Learning

Sparse reconstruction techniques from compressed sensing have been successfully applied to many application areas, including signal processing, inverse problems in imaging, and approximation of solutions to parameterized partial differential equations (PDE). Such approaches are capable of exploiting the sparsity of the signal to achieve highly accurate approximations with minimal sample complexity. For problems whose solutions possess a great deal of structure, their recovery properties can be further enhanced through a combination of carefully selected weighting or structured sampling schemes. Recently connections between compressed sensing and deep learning have been explored, and the existence of deep neural network (DNN) architectures which achieve the same sample complexity and accuracy as compressed sensing on function approximation problems have been established. In this work, we further explore these connections and sparse neural network approximation in the context of high-dimensional parameterized PDE problems. We provide a full error analysis for such problems, explicitly accounting for the errors of best approximation (describing DNN expressibility), spatial discretization of the PDE, and the algorithm used in solving the underlying optimization problem. We complement our theoretical contributions with detailed numerical experiments, demonstrating the potential for sparse neural network approximation in scientific machine learning contexts.

Nick Dexter, Ben Adcock, Juan M. Cardenas
Simon Fraser University
ndexter@sfu.ca, ben_adcock@sfu.ca, jcardena@sfu.ca

Simone Brugiapaglia
Department of Mathematics and Statistics
Concordia University
simone.brugiapaglia@concordia.ca

Sabastian Moraga
Department of Mathematics
Simon Fraser University, Canada
sebastian_moraga_scheuermann@sfu.ca

MS153

Level Set Learning with Pseudo-Reversible Neural Networks for Nonlinear Dimension Reduction in Function Approximation

Inspired by the Nonlinear Level set Learning (NLL) method that uses the reversible residual network (RevNet), in this paper we propose a new method of Dimension Reduction via Learning Level Sets (DRiLLS) for function approximation. Our method contains two major components: one is the pseudo-reversible neural network (PRNN) module that effectively transforms high-dimensional input variables to low-dimensional active variables, and the other is the synthesized regression module for approximating function values based on the transformed data in the low-dimensional space. The PRNN not only relaxes the invertibility constraint of the nonlinear transformation present in the NLL method due to the use of RevNet, but also adaptively weights the influence of each sample and controls the sensitivity of the function to the learned active variables. The synthesized regression uses Euclidean distance in the input space to select neighboring samples, whose projections on the space of active variables are used to perform local least-squares polynomial fitting. This helps to resolve numerical oscillation issues present in traditional local and global regressions. Extensive experimental results demonstrate that our DRiLLS method outperforms both the NLL and Active Subspace methods, especially when the target function possesses critical points in the interior of its input

domain.

Yuankai Teng
University of South Carolina
yteng@email.sc.edu

Zhu Wang
Department of Mathematics
University of South Carolina
wangzhu@math.sc.edu

Lili Ju
University of South Carolina
Department of Mathematics
ju@math.sc.edu

Anthony Gruber
Florida State University
anthony.gruber@fsu.edu

Guannan Zhang
Oak Ridge National Laboratory
zhangg@ornl.gov

MS154

Coherent Risk Assessment for Nonlinear Structural Analysis via Reduced-Order Models

Quantifying and accounting for risk in engineering design is a mission-critical task. Coherent risk measures satisfy mathematical axioms that make them advantageous for optimization, and also provide quantitative advantages as they take into account the magnitude of failure of a system. Coherent risk therefore measures an average of high-loss scenarios, and this tail integral is challenging to compute, especially for nonlinear dynamical systems. In this work, we present a multifidelity approach to coherent risk measure estimation, using error estimates to drive the reduced-order model construction. We illustrate our results on a nonlinear structural response analysis.

Dongjin Lee
University of California San Diego
Mechanical and Aerospace Engineering
dongjin-lee@uiowa.edu

Boris Kramer
University of California San Diego
bmkramer@ucsd.edu

MS154

Non-Intrusive Surrogate Modeling of Parametric Frequency Response Problems with Applications in Forward UQ

Numerical methods for time-harmonic wave propagation phenomena are often computationally intensive, expecially in mid- and high-frequency regimes, thus making a direct frequency response analysis prohibitively expensive. In this framework, model order reduction (MOR) methods are very promising: starting from few expensive solves of the problem, they can provide a reliable approximation of the frequency response of the system, very cheap to evaluate in a whole range of frequencies. In this talk, we describe an MOR approach for *parametric* frequency response problems, where the high-fidelity problem models not only the impact of the frequency on the system response, but also that of additional design and/or uncertain parameters. Our proposed method relies on *minimal rational interpolation* for the surrogate modeling of the frequency dependence, for few *fixed* values of the parameters. Then, the different surrogates are combined to obtain a global approximation with respect to both frequency and parameters. Our approach is *non-intrusive*, i.e., we do not require access to the matrices/operators defining the underlying high-fidelity problem, and allows for an *adaptive* selection of the sampled frequencies and parameters. Numerical examples in electrical circuit modeling and elasto-dynamics are also included, providing evidence of the approximation quality and computational efficiency of the surrogate model obtained with the proposed technique.

Davide Pradovera
EPFL Switzerland
davide.pradovera@epfl.ch

Fabio Nobile
EPFL, Switzerland
fabio.nobile@epfl.ch

MS154

Stabilized Reduced Order Methods for Transport Control Problems with Random Inputs

This talk focuses on weighted reduced order methods (w-ROMs) for parametrized advection-dominated Optimal Control Problems (OCP(μ)s) governed by stochastic Partial Differential Equations. This framework is a powerful tool to reliably fill the gap between the model and an observed solution. Studying such a complicated framework might face two issues: (I) the numerical instabilities due to the advection-dominated context, (II) the computational costs needed for statistical analysis. We tackle them through stabilized offline-online w-ROMs. They exploit the probability distribution of the parameters and the stabilized simulations to build a reduced space where faster (and stabilized) simulations are performed. [L. Venturi, D. Torlo, F. Ballarin, and G. Rozza, Weighted Reduced Order Methods for Parametrized Partial Differential Equations with Random Inputs. Uncertainty Modeling for Engineering Applications, F. Canavero (ed.), Springer International Publishing, pp. 2740, 2019] [D. Torlo, F. Ballarin, and G. Rozza, Stabilized Weighted Reduced Basis Methods for Parametrized Advection Dominated Problems with Random Inputs. SIAM/ASA Journal on Uncertainty Quantification, 6(4), pp. 1475-1502, 2018]. The methodology is validated employing several numerical test cases in the steady and time-dependent framework [F. Zoccolan, M. Strazzullo and G. Rozza. Stabilized Reduced Order Methods for Advection-Diffusion Optimal Control Problems with random inputs. In preparation, 2021].

Maria Strazzullo
mathLab, Mathematics Area, SISSA International School for Advanced Studies
mstrazzu@sissa.it

Fabio Zoccolan
Univeristy of Trieste
fabio.zoccolan@studenti.units.it

Gianluigi Rozza
SISSA, International School for Advanced Studies, Trieste, Italy
grozza@sissa.it

Davide Torlo

SISSA, International School for Advanced Studies,
Trieste, I
dtorlo@sissa.it

MS154

Convolutional Autoencoders for Reduced Order Modeling

In the construction of reduced-order models for dynamical systems, linear projection methods, such as proper orthogonal decompositions, are commonly employed. However, for many dynamical systems, the lower dimensional representation of the state space can most accurately be described by a *nonlinear* manifold. Previous research has shown that deep learning can provide an efficient method for performing nonlinear dimension reduction, though they are dependent on the availability of training data and are often problem-specific (see Lee and Carlberg, 2020). Here, we utilize randomized training data to create and train convolutional autoencoders that perform nonlinear dimension reduction for the wave and Kuramoto-Shivasinsky equations. Moreover, we present training methods that are independent of full-order model samples and use the manifold least-squares Petrov-Galerkin projection method to define a reduced-order model for the heat, wave, and Kuramoto-Shivasinsky equations using the same autoencoder.

Sreeram R. Venkat
University of Texas Austin
201 E 24th St, Austin, TX 78712
srvenkat@utexas.edu

Ralph Smith
North Carolina State University
rsmith@ncsu.edu

C.T. Kelley
North Carolina State Univ
Department of Mathematics
tim_kelley@ncsu.edu

MS155

ACE: Atomic Cluster Expansion

The traditional multi-scale modelling tools, such as densitity functional theory, empirical interatomic potentials, and various coupling and coarse-graining schemes have largely proven inadequate in adressing the need of computational materials and molecular modelling to bridge the scales from electronic structure to the meso-scale. Machine-learned interatomic potentials (MLIPs) provide maybe the first realistic scheme to bridge this gap. In this talk I will review a specific MLIP model, the Atomic Cluster Expansion (ACE), explain how it is ideally suited for a systematic approach to "learning" interatomic potentials. I will then outline our current efforts in integrating our ACE with statistical (Bayesian) modelling approaches.

Christoph Ortner
The University of British Columbia
ortner@math.ubc.ca

Matthias Sachs
University of British Columbia
m.sachs@bham.ac.uk

MS155

Uncertainty-Quantification-Driven Calculation of Long-Time Properties from Atomistic Simulations

A common way of representing the long-time dynamics of materials is in terms of a Markov chain over metastable states. This chain can either be used to generate trajectories using kinetic Monte Carlo, or analyzed directly, e.g., in terms of first passage times between distant states. While a number of approaches have been proposed to infer such a representation from direct molecular dynamics (MD) simulations, challenges remain. For example, chains inferred from a finite amount of MD will in general be incomplete, leading to uncertainties in the inferred quantities of interest (QoI). We quantify the local completeness of the chain in terms of Bayesian estimators of the yet unobserved rate, and its global completeness in terms of the residence time of trajectories within the explored subspace. This allows ones to characterize the uncertainty on the measured QoI, as well as to develop optimal sampling strategies to further improve the accuracy of the target QoI at the smallest computational cost. We illustrate our approach with examples from materials science, including the computation of breakup times of defect clusters, as well as of transport coefficients.

Danny Perez
Theoretical Division, T-1
Los Alamos National Laboratory
danny_perez@lanl.gov

Thomas Swinburne
Université Aix-Marseille, CNRS
swinburne@cinam.univ-mrs.fr

MS155

HAL: Hyperactive Bayesian Learning for Molecular Force Fields

We present a novel method for the efficient generation of data sets of atomic configurations for the purpose of fitting an atomistic force-field model. The method combines elements of active learning and adaptive biasing, and is formulated in a Bayesian framework. We explain the underlying ideas for the design of the method in the context of a simple synthetic toy problem. We demonstrate the practical applicability of the method in real data examples, and show dramatic performance gain in the assembly of relevant data sets in material science applications.

Cas Van Der Oord
University of Cambridge
casv2@eng.cam.ac.uk

Matthias Sachs
University of British Columbia
m.sachs@bham.ac.uk

Christoph Ortner
The University of British Columbia
ortner@math.ubc.ca

Gabor Csanyi
University of Cambridge
gc121@cam.ac.uk

MS155

Quantification and Propagation of Uncertainties in Machine Learning Interatomic Potentials for

Molecular Dynamics

Molecular dynamics (MD) simulations are often done using machine-learned interatomic potentials (MLIAPs) that are constructed from empirical and physical considerations, and fitted to data available from expensive ab initio quantum chemistry computations. These MLIAPs encapsulate the functional relationship between atomic configuration and potential energy of an atomic system, and are trained in a supervised machine learning context. Uncertainty quantification (UQ) for MLIAPs is useful for both training data selection in an active learning context, and the selection of MLIAP models of optimal complexity. Furthermore, MLIAPs equipped with UQ enable the propagation of uncertainty through MD simulations, thereby providing uncertainty estimates on MD simulation outputs. In this talk, we will discuss our work on a range of UQ approaches for MLIAPs and subsequent propagation of uncertainties through MD simulations. This includes Bayesian inference of MLIAP parameters via Markov chain Monte Carlo sampling, as well as approximate versions including variational inference and approximate Bayesian computation to help in the handling of highly overparameterized MLIAPs, such as those based on neural network forms. We will also explore ensemble methods such as query-by-committee, as a means of extracting MLIAP predictive uncertainties. We will demonstrate the results on material systems of interest, driven by fusion energy science applications.

Khachik Sargsyan
Sandia National Laboratories
ksargsy@sandia.gov

Habib N. Najm
Sandia National Laboratories
Livermore, CA, USA
hnnajm@sandia.gov

MS156

Learning Optimal Power Flow Solutions for Low-Carbon Power Grids

As power grids move towards integrating higher levels of low-carbon renewable energy sources, levels of intermittency and uncertainty in the power supply increase. Traditional methods of optimizing grid operations must adapt to these faster dynamics in order to maintain system stability and reliability. Recent advancements in machine learning applied to learning optimal power flow (OPF) solutions for grid dispatch, bypassing solving optimization problems directly, have shown great promise. However, many of these techniques only consider "snapshot' OPF solutions - ignoring practical power system constraints such as generator ramp limits and other intertemporal constraints. In this talk, we discuss how recurrent neural networks (RNNs) can be used to capture the time dependency of optimal generation dispatch solutions, incorporating constraints that span across time steps to predict solutions to OPF problems with high accuracy. The framework can provide grid operators with a potential solution to optimizing large power networks with complex constraints in near real-time as more intermittent energy sources come online in the coming decade.

Kyri Baker
University of Colorado, Boulder
kyri.baker@colorado.edu

Mostafa Mohammadian
University of Colorado Boulder
mostafa.mohammadian@colorado.edu

MS156

Predicting Power System Dynamics and Transients Using Fourier Neural Operators

The dynamics of a power grid are governed by a large number of nonlinear ordinary differential equations (ODEs). To safely operate the system, operators need to check that the states described by this set of ODEs stay within prescribed limits after various potential faults. But solving these ODEs are very time-consuming using numerical solvers, and machine learning approaches have been proposed to reduce computational times. Existing learning methods generally suffer from overfitting and failures to predict unstable behaviors. We present a framework for power system dynamic simulation by learning in the frequency domain, where the Fourier transform is across both time and the buses in the network. This allows us to predict the dynamic behavior of large systems with relatively small amount of training data. The system topology and fault information are encoded through a 3D Fourier transform. We show that the proposed approach can speed up the computation by orders of magnitude while also provide highly accurate simulations for different fault types.

Baosen Zhang
University of Washington
zhangbao@uw.edu

Wenqi Cui
University of Washington, U.S.
wenqicui@uw.edu

MS156

Fast Covariance Parameter Estimation of Spatial Gaussian Process Models Using Neural Networks

Gaussian processes are a popular model for spatially referenced data and allow descriptive statements, predictions at new locations, and simulation of new fields. Often, a few parameters are sufficient to parameterize the covariance function, and maximum likelihood (ML) methods can be used to estimate these parameters from data. Maximizing the likelihood, however, is computationally demanding. For example, in the case of local likelihood estimation, even fitting covariance models on modest size windows can overwhelm typical computational resources for data analysis. This limitation motivates the idea of using deep neural networks (DNNs) to approximate ML estimates. We train DNNs to take moderate size spatial fields or variograms as input and return the range and noise-to-signal covariance parameters. Once trained, the DNNs provide estimates with a similar accuracy compared to ML estimation and at a speedup by a factor of 100 or more. Although we focus on a specific covariance estimation problem motivated by a climate science application, this work can be easily extended to other, more complex, spatial problems and provides a proof-of-concept for this use of DNNs in computational statistics.

Florian Gerber
University of Zurich
florxgerber@gmail.com

Douglas Nychka
National Center for Atmospheric Research

nychka@ucar.edu

MS156

Learning to Optimize for Wireless Communications

We discuss applications of modern optimization techniques to new wireless communications tasks such as channel estimation, hybrid precoding, and resource management, particularly to 5G scenarios. In this research area, techniques such as "deep unfolding" prove very popular to improve the performance of communications systems as measured by Signal to Noise Ratio (SNR) even at very low SNRs (a regime of interest for 5G communications). Learning to optimize seems to be a promising route in the quest to have real-time processing in modern communication systems.

Cristian Rusu
University of Bucharest
cristian.rusu@unibuc.ro

MS157

Certifiable Risk-Based Engineering Design Optimization Using Trust-Regions

Reliable, risk-averse design of complex engineering systems with optimized performance requires dealing with uncertainties. A conventional approach is to add safety margins to a design that was obtained from deterministic optimization. Safer engineering designs require appropriate cost and constraint function definitions that capture the *risk* associated with unwanted system behavior in the presence of uncertainties. The work proposes two notions of certifiability. The first is based on accounting for the magnitude of failure to ensure data-informed conservativeness. The second is the ability to provide optimization convergence guarantees by preserving convexity. Satisfying these notions leads to certifiable risk-based design optimization (CRiBDO). In the context of CRiBDO, risk measures based on superquantile and buffered probability of failure are analyzed. CRiBDO is contrasted with reliability-based design optimization (RBDO), where uncertainties are accounted for via the probability of failure, through a thermal design problem. The CRiBDO formulations capture more information about the problem to assign the appropriate conservativeness, exhibit superior optimization convergence by preserving properties of underlying functions, and alleviate the adverse effects of choosing hard failure thresholds required in RBDO. The work further shows a CRiBDO reformulation approach that leads to convex risk-based design optimization using trust regions.

Anirban Chaudhuri
MIT
anirbanc@oden.utexas.edu

MS157

Iterative Sampling Methods for Reinforcement Learning

Reinforcement learning (RL) seeks to develop an optimal policy through which an agent can interact with their environment to maximize their cumulative reward. This policy is generated under uncertainty - the agent explores without knowledge of the best decisions in hopes of converging to an optimal policy. This is one reason that learning an optimal policy is challenging, especially when the policy is given by a function approximator, which is critical for high-dimensional problems. In this talk, we propose an approach to make RL policy learning easier. We leverage recently-developed iterative sampling approaches that use past explorations to inform subsequent decisions and automatically select appropriate hyperparameters to mitigate poor decisions. We will support our method empirically using some classic RL control examples.

Elizabeth Newman
Emory University
elizabeth.newman@emory.edu

Lars Ruthotto
Department of Mathematics and Computer Science
Emory University
lruthotto@emory.edu

Deepanshu Verma
George Mason University
dverma4@emory.edu

Bas Peters
Emory University
bas.peters@emory.edu

MS157

Propagation of Uncertainty in Operator Networks for Partial Differential Equations

In this talk, we present a neural network training procedure for solving time-dependent partial differential equations with an automated form of uncertainty quantification. The light-weight uncertainty extension requires minimal changes to the underlying network architecture, and the resulting models are capable of providing real-time predictions along with simultaneous uncertainty estimates to help identify potential inaccuracies in the network predictions. We provide an analysis of several architecture/training variations for propagating uncertainty forward in the context of time-dependent problems, and assess the quality of the uncertainty estimates based on the empirical distributions of network errors observed on both in-distribution and out-of-distribution validation data.

Nick Winovich
Purdue University
nwinovi@sandia.gov

Bart G. van Bloeman Waanders
Sandia National Laboratories
Optimization and Uncertainty Quantification Department
bartv@sandia.gov

MS158

Extreme Event Quantification for Fluid Systems: Rogue Waves and Turbulence

A central problem in uncertainty quantification is how to characterize the impact that our incomplete knowledge about models has on the predictions we make from them. It naturally lends itself to a probabilistic formulation, by making the unknown model parameters random with given statistics. This approach can be used in concert with tools from large deviation theory (LDT) and optimal control to estimate the probability that some observables in a dynamical system go above a large threshold after some time, given the prior statistical information about the system's parameters and its initial conditions. We use it to quantify

the likelihood of extreme surface elevation events for deep sea waves, so-called rogue waves, and compare the results to experimental measurements. We then explore how this procedure generalizes to strongly coupled stochastic partial differential equations as encountered in fluid dynamics.

Tobias Grafke
Warwick Mathematics Institute
University of Warwick
T.Grafke@warwick.ac.uk

MS158

Adaptive Importance Sampling for Efficient Stochastic Root Finding and Quantile Estimation

In solving simulation-based stochastic root-finding or optimization problems that involve rare events, such as in extreme quantile estimation, running crude Monte Carlo can be prohibitively inefficient. To address this issue, importance sampling can be employed to drive down the sampling error to a desirable level. However, selecting a good importance sampler requires knowledge of the solution to the problem at hand, which is the goal to begin with and thus forms a circular challenge. We investigate the use of adaptive importance sampling to untie this circularity. Our procedure sequentially updates the importance sampler to reach the optimal sampler and the optimal solution simultaneously, and can be embedded in both sample average approximation and stochastic approximation-type algorithms. Our theoretical analysis establishes strong consistency and asymptotic normality of the resulting estimators. We also demonstrate, via a minimax perspective, the key role of using adaptivity in controlling asymptotic errors. Finally, we illustrate the effectiveness of our approach via numerical experiments.

Michael Fu
University of Maryland,
mfu@umd.edu

Shengyi He
Columbia University, U.S.
sh3972@columbia.edu

Guangxin Jiang
Harbin Institute of Technology
gxjiang@hit.edu.cn

Henry Lam
Columbia University
henry.lam@columbia.edu

MS158

Using Large Deviation Theory and Bilevel Optimization for Controlling Cascading Failures in the Power Grid

Despite cascading failures being the central cause of blackouts in power transmission systems, existing operational and planning decisions are made largely by ignoring their underlying cascade potential. This paper posits a reliability-aware AC Optimal Power Flow formulation that seeks to design a dispatch point which has a low operator-specified likelihood of triggering a cascade starting from any single component outage. By exploiting a recently developed analytical model of the probability of component failure, our Failure Probability-constrained ACOPF (FP-ACOPF) utilizes the system's expected first failure time as a smoothly tunable and interpretable signature of cascade risk. We use techniques from bilevel optimization and numerical linear algebra to efficiently formulate and solve the FP-ACOPF using off-the-shelf solvers. Extensive simulations on the IEEE 118-bus case show that, when compared to the unconstrained and N-1 security-constrained ACOPF, our probability-constrained dispatch points can significantly lower the probabilities of long severe cascades and of large demand losses, while incurring only minor increases in total generation costs.

Anirudh Subramanyam
Argonne National Laboratory
asubramanyam@anl.gov

Jacob Roth
University of Minnesota
roth0674@umn.edu

Albert Lam
Argonne National Laboratory
klam@anl.gov

Mihai Anitescu
Argonne National Laboratory
Mathematics and Computer Science Division
anitescu@mcs.anl.gov

MS158

LDT-Based Importance Sampling for Probability Estimation of Extreme Events

We propose a method for estimating tail probabilities in complex systems that depend on high-dimensional random parameters. Our approach combines ideas from large deviation theory (LDT), optimization, dimension reduction and importance sampling. Following the LDT approach, we first compute the least unlikely point in the extreme event set, which holds crucial information about the event set probability. This point and its local derivative information are used to find a low-dimensional subspace that dominates the probability. This subspace is used in the construction of the biasing probability for importance sampling.

Shanyin Tong
New York University
shanyin.tong@nyu.edu

Georg Stadler
Courant Institute for Mathematical Sciences
New York University
stadler@cims.nyu.edu

Eric Vanden-Eijnden
Courant Institute
New York University
eve2@cims.nyu.edu

MS159

Probabilistic Sensitivity with Optimal Transport

The theory of optimal transport and the use of Wasserstein distances are attracting increasing attention in statistics and machine learning. At the same time, the definition of measures of statistical association for multivariate responses is topical research subject. This work examines the construction of probabilistic sensitivity measures using the theory of optimal transport. We obtain a new family

of indicators that are global, well posed in the presence of correlations and possess the zero-independence property. Closed form expressions are derived for the family of elliptical distributions ans a connection between dependence measures based on the Wasserstein-Bures approximation and previously introduced generalized variance-based indicators. For estimation, we employ a one-sample strategy that keeps computational burden under control. We prove the asymptotic consistency of the estimators. We test estimators based on alternative algorithmic approaches developed in the machine learning literature for optimal transport problems. Findings show that consistent estimates are obtained at reasonable sample sizes and fast execution times.

Emanuele Borgonovo
Bocconi University (Milan)
Department of Decision Sciences
emanuele.borgonovo@unibocconi.it

Alssio Figalli
ETH Zurich
alessio.figalli@math.ethz.ch

Elmar Plischke
Clausthal Univeirsity of Technology
elmar.plischke@tu-clausthal.de

Giuseppe Savare
Bocconi University
guyseppe.savare@unibocconi.it

MS159

Kernel-Based Anova Decomposition and Shapley Effects-Application to Global Sensitivity Analysis

Global sensitivity analysis is the main quantitative technique for identifying the most influential input variables in a numerical model. In particular when the inputs are independent, Sobol sensitivity indices attribute a portion of the output variance to each input and all possible interactions in the model, thanks to a functional ANOVA decomposition. On the other hand, moment-independent sensitivity indices focus on the impact of inputs on the whole output distribution instead of the variance only, thus providing complementary insight on the inputs/output relationship. But they do not enjoy the nice decomposition property of Sobol indices and are consequently harder to analyze. In this talk, we introduce two moment-independent indices based on kernel-embeddings of probability distributions and show that the RKHS framework makes it possible to exhibit a kernel-based ANOVA decomposition. This is the first time such a desirable property is proved for sensitivity indices apart from Sobol ones. With dependent inputs, we also use these new sensitivity indices as building blocks to design kernel-embedding Shapley effects which generalize the traditional ones. Several estimation procedures are discussed and illustrated on test cases with various output types such as categorical variables and probability distributions. All these examples show their potential for enhancing sensitivity analysis with a kernel viewpoint.

Sébastien Da Veiga
Safran Tech, FRANCE
sebastien.da-veiga@safrangroup.com

MS159

Global Sensitivity Analysis of Models Described by Hypoelliptic Systems of Stochastic Differential Equations

Global sensitivity analysis aims to identify the parameters whose uncertainty has the largest impact on the variability of a quantity of interest (QoI). Here we consider models described by $Z = (X, Y)$ solution of the following Stochastic Differential Equation (SDE), whose coefficients depend on some uncertainty parameter $\xi = (\xi_1, \ldots, \xi_d) \in \mathbb{R}^d$

$$dX_t = Y_t dt \quad \text{and} \quad dY_t = \sigma(\xi) dW_t - (c(\xi, X_t, Y_t) Y_t + \nabla V(\xi, X_t)) dt.$$

Under some conditions on the damping coefficient c and the potential V the process is hypoelliptic and is ergodic with a unique invariant probability measure $\mu(dxdy, \xi) = p(x, y, \xi) dxdy$. We aim at studying the influence of the ξ_i's on QoI defined from the density $p(x, y, \xi)$ of $\mu(dxdy, \xi)$. This density solves the stationary parametrized Fokker-Planck Partial Differential Equation (FKPDE) involving the adjoint of the generator of Z, with the constraint $\int p(x, y, \xi) dxdy = 1$. Note that this FKPDE is hypoelliptic. Sensitivity analysis for models driven by SDEs were presented, e.g., in [Le Matre and Knio 2015, Etor, Prieur et al. 2020]. The main challenges of the present work are to find a numerical method for solving the hypoelliptic parametrized FKPDE (for this point we are inspired by [Langtangen 1991]), and the specific nature of the quantities of interest for sensitivity analysis (that we handle using indices proposed in [Da Veiga 2021]).

Pierre Etoré
Université Grenoble Alpes, France
pierre.etore@univ-grenoble-alpes.fr

Clémentine Prieur
Grenoble Alpes University
Jean Kunzmann Lab, INRIA project/team AIRSEA
clementine.prieur@univ-grenoble-alpes.fr

Joel Andrepont
Université Grenoble Alpes
joel.andrepont@grenoble-inp.org

MS159

Information Density in Global Sensitivity Analysis

In this work, we discuss the notation and graphical representation of information density in simulation experiments. The proposed tool complements uncertainty quantification revealing the region of the support of an input where such input becomes important. We formulate information density in such a way that the definition remains well-posed for any sensitivity measure defined as the expected separation between a marginal and a conditional distribution. We discuss a one-sample estimation strategy that keeps the computational burden under control for individuals as well as joint analysis. The method is applied to the study of an epidemic model developed for risk management within the COVID-19 pandemic.

Emanuele Borgonovo
Bocconi University (Milan)
Department of Decision Sciences
emanuele.borgonovo@unibocconi.it

Gordon B. Hazen
Northwestern University
northwesterngbh305@northwestern.edu

Xuefei Lu

SKEMA Business School, Université Côte dAzur, Paris, France
xuefei.lu@skema.edu

Elmar Plischke
Clausthal Univeirsity of Technology
elmar.plischke@tu-clausthal.de

MS160

Bayesian Additive Regression Trees for Stochastic Data-Driven RANS Turbulence Modelling

Turbulent flows are commonly encountered in engineering, where the only computationally tractable paradigm is Reynolds-averaged Navier-Stokes (RANS). RANS requires a closure to model the effect of turbulence on the mean-flow, many of which perform poorly in common situations. Recent data-driven derivations of RANS closures, use LES data and machine-learning tools. These techniques allow practitioners to generate custom closures for flows, provided reference data is available. However in these problems there are an extremely large number of parameters being estimated, and relatively thin data. Therefore there is very likely a significant region of model-space that contains viable models under the data. In this work, we explore this space using Bayesian statistics. We consider a class of closure models which predict the Reynolds anisotropy tensor based on Bayesian Additive Regression Trees (BARTs). We obtain a posterior probability density over models, and therefore also meaningful model-uncertainty estimates on RANS predictions of the mean-flow. We study a variety of simple flows: square- and rectangular-ducts, backward-facing steps and periodic-hills. The variance predictions of BART are compared with (infinitesimal)-Jackknife variance estimates of existing tensor-basis random forest models, and the two methods broadly agree. The investigation demonstrates that the variance of data-driven closures is extremely large, and should not be neglected in future studies.

Richard P. Dwight
Delft University of Technology
r.p.dwight@tudelft.nl

MS160

Half-Spectral Covariance Models for Nonstationary Spatio-Temporal Flows

Gaussian process (GP) models are ubiquitous in the study of dependent processes. Particularly in the space-time setting, however, it is challenging to specialize parametric models for covariance structure to specific data and problem settings. As a result, many practitioners fall back to very generic covariance functions or statistically suboptimal approximations such as low-rank covariance matrices or separable covariance functions. These problems are further exacerbated in the case of nonstationarity, which will hold for most real-world processes. In this work, we introduce a class of covariance functions that are easily specialized to accommodate a broad range of complex parametric forms while maintaining validity and expressiveness. We demonstrate the value of such functions with an application to very high-frequency Doppler LIDAR flow velocity measurements in the low atmosphere, directly modeling and obtaining uncertainty estimates for physical quantities such as the height of the atmospheric boundary layer.

Christopher Geoga, Michael Stein
Rutgers University
christopher.geoga@rutgers.edu, ms2870@stat.rutgers.edu

MS160

PDE Modeling of a Rising Bubble for Efficient Uncertainty Quantification Computations

Uncertainty quantification computations may be prohibitively expensive. Sampling a range of values of a specific parameter with an associated uncertainty leads to solving augmented linear systems. Such systems can be solved more efficiently by batching the operations and invoking block algebra algorithms. The underlying PDE model used may yield, upon discretization, parameter dependency either for matrices or right-hand sides. From a computational efficiency perspective, it may be convenient to recast the differential equation model to achieve the highest computational performance. We will present and analyze several options for computing the uncertainty of surface tension for a rising bubble in Stokes flow. The Stokes flow assumption, albeit restrictive, is valid and widely used in micro-fluidics where the dimension of the problem renders the Reynolds number infinitesimally small. The goal of this work is to put forth the best scenarios for more accurate surface tension computations and quantifying measurement uncertainties.

Oana Marin
Argonne National Laboratory
oanam@anl.gov

MS160

Inference Design for the Uncertainty Quantification of Extreme-Scale Fluid Dynamics Simulations

The quantification of uncertainties in modern computational fluid dynamics solvers poses significant computational challenges with individual simulations costing 100s of thousands of CPU or GPU hours. Approaching such large simulation workflows with Bayesian Uncertainty Quantification, Bayesian inference at its core, hence requires the commitment of inordinate computational resources to the inference routine. To enable such large routines we hence have to exploit inherent model hierarchies to the fullest extent possible. To accelerate the sampling we build on advances in fields adjacent to uncertainty quantification, such as machine-learning based design, sequential decision making, and reinforcement learning to pose our problem as the design of a Multifidelity inference routine with intelligent sampling agents at its core. The intuition here is that the sampling agents learn the task- and problem-structure in their attention-based policy networks, which can then later be fine-tuned to similar inference problems. In this work we present a machine-learning based design approach to Bayesian Uncertainty Quantification inference routines, which is based on a graph network representation of the inference routine in combination with a Transformer-based placement network to exploit the available model hierarchies. The excessive training costs of the learned inference routine are amortized across later inference studies.

Ludger Paehler, Nikolaus Adams
Technical University of Munich
ludger.paehler@tum.de, nikolaus.adams@tum.de

MS161

Variance-Based Sensitivity Analysis for History

Matching

Variance-Based Sensitivity Analysis is a method used to apportion the sensitivity of the output of a numerical model to small changes in its inputs individually and together. This allows one to see which input variables have the most influence on the variability of the output. This has its uses in history matching (i.e. calibrating the model towards a set of output values by ruling out regions on input space that don't evaluate to those output values) as knowing which variables have the most influence means we focus our search on those variables. History matching works by constructing an emulator from a set of design points sampled within the space that is Not Ruled Out Yet (NROY). Traditionally, sensitivity analysis is carried out before the calibration but would be more efficient to conduct sensitivity analysis at each wave. We explore how sensitivity analysis can be conducted in these regions of NROY space. We know very little about their shape. We sample in parts of space with both a low implausibility measure and low uncertainty as we know the emulator has more confidence in those regions. We use these samples to conduct sensitivity analysis which informs us of the most and least influential variables. We use this knowledge to reconfigure the model for the emulator by treating variables with negligible effect as noise. This reduces the dimensionality of the emulator and thus makes history matching easier to interpret.

Michael Dunne
University of Exeter
md624@exeter.ac.uk

Peter Challenor
College of Engineering, Mathematics and Physical Sciences
University of Exeter
p.g.challenor@exeter.ac.uk

MS161

Efficient Calibration for Spatio-Temporal Models Using Basis Methods

Calibration of expensive computer models can be approached via history matching. Many such models (e.g., in climate, engineering) have high-dimensional spatial and/or temporal outputs, all of which we may wish to predict for unseen regions of parameter space, and compare to real-world observations, rather than only considering summaries of the output. We demonstrate how to history match high-dimensional output fields efficiently and effectively, with computational savings at both the emulation and history matching stages, among other benefits given by the basis approach, such as more physically-coherent predictions. We consider its application to real-world examples, including for a non-connected target space and for calibrating an environmental model, and consider how exploring model discrepancy (missing processes in the model) fits into the framework.

James Salter
University of Exeter
j.m.salter@exeter.ac.uk

Danny Williamson
Durham University
d.williamson@exeter.ac.uk

MS161

Approximate History Matching of Paleo Ice Sheet Evolution: Issues, Lessons, and Outstanding Questions for the UQ Community

This presentation offers the perspective of a modeller and lessons learned from two decades of trying to meaningfully assess uncertainties in paleo ice sheet modelling on glacial cycle (one hundred thousand year) time scales. As ice sheet evolution is highly sensitive to the associated climate, the potential phase space is large. Simulator parameter vector dimensions are order 50. Depending on the climate representation, computationally-achievable ensemble sizes with the full simulator can range from a few hundred to over 10,000. As I'll show, the latter enables adequate emulation by Bayesian Artificial Neural Networks (BANNs) for most critical scalar and some critical space-time dependent quantities. A key feature is that the required product, the complete space-time evolution of the ice sheet and/or associated climate, cannot (to date) be adequately emulated and must come from the full simulator. I will outline the physical and data context of paleo ice sheet modelling and my current approach (and ongoing challenges) to simulator calibration and state-space estimation. I will share some communication strategies gleaned from discussions with colleagues and from teaching graduate students in the modelling field about uncertainty quantification. I will also raise a number of methodological research questions, whose answers would significantly ease the adoption and advancement of meaningful uncertainty quantification in earth system modelling contexts.

Lev Tarasov
Memorial University of Newfoundland
lev@mun.ca

MS162

Data-Driven Modeling, Learning, Prediction, and Optimization for Decision Making under Uncertainty: Application to COVID-19

We propose an end-to-end computational framework for data-driven learning and decision-making under uncertainty. The framework consists of data-driven mathematical modeling, statistical learning, simulation-based prediction, and stochastic optimization, all under various inevitable uncertainties due to imperfect or unknown information, e.g., data noise, model inadequacy, parameter uncertainty, stochastic environment. In particular, to address the challenge of the curse of dimensionality in the learning by Bayesian inference and optimization with chance constraint, often arising from heterogeneity in data, inference parameters, and optimization variables, and to facilitate a seamless process from data to decision making, we propose to use a projected variational inference method in a combination of the sample average approximation for stochastic optimization. We apply the framework for optimal mitigation decision-making given corrupted data to fight against COVID-19, accounting for the strong heterogeneity of severity and transmission inside and outside of long-term care facilities.

Peng Chen
University of Texas in Austin
peng@oden.utexas.edu

Keyi Wu

University of Texas Austin
keyiwu@math.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

MS162

Graph Convolutional Neural Networks for Microstructure Homogenization with Quantified Uncertainty

Data-driven models for materials with complex microstructure have surged in popularity in recent years due to their ability to capture the structure-response relationship directly. Deep neural networks that act directly on the structure attain great accuracy in predicting stress-strain relations of the homogenized material. This accuracy comes at the cost of high model complexity and expense of training, which makes uncertainty quantification increasingly difficult. Gold standard methods like Hamiltonian Monte Carlo perform slowly due to the high dimensional parameter space, while mean field variational inference can overpredict model uncertainty due to missing correlations in the parameters. In this work we explore the use of a novel technique called Stein Variational Gradient Descent, which can effectively represent posterior distributions with a limited number of samples. We demonstrate this approach on a graph convolutional neural network for predicting the stiffness of polycrystalline structures and compare the computational efficiency and accuracy of the predictive distributions against a Hamiltonian Monte Carlo approach. SNL is managed and operated by NTESS under DOE NNSA contract DE-NA0003525

Ari Frankel, Cosmin Safta
Sandia National Laboratories
alfrank@sandia.gov, csafta@sandia.gov

Reese Jones
Co-author
rjones@sandia.gov

MS162

Bayesian Parameter Estimation from Sparse and Noisy Measurement Data in Systems Biology

We propose a Bayesian parameter estimation framework to facilitate model calibration and uncertainty analysis for predictive models of biological systems. In general, the available experimental data for these systems are sparse and noisy and introduce uncertainties into the model. Despite this, systems biologists often overlook rigorous uncertainty quantification. We present a new modeling paradigm that bridges the gap between standard practices in systems biology and more thorough uncertainty quantification. Our proposed method accounts for uncertainties in the data by adapting an approximate marginal Markov chain Monte Carlo (MCMC) method for Bayesian parameter estimation. This method enables us to characterize the parameter uncertainty and analyze the distribution of possible model outputs. We use this method to study the effects of experimental assumptions on parameter inference for a model of a well-known biological system; the mitogen-activated protein kinase signaling pathway. Our work demonstrates uncertainty quantification in systems biology and outlines the considerations and necessary steps to apply this method to future modeling efforts.

Nate Linden
University of California, San Diego
nlinden@ucsd.edu

Padmini Rangamani
Department of Mechanical and Aerospace Engineering
University of California San Diego
padmini.rangamani@eng.ucsd.edu

Boris Kramer
University of California San Diego
bmkramer@ucsd.edu

MS162

Post-Optimality Sensitivity of Model-Form Error to Support Decision-Making with Reduced Order Models

Model form error is an important source of uncertainty when high-fidelity models are represented with surrogates or reduced order models. In this work, we develop post-optimality sensitivities of model-form error in the context of an overarching decision goals (inversion, control, design). This reveals which errors are the most influential to the optimization solution and can guide training and use of the surrogates in decision-informed manner. A computational bottleneck is addressed by exploiting Kronecker product structure in the model error operator. The approach is demonstrated on non-trivial PDE-constrained optimization problems.

Bart G. Van Bloemen Waanders, Joseph Lee Hart
Sandia National Laboratories
bartv@sandia.gov, joshart@sandia.gov

MS163

Computer Using Deep Gaussian Processes to Emulate Binary Black-Hole Mergers

Simulator models are used to explore physical systems in many branches of science and engineering. Computer model emulation is usually done when the simulator is expensive to run and a surrogate model, with uncertainty, is desired. More and more, computer models are fast and a large suite of simulator runs can be made available, but the code is not readily accessible to scientists. In these cases, an emulator is also used to stand in for the computer model. This work was motivated by a simulator for the chirp mass of binary black hole mergers where no output is observed for large portions of the input space and more that one million simulator evaluations are available. This poses a problem insofar as the usual emulation approaches do not accommodate the discontinuity when observing no chirp mass. In these cases, we propose using deep Gaussian process (GP) models. We explore the impact of the model choices when setting up a deep GP on posterior inference and apply the proposed approach to the real application.

Derek Bingham, Faezeh Yazdi
Simon Fraser University
derek_bingham@sfu.ca, fyazdi@sfu.ca

Danny Williamson
Durham University
d.williamson@exeter.ac.uk

Ilya Mandel

Monash University
ilya.mandel@monash.edu

MS163

BdryGP: a New Gaussian Process Model for Incorporating Boundary Information

Gaussian processes (GPs) are widely used as surrogate models for emulating computer code, which simulate complex physical phenomena. In many problems, additional boundary information (i.e., the behavior of the phenomena along input boundaries) is known beforehand, either from governing physics or scientific knowledge. While there has been recent work on incorporating boundary information within GPs, such models do not provide theoretical insights on improved convergence rates.To this end, we propose a new GP model, called BdryGP, for incorporating boundary information. We show that BdryGP not only has improved convergence rates over existing GP models (which do not incorporate boundaries), but is also more resistant to the "curse-of-dimensionality' in nonparametric regression. Our proofs make use of a novel connection between GP interpolation and finite-element modeling.

Simon Mak
Duke University
sm769@duke.edu

Liang Ding
Hong Kong University of Science and Technology
*

Jeff Wu
Georgia Institute of Technology
jeffwu@isye.gatech.edu

MS163

Linearly Constrained Gaussian Processes

In this talk we want to show that the combined use of data-driven modelling and existing scientific knowledge can be quite rewarding. We briefly illustrate this using concrete examples from physics, including modelling the ambient magnetic field, neutron diffraction experiments aiming to reconstruct the strain field, and computed tomographic (CT) reconstruction. These are all concrete examples where physics provide us with linear operator constraints that needs to be fulfilled or alternatively measurements constituted by line integrals. The reason for the usefulness of the Gaussian process (GP) is that it offers a probabilistic and non-parametric model of nonlinear functions. When these properties are combined with basic existing scientific knowledge of the phenomenon under study we have a useful mathematical tool capable of fusing existing knowledge with new measured data. We will show how the GP can be adapted so that it obeys linear operator constraints (including ODEs, PDEs and integrals), motivated by the specific examples above. Towards the end we will also show how this work can be extended to one specific instance of non-linearly constrained GPs, where the constraints take the form of a sum over the outputs. These developments open up for the use of basic scientific knowledge within one of our classic machine learning models.

Niklas Wahlström, Carl Jidling, Philipp Pilar
Uppsala University
niklas.wahlstrom@it.uu.se, carl.jidling@it.uu.se, philipp.pilar@it.uu.se

Thomas Schön
Uppsala University
Department of Information Technology
thomas.schon@it.uu.se

Adrian Wills
University of Newcastle
adrian.wills@newcastle.edu.au

MS163

Gaussian Process Subspace Regression for Dimension Reduction of Computational Models

Subspace-valued functions arise in a wide range of problems including parametric reduced order modeling (PROM). In PROM, each parameter point can be associated with a subspace, which is used for Petrov-Galerkin projections of system matrices. Previous efforts to approximate such functions use interpolations on manifolds, which can be inaccurate and slow. We propose a novel Bayesian nonparametric model for subspace prediction: the Gaussian Process Subspace (GPS) model. This method is extrinsic and intrinsic at the same time: with multivariate Gaussian distributions on the Euclidean space, it induces a joint probability model on the Grassmann manifold, the set of fixed-dimensional subspaces. The GPS adopts a simple yet general correlation structure, and a principled approach for model selection. Its predictive distribution admits an analytical form, which allows for efficient subspace prediction over the parameter space. For PROM, the GPS provides a probabilistic prediction at a new parameter point that retains the accuracy of local reduced models, at a computational complexity that does not depend on system dimension, and thus is suitable for online computation. We give four numerical examples to compare our method to subspace interpolation, as well as two methods that interpolate local reduced models. Overall, GPS is the most data efficient, more computationally efficient than subspace interpolation, and gives smooth predictions with uncertainty quantification.

Ruda Zhang, Simon Mak, David Dunson
Duke University
ruda.zhang@duke.edu, sm769@duke.edu, dunson@duke.edu

MS164

Generating Random Field Using GAN for Stochastic Simulations

Many stochastic simulations in material modeling involve a random field that represents the property of the material such as Youngs modulus. In existing studies, a conventional way of constructing such random fields is to use a Gaussian random field which typically assumes a homogeneous kernel function. A significant drawback of this approach is that it does not exploit available data, e.g., a few images from experiments or experts' knowledge such as known features of the material. We propose to use a variant of generative adversarial network (GAN) to generate such random fields based on available data and known (or desired) features. Unlike the conventional GAN which requires a large amount of training data, this approach can generate new samples using limited data. It can also generate 3D material structure based on 2D images scanned at multiple layers.

Xiu Yang
Lehigh University

xiy518@lehigh.edu

MS164

Multiscale Coupling with Operator-Learning Neural Network

Multiscale modeling is an effective approach for studying mechanics of solids when involving microsctructure and strain localization which makes the computational cost of high-fidelity, first-principle modeling in the global scale prohibitively high. In multiscale modeling, microscopic features are described by expensive, fine models and refined discretization, whereas the system behavior in the region far from the microstructure is emulated with a cheaper coarse model. In this work, we develop a novel multiscale framework by coupling a powerful operator-learning neural network, Deep Operator Network (DeepONet), as the fine model with the finite-element method (FEM) as the coarse model for solving static and dynamic problems. The framework includes a offline training part, where a DeepONet is trained to learn the hidden mechanism of the microstructure and serves as a surrogate model for this fine model. With the trained DeepONet, multiscale problems are solved in an online approach by coupling the FEM as the coarse model and the DeepONet as the surrogate fine model. Our method significantly reduces the computational cost of multiscale problems thanks to the fast prediction of trained DeepONet, which bypasses the needs of explicitly involving the governing mechanisms upon the online solution of multiscale problems. We demonstrate by examples that our framework is flexible with the coupling scheme: both overlapping and non-overlapping schemes are compatible with our method.

Minglang Yin, Enrui Zhang
Brown University
minglang_yin@brown.edu, enrui_zhang@brown.edu

Yue Yu
Department of Mathematics, Lehigh University
yuy214@lehigh.edu

George E. Karniadakis
Brown University
Division of Applied Mathematics
george_karniadakis@brown.edu

MS164

Pragmatic Stochastic Fractional PDEs and UQ

Holistic, pragmatic and predictive mathematical modeling is of great practical and theoretical importance. We present a series of data-infused and generalized mathematical modeling, where the underlying physics and mathematical language meet each other at the expense of pragmatic usability of data, in which data-errors, model-errors, and simulations errors are regarded in an integrative fashion parsimoniously

Mohsen Zayernouri
Department of Computational Mathematics, Science, & Eng.
Michigan State University
zayern@msu.edu

MS164

Meta-Learning for Heterogeneous Materials: A Provable Nonlocal Operator Regression Approach

In this talk, we propose a meta-learnt approach for learning mappings between function spaces (operators), Meta-NOR, based on the nonlocal operator regression, to efficiently provide accurate model surrogates for in new and unknown PDE-learning tasks. The proposed provably sample-efficient meta-learning algorithm uses a multi-task nonlocal operator regression model in the kernel space, which consists of two phases: (1) learning a common nonlocal kernel representation from existing tasks; and (2) transferring the learned knowledge to rapidly learn surrogate models for new tasks with different governing PDEs (such as with different constitutive laws or material parameters), where the governing PDEs could be possibly unknown and only a few test samples are provided. Under the linear kernel regression setting, a provable optimization-based approach is provided, with theoretically guaranteed transfer-learning error bounds. We apply the proposed technique to model the wave propagation within 1D meta-materials with both periodic and random microstructures, showing that the meta-learnt kernel representation would greatly improve the sampling efficiency in new and unseen microstructures, compared to existing baselines.

Lu Zhang
Lehigh University
luz319@lehigh.edu

MS165

Learning Data Shifts in Spectroscopy via Structured Normalizing Flows

Since landing at Gale Crater in 2012, the ChemCam laser-induced breakdown spectroscopy (LIBS) instrument onboard the Mars rover Curiosity has obtained spectral measurements of thousands of rock and soil analysis targets. The compositions of the major elements may be predicted using models trained on samples with known compositions measured by a laboratory instrument under Mars-like atmospheric conditions. However, laboratory measurements and rover measurements on identical sets of *calibration targets* still display some notable differences, prompting development of an *Earth-Mars* correction. Currently, this correction is computed using the ratios of Mars and laboratory spectra on a few of the calibration targets, but this calculation is sensitive to the small amount of data and may not generalize well on new samples. In this work, we explore the Earth-Mars spectral difference in a probabilistic framework by investigating how the spectral probability densities differ on Mars and Earth. On a structured latent space obtained by standard dimension reduction methods, we construct a *composition operator* with Normalizing Flows that learns how the latent spaces for Mars spectral data and Earth spectral data differ in distribution. We arrive at a structured approach for learning spectral data shifts between Earth and Mars that leverages all of the available measurements and provides a richer and potentially more robust and interpretable Earth-Mars correction.

Natalie Klein
Los Alamos National Lab
neklein@lanl.gov

Nishant Panda
Colorado State University

nishpan@lanl.gov

MS165

A Bayesian Deep Learning Approach to Near-Term Climate Prediction

The unparalleled availability of high-fidelity data from numerical simulations and experimental measurements enables the climate community to tackle complex classification and regression tasks using deep neural networks. It is critical to represent uncertainty since these deeply structured models are ultimately utilized to make decisions. The current paper presents a Bayesian deep learning-based approach for estimating and forecasting sea surface temperature, as well as for obtaining practical uncertainty estimates in model projections. We investigated several deterministic machine learning and deep learning techniques. The findings demonstrate that Bayesian deep learning models outperform deterministic deep learning models in terms of predictive ability.

Xihaier Luo
Brookhaven National Lab
xluo@bnl.gov

Balu Nadiga
Los Alamos
balu@lanl.gov

Yihui Ren
Brookhaven National Laboratory
yren@bnl.gov

Ji Hwan Park
Brookhaven National Lab
parkj@bnl.gov

Wei Xu
Brookhaven National Laboratory
xuw@bnl.gov

Shinjae Yoo
Brookhaven National Lab
sjyoo@bnl.gov

MS165

An Uncertainty Quantification Enabled Semi-Supervised Paired Neural Network for Few-Shot Classification

In this work we introduce Bootstrapped Paired Neural Networks (BPNN), a semi-supervised, low-shot model with uncertainty quantification (UQ). When collecting imaging data, there is often large amounts of data which can be costly to label, so we would like to supplement labeled data with the vast unlabeled data (often $> 90\%$ of the data) available. Exponential average adversarial training uses unlabeled data with a combination of mean teacher semi-supervised learning and adversarial examples as in virtual adversarial training. Often, it is difficult and costly to obtain the sample size necessary to train a deep learning model to a new class or target; using paired neural networks (PNN), our model is generalized to low- and no-shot learning by learning an embedding space for which the underlying data population lives, this way additional labeled data may not be necessary to detect for targets or classes which weren't originally trained on. Finally, by bootstrapping the PNN, the BPNN model gives an uncertainty score on predicted classifications with minimal statistical distributional assumptions. The model's ability to provide uncertainty for its own predictions can be used to reduce false alarms rates, provide explainability to black box models, and help design efficient future data collection campaigns. Although models exist to contain two of these three qualities, to our knowledge no model contains all three: semi-supervised, low-shot, and uncertainty quantification.

Kathryn Gray, Daniel Ries
Sandia National Laboratories
kgray1@sandia.gov, dries@sandia.gov

MS166

Robust Convolutional Autoencoder in Learning Missing Physics Between Multifidelity Models

Heat conduction simulation plays an important role in predicting solidification dynamics that determine microstructure properties of additively manufactured products. High-fidelity heat conduction simulators, e.g., `Truchas` [**?**], can capture many desired physics, but they are often too computationally expensive to simulate the solidification dynamics of a printing process at a practical temporospatial scale. In this work, we propose to formulate a model approximation problem which can be solved by leveraging blackbox optimization methods. The analytical model with calibrated parameters enables a cost effective approximation to the ground truth data produced by the high-fidelity simulation model. To improve the calibration performance, we leverage the analytical model with multiple heat sources to increase the model complexity and flexibility enabling a better fit the high-fidelity model. Two blackbox optimization methods, i.e., Bayesian optimization and directional Gaussian smoothing methods, are employed to address the challenge that the gradient of the loss function is inaccessible during optimization. Our results show that the single-beam analytical model has limitations in fitting the high-fidelity model but the multi-beam analytical model provides satisfactory approximations to the high-fidelity temperature and melt pool fields.

Sirui Bi
Oak Ridge National Laboratory
bis1@ornl.gov

MS167

Optimal Regularization of Inverse Problems Using Neural Networks

Emerging fields such as data analytics, machine learning, and uncertainty quantification heavily rely on efficient computational methods for solving inverse problems. With growing model complexities and ever-increasing data volumes, state-of-the-art inference method exceeded their limits of applicability and novel methods are urgently needed. In recent years new approaches for optimal experimental design for inverse problems have been investigated. Optimal experimental design for inverse problems require well suited prior to obtaining meaningful solutions. In this talk, we will discuss and utilize new optimal experimental design frameworks for obtaining optimal priors, where priors are designed for their experiments. We also discuss alternative neural network learning techniques. In various numerical experiments, such as medical tomography, we illustrate the advantages and limitations of our methods.

Matthias Chung

Department of Mathematics
Virginia Tech
mcchung@vt.edu

Afkham Babak
Technical University of Denmark
bmaaf@dtu.dk

Julianne Chung
Department of Mathematics
Virginia Tech
jmchung@vt.edu

MS167

Goal-Oriented Optimal Experimental Design for Nonlinear Models using MCMC

Optimal experimental design (OED) provides a systematic approach to quantify and maximize the value of experimental data. Under a Bayesian approach, conventional OED maximizes the expected information gain (EIG) on model parameters. However, we are often interested not the parameters, but predictive quantities of interest (QoIs) that depend on the parameters in a nonlinear manner. We present a framework of goal-oriented optimal OED (GOOED) to find the experimental design that provides the greatest EIG on the QoIs. Not only does GOOED avoid unnecessary shrinkage of parameter posterior in regions that are unimportant to the QoIs, it also focuses the computation on the posterior pushforward distribution that is generally lower dimensional than the parameter space. In this work, we use a double-nested Monte Carlo algorithm. In the outer loop, observations are first generated from the marginal likelihood. For each observation, Markov chain Monte Carlo then draws samples from the corresponding parameter posterior, which are then evaluated in the prediction model to obtain QoI samples from the posterior pushforward. Kernel density estimation is then used to obtain the probability density of the posterior pushforward, in order to compute the Kullback-Leibler divergence information gain. We demonstrate the effectiveness of nonlinear GOOED on a number of applications, and illustrate the difference between optimal designs from conventional OED versus GOOED.

Shijie Zhong
Shanghai Jiaotong University
18-zsj@sjtu.edu.cn

Wanggang Shen, Xun Huan
University of Michigan
wgshen@umich.edu, xhuan@umich.edu

MS167

Optimal Experimental Design Using Data About Bifurcation Points

The use of mathematical optimization and computerized simulation to gain knowledge of a dynamical process is a challenging task of modern sciences. To perform a simulation and optimization of a dynamical process, a mathematical model is required which is assumed to be a sufficient explanation of the underlying process. Usually the model contains unknown natural quantities and one of the principal challenges is to adapt the model to the process behavior using measurement data. In special case of chemical or bio-chemical dynamical models there are often several possibilities to measure the process behavior. If underlying process shows temporal changes, it is obvious that an observation at an additional point in time delivers additional information. If the underlying process possesses bi-stability property, valuable information can be gained using measurement data at the so called bifurcation points. Such bifurcation points can be located experimentally studying hysteretic curves. In this talk we discuss novel numerical methods for parameter estimation for bi-stable dynamic processes and design of optimal experiments to support parameter estimation using measurement data from bifurcation points.

Ekaterina Kostina
Heidelberg University, Germany
ekaterina.kostina@iwr.uni-heidelberg.de

Juergen Gutekunst
IWR, University of Heidelberg
juergen.gutekunst@iwr.uni-heidelberg.de

Hridya Vinod Varma
IWR, Heidelberg University
hridya.varma@iwr.uni-heidelberg.de

MS168

Operator Inference for Non-Intrusive Model Reduction with Nonlinear Manifolds

Nearly all model-reduction techniques project the governing equations onto a linear subspace of the original state space. Unfortunately, restricting the state to evolve in a linear subspace does not always produce low-dimensional reduced-order models (ROMs). To address this, we propose a novel framework for projecting dynamical systems onto nonlinear manifolds using minimum-residual formulations. We introduce a nonlinear mapping characterized by a polynomial structure for our purpose of reduction. Compared to learning approaches based on convolutional autoencoders from deep learning, the proposed framework does not require large amounts of training data or the calibration of hyperparameters. Instead, the nonlinear mapping can be constructed in a non-intrusive fashion using regression methods and is driven by physics-based training data. The mapping can thus be inferred efficiently using standard numerical solvers. The nonlinear manifolds are then applied to the operator inference approach for projection-based model reduction of complex PDE models. The potential of the proposed approach is investigated in a numerical study, where several variations of the approach are compared on different examples, giving a clear indication of where the proposed approach is applicable.

Rudy Geelen
Oden Institute for Computational Engineering and Sciences
The University of Texas at Austin
rudy.geelen@austin.utexas.edu

MS168

Reduced Order Models for Risk-Averse Optimization

This talk discusses the integration of reduced order models (ROMs) into risk-averse optimization governed by partial differential equation constraints with uncertain parameters. Risk-averse optimization formulations use a risk measure to penalize high-cost, rare events, but are computationally difficult to solve as many risk measures such as the Conditional Value-at-Risk (CVaR) or buffered probability

of failure are non-smooth and require sampling in the tail of a complex, unknown probability distribution that depends on the optimization variables. ROMs are used to approximate the original quantity of interest, as well as to identify samples in the tail of the unknown probability distribution to generate biasing distributions for importance sampling. In the optimization context it is important that key quantities that need to be approximated for CVaR objective function evaluations are also the key quantities arising in gradient computations. We adapt optimization approaches that can incorporate inexact function and gradient evaluations to adapt the accuracy of the ROMs and sampling throughout the optimization iterations.

Matthias Heinkenschloss
Department of Computational and Applied Mathematics
Rice University
heinken@rice.edu

Mae Markowski
Rice University
mae.markowski@rice.edu

MS168

Transport-Based Offline/Online Approach for Sequential Bayesian Inference

In this work aims to address sequential Bayesian inference problems where the objective is to characterize the posterior distribution of some static parameters of a model each time a set of new observations is available. The approach developed relies on the extraction of conditional distributions (e.g., likelihood functions) from the *joint prior* distribution of parameters and data, via the estimation of structured (e.g., block triangular) transport maps. Once the estimation of such maps is found, closed forms of the underlying likelihood functions can be extracted in order to perform Bayesian inference. These computations are performed in the *offline* phase as the knowledge of the observations is not required. Then, in the *online* phase new maps are computed to characterize the posterior distribution defined as the product of the pre-computed surrogate likelihood functions and the prior. This, allow model-free computation in the online phase and makes Bayesian inference in real time possible. As only simulations from the model are required, this approach is also particularly well suited when facing black-box models. The performance of the method will be illustrated on different numerical examples.

Paul-Baptiste Rubio
MIT
rubiop@mit.edu

Youssef M. Marzouk
Massachusetts Institute of Technology
ymarz@mit.edu

MS168

Context-Aware Learning of Stabilizing Controllers from Data

Learning controllers from data is typically a two step process. First, a model of the system is learned from data (system identification). Then a controller is constructed based on the learned model. However, learning a model first can be expensive in terms of number of data points and training costs. In this work, we demonstrate that stabilizing controllers can be identified directly from data without learning models of the underlying systems first. In particular, directly learning controllers requires fewer data points than identifying models in certain situations. Numerical experiments with science and engineering problems show that the proposed approach learns stabilizing controllers even when the classical learn-model-then-stabilize approach fails due to too few data points.

Steffen W. R. Werner, Benjamin Peherstorfer
Courant Institute of Mathematical Sciences
New York University
steffen.werner@nyu.edu, pehersto@cims.nyu.edu

MS169

Machine Learning Models with Uncertainty Estimation for Molecular Dynamics and Sampling

Machine-learning models have emerged as a very effective strategy to sidestep time-consuming electronic-structure calculations, enabling accurate simulations of greater size, time scale, and complexity. Given the interpolative nature of these models, the reliability of predictions depends on the position in phase space, and it is crucial to obtain an estimate of the error that derives from the finite number of reference structures included during model training. When using a machine-learning potential to sample a finite-temperature ensemble, the uncertainty on individual configurations translates into an error on thermodynamic averages and leads to a loss of accuracy when the simulation enters a previously unexplored region. I will discuss how uncertainty quantification can be used, together with a baseline energy model, or a more robust but less accurate interatomic potential to obtain more resilient simulations and to support active-learning strategies, and introduce an on-the-fly reweighing scheme that makes it possible to estimate the uncertainty in thermodynamic averages extracted from long trajectories.

Michele Ceriotti
EPFL
michele.ceriotti@epfl.ch

MS169

Finding and Filling 'Holes' in ACE (Atomic Cluster Expansion) Using Hyperactive Learning (HAL)

We present HAL which aims to find holes (regions of instability) in ACE potentials by driving Molecular Dynamics (MD) simulations towards regions of high uncertainty. Using a Bayesian committee the relative force error is estimated during these biased MD simulations and is closely monitored in order to find holes in the ACE potential. By using this scheme iteratively it is shown that training databases can efficiently be assembled reducing the number of DFT calculations required to determine macroscopic properties such as alloy bulk/elastic constants and melting temperatures.

Cas Van Der Oord
University of Cambridge
casv2@eng.cam.ac.uk

MS170

Solving Nonlinear Pdes with Gaussian Processes

In this talk I present a simple, rigorous, and interpretable framework for solution of nonlinear PDEs based on the framework of Gaussian Processes. The proposed approach

provides a natural generalization of kernel methods to nonlinear PDEs; has guaranteed convergence; and inherits the state-of-the-art computational complexity of linear solvers for dense kernel matrices. I will outline our approach by focusing on an example nonlinear elliptic PDE followed by further numerical examples. I will also briefly comment on extending our approach to solving inverse problems.

Bamdad Hosseini
California Institute of Technology
bamdadh@uw.edu

MS170

Learning Orbital Dynamics of Binary Black Hole Systems from Gravitational Wave Measurements

We introduce a gravitational waveform inversion strategy that discovers mechanical models of binary black hole (BBH) systems. We show that only a single time series of (possibly noisy) waveform data is necessary to construct the equations of motion for a BBH system. Starting with a class of universal differential equations parameterized by feed-forward neural networks, our strategy involves the construction of a space of plausible mechanical models and a physics-informed constrained optimization within that space to minimize the waveform error. We apply our method to various BBH systems including extreme and comparable mass ratio systems in eccentric and non-eccentric orbits. We show the resulting differential equations apply to time durations longer than the training interval, and relativistic effects, such as perihelion precession, radiation reaction, and orbital plunge, are automatically accounted for. The methods outlined here provide a new, data-driven approach to studying the dynamics of binary black hole systems.

Brendan Keith
Brown University
brendan_keith@brown.edu

Scott Field
Mathematics Department, UMASS Dartmouth
Dartmouth, MA , USA
sfield@umassd.edu

Akshay Khadse
The University of Mississippi
akhadse@go.olemiss.edu

MS170

Rapid Emulation of Parametric PDEs via Self-Supervised Learning with Deep Operator Networks

Design and optimal control problems are among the fundamental, ubiquitous tasks we face in science and engineering. In both cases, we aim to represent and optimize an unknown (black-box) function that associates a performance/outcome to a set of controllable variables through an experiment. In cases where the experimental dynamics can be described by partial differential equations (PDEs), such problems can be mathematically translated into PDE-constrained optimization tasks, which quickly become intractable as the number of control variables and the cost of experiments increases. In this work we leverage physics-informed deep operator networks (DeepONets) – a self-supervised framework for learning the solution operator of parametric PDEs – to build fast and differentiable surrogates for rapidly solving PDE-constrained optimization problems, even in the absence of any paired input-output training data. The effectiveness of the proposed framework will be demonstrated across different applications involving continuous functions as control or design variables, including time-dependent optimal control of heat transfer, and drag minimization of obstacles in Stokes flow. In all cases, we observe that DeepONets can minimize infinite dimensional cost functionals in a matter of seconds, yielding a significant speed up compared to traditional PDE-constrained optimization approaches that employ adjoint solvers to optimize over finite-dimensional state-spaces.

Sifan Wang, Paris Perdikaris
University of Pennsylvania
sifanw@sas.upenn.edu, pgp@seas.upenn.edu

MS170

Neural Network Approximation of Analytic Functions with Respect to Gaussian Measures

We prove expression rates for the approximation of analytic functions on $\mathbb{R}^d$ with ReLU neural networks. The error is measured in the L^2-norm with respect to the Gaussian product probability measure. In the finite dimensional case, exponential convergence rates are shown. In the infinite dimensional case, under suitable smoothness and sparsity assumptions, we prove algebraic expression rates. As an application we consider the approximation of response surfaces of elliptic PDEs with log-Gaussian random field inputs.

Jakob Zech
Heidelberg University
jakob.zech@uni-heidelberg.de

Christoph Schwab
ETH-Zurich
schwab@sam.math.ethz.ch

MS171

Active Subspaces Without Gradients: Non-Linear Ordered Design Manifolds for Deep Learning Model Reduction in Design Optimization and UQ

We will describe some of our ongoing work for the Department of Energy and National Institutes of Health on recovering low-dimensional reduced-order models for performing accelerated design optimization, robust optimization, and UQ. The key idea lies in using machine-learned generative models on a subset of optimal designs to recover low dimensional representation when access to high fidelity gradient-based solvers are not available, such that approaches like traditional Active Subspaces are not viable. We'll present some pedagogical examples on known test cases, where we can compare the spaces obtained from design manifolds vs those found from competing methods. Then we will demonstrate this on a few real-world test cases we are using in ongoing work including robust optimization applications in medical devices, aerodynamic surfaces, and heat exchangers.

Mark Fuge
Maryland University
fuge@umd.edu

MS171

Automatic Differentiation for Stellarator Opti-

mization

Stellarators are a class of toroidal fusion reactors which use geometrically complex magnetic fields and current-carrying coils to contain a high-temperature plasma in a stable magnetic equilibrium. Creating a high-performance stellarator equilibrium has required the use of computational optimization techniques. For example, the world-leading Wendelstein 7-X (W7X) stellarator experiment was completed in 2015 but designed using computational optimization in the 1990s. The W7X experiment has driven renewed interest in the stellarator approach to magnetic fusion. This interest has led to an explosion of theoretical advances into the stellarator concept and renewed computational work on stellarator optimization. In this talk, I discuss the use of automatic differentiation (AD) in stellarator optimization. I highlight my own work introducing AD to stellarator coil design, and discuss follow-on work which has used AD in the optimization of the magnetic equilibrium. Overall, I argue that the stellarator optimization community would benefit from the adoption of AD today in much the same way the machine learning community has benefited from the adoption of AD in the early 2010s.

Nick McGreivy
Princeton University
mcgreivy@princeton.edu

MS171

Bayesian Techniques for Particle Accelerator Modeling and Control

Accelerators and other large experimental facilities are complex, noisy systems that are difficult to characterize and control efficiently. Bayesian statistical modeling techniques are well suited to this task, as they minimize the number of experimental measurements needed to create robust models, by incorporating prior, but not necessarily exact, information about the target system. Furthermore, these models inherently consider noisy and/or uncertain measurements and can react to time-varying systems. Here we will describe several advanced methods for using these models in accelerator characterization and optimization. First, we describe a method for rapid, turn-key exploration of input parameter spaces using little-to-no prior information about the target system. Second, we highlight how these models can take hysteresis effects into account and create in-situ models of individual magnetic elements.

Ryan Roussel
University of Chicago
rroussel@slac.stanford.edu

Auralee Edelen
SLAC National Accelerator Laboratory
edelen@slac.stanford.edu

Kabir Dubey
University of Chicago
kabirdubey@uchicago.edu

MS172

Spatial Models with Boundary Constraints with Application to Probabilistic Numerical Methods

Spatio-temporal evolution of environmental variables, from the diffusion of a pollutant to the movement of animals, is often understood implicitly as a function of their derivatives. These models take the form of partial differential equations defined implicitly over a spatio-temporal domain with constraints applied at the boundary. Highly structured discretization uncertainty for PDEs has recently been modeled probabilistically via Bayesian error modeling. This talk will introduce a new model for discretization uncertainty for partial differential equations and describe related advances to nonparametric modeling of states on bounded domains with known constraints.

Oksana A. Chkrebtii
The Ohio State University
oksana@stat.osu.edu

MS172

Parameter Estimation and Uncertainty Quantification Using Dynamic Survival Analysis (DSA) of Epidemiological Models

This talk will introduce the notion of dynamic survival analysis (DSA). We show that solutions to ordinary/partial differential equations (ODEs/PDEs) describing the large-population limits of stochastic epidemic models can be interpreted as survival or cumulative hazard functions when analysing data on individuals sampled from the population. We refer to the individual-level survival and hazard functions derived from population-level equations as a survival dynamical system (SDS). To illustrate how population-level dynamics imply probability laws for individual-level infection and recovery times that can be used for statistical inference and uncertainty quantification, we show numerical examples based on synthetic data as well as the COVID-19 and the Ebola outbreak data.

Wasiur Rahman Khuda Bukhsh
University of Nottingham
wasiur.rahman.kb@gmail.com

MS172

A Sequential Approach to Calibration of a Computationally Intensive Model

Calibration is the determination of the best parameter settings for a simulator to predict futurewith well-quantified uncertainties. In a general calibration process, one builds a cheap emulator using data from the simulations, and then performs calibration using different approaches such as well-known Bayesian or optimization-based calibration approaches. While an emulation-based calibration inference can be considered as a one-time process, a sequential model calibration framework can result in correct and computationally efficient inference for the calibration of computationally challenging computer models. Our approach sequentially updates the model calibration depending on new surrogate-assisted points from simulations, and terminates once it reaches to certain uncertainty level. To illustrate this approach, we use surmise-a Python package that is designed to provide a surrogate model interface for calibration, uncertainty quantification, and sensitivity analysis. We test our approach and report our results on a computationally expensive low-energy physics code for coupled-channels calculations.

Ozge Surer
Northwestern University
ozgesurer2019@u.northwestern.edu

MS173

Reduced order models for Lagrangian hydrody-

namics

As a mathematical model of high-speed flow and shock wave propagation in a complex multimaterial setting, Lagrangian hydrodynamics is characterized by moving meshes, advection-dominated solutions, and moving shock fronts with sharp gradients. These challenges hinder the existing projection-based model reduction schemes from being practical. We present several variations of projection-based reduced order model techniques for Lagrangian hydrodynamics by introducing three different reduced bases for position, velocity, and energy fields, with a time-windowing approach to address the challenge imposed by the advection-dominated solutions. Lagrangian hydrodynamics is formulated as a nonlinear problem, which requires a proper hyper-reduction technique. Therefore, we present the over-sampling DEIM and SNS approaches to reduce the complexity due to the nonlinear terms. Finally, we also present both a posteriori and a priori error bounds associated with our reduced order model, and the performance comparison of the spatial and time-windowing reduced order modeling approaches in terms of accuracy and speed-up with respect to the corresponding full order model for several numerical examples, namely Sedov blast, Gresho vortices, Taylor-Green vortices, and triple-point problems.

Siu Wun Cheung, Youngsoo Choi
Lawrence Livermore National Laboratory
cheung26@llnl.gov, choi15@llnl.gov

Dylan M. Copeland
TBA2
copeland11@llnl.gov

Kevin Huynh
Lawrence Livermore National Laboratory
huynh24@llnl.gov

MS173

AI for Materials Science: Tuning Laser-Induced Graphene Production and Beyond

AI and machine learning have advanced the state of the art in many application domains. We present an application to materials science; in particular, we use surrogate models with Bayesian optimization for automated parameter tuning to optimize the fabrication of laser-induced graphene. This process allows to create microscopic conductive lines in thin layers of insulating material, enabling the development of next-generation nano-circuits. We are able to achieve improvements of up to a factor of two compared to existing approaches in the literature and to what human experts are able to achieve. Our implementation is based on the open-source mlr and mlrMBO frameworks and can be applied in many other contexts; we explore the generation of advanced materials based on the same framework in a preliminary study.

Lars Kotthoff, Patrick Johnson, Hud Wahab, Vivek Jain, Alexander Tyrrell, Sourin Dey
University of Wyoming
larsko@uwyo.edu, pjohns27@uwyo.edu, hwahab@uwyo.edu, vjain2@uwyo.edu, atyrrel1@uwyo.edu, sdey2@uwyo.edu

MS173

Autonomous Learning of Reduced-Order PDF Equations

We consider systems whose dynamics can be described by high-dimensional stochastic differential equations (SDEs). Such equations admit a distribution of solutions (partially) characterized by a single-point (in time) joint probability density function (PDF) of system states that satisfies exactly a Fokker-Planck equation (FPE). Since high-dimensional phase space usually renders the FPE computationally intractable, reduced-order PDF equations can be used to quantify uncertainty for specific quantities of interest. However, these reduced-order equations require a closure approximation. We propose a data-driven method for closing such equations that is independent of the high-dimensional phase space. It involves estimating conditional expectations from experimental data or sample trajectories. We demonstrate the methods efficacy for high-dimensional power system models in presence of cross-correlated Ornstein-Uhlenbeck noise, random parameters, and random initial conditions. Accuracy is successfully tested against Monte Carlo simulations.

Vishwas Rao
Argonne National Laboratory
vhebbur@anl.gov

Daniel A. Maldonado
Argonne National Laboratory
9700 S. Cass Avenue Lemont, IL 60439
maldonadod@anl.gov

Tyler Maltba
University of California, Berkeley
tyler_maltba@berkeley.edu

MS173

Gaussian Process Surrogate Models for Efficient Computations in Bayesian Inverse Problems

A major challenge in the application of sampling methods to large scale Bayesian inverse problems, is the high computational cost associated with solving the forward model for a given set of input parameters. To overcome this difficulty, we consider using a surrogate model that approximates the solution of the forward model at a much lower computational cost. We focus in particular on Gaussian process emulators, and discuss issues in the efficent goal-oriented construction of the emulator.

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

PP1

Disentangling Intracellular Behavior from Extracellular Data

Bacteria can utilize a biodiesel waste product, glycerol, to produce 1,3-propanediol (1,3-PDO), a common commercial solvent. To study the viability of 1,3-PDO producing bacteria, the cell membrane transport of reactant species and the in-vivo kinetics of the pathway enzymes (DhaT DhaB) need to be quantified. The kinetics of purified DhaT have been studied, but the reaction environment and kinetics in cells may differ. Additionally, the kinetics of the DhaB enzyme and cell membrane transport mechanisms of the reactants are undetermined. With published consumption and production time series data of bacteria feeding on different initial glycerol concentrations, I used Bayesian Cal-

ibration on an ODE model of the system to infer the unknown parameters. The resulting data distribution closely matched the experimental results. The resulting parameter distributions revealed that glycerol and 1,3-PDO are transported via facilitated diffusion; the maximum reaction rate of DhaT is larger in-vivo than in-vitro; and the permeability of 3-HPA, a toxic intermediate, is structurally unidentifiable under the experimental conditions.

Andre Archer, Taylor Nichols
Northwestern University
andrearcher2017@u.northwestern.edu,
tay.nichols@u.northwestern.edu

Niall M. Mangan
Dept. Eng. Sci. and Applied Mathematics
Northwestern University
niall.mangan@northwestern.edu

Danielle Tullman-Ercek
Northwestern University
ercek@northwestern.edu

PP1

Robust Initializations of Variational Inference with Gaussian Mixtures Through Global Optimization and Laplace Approximations

In variational inference (VI), mean-field approximations are often too restrictive to provide useful approximations of intractable Bayesian posteriors exhibiting multimodal and non-Gaussian behavior. High-fidelity surrogate posteriors can be obtained by considering the family of Gaussian mixtures, capable of capturing multiple modes and approximating any distribution. VI with Gaussian mixtures can suffer from an explosion in the number of parameters as the dimensionality of the underlying problem increases. Coupled with the existence of multiple local minima due to strong nonconvex trends in the loss functions associated with VI, these challenges motivate the need for robust initialization procedures to improve the performance of VI with mixture models. In this work, we propose a method for constructing a Gaussian mixture model approximation that can be used to initialize VI. The procedure begins with a global optimization stage to find a set of local maxima, which we take to approximate the mixture component centers. Around each mode, a local Gaussian approximation is constructed via the Laplace approximation. Finally, the mixture weights are determined through constrained least squares regression. The procedure is subjected to a variance based sensitivity analysis to investigate its robustness across various aspects of the posterior distributions. The proposed methodology is also analyzed as an initialization procedure for VI on a multimodal toy problem in two dimensions.

Wyatt H. Bridgman
Sandia National Laboratories
whbridg@sandia.gov

Ahmad A. Rushdi
Stanford University
rushdi@stanford.edu

Mohammad Khalil
Sandia National Laboratories
mkhalil@sandia.gov

PP1

Uncertainty Quantification in Multiscale Modeling of Markovian Type Charge Dynamics

First-principle multiscale models of charge dynamics in disordered material typically combine classical methods for structure simulations with a quantum description of the charge carriers[**?**]. Rates for transitions are determined and the dynamics modeled as a random walk in an electric network. The Markovian transition matrix is subject to uncertainties from the multi-scale procedure, and large statistical samples are required for obtaining device properties, such as charge absorption time and effective conductance. As an alternative to expensive Monte Carlo (MC) sampling, we study the dynamical properties from an analysis of the uncertainty within a Gaussian disorder model. Specifically, means and standard deviations of absorption time and effect resistance are calculated from polynomial chaos expansion (PCE)[**?**]. Sensitivity analysis is performed to identify the material regions prone to perturbation and most influential on the macroscopic observables. We find that as the PCE terms go to infinities, the effective resistance approaches the Monte Carlo sample means. From the uncertainty quantification achieved by PCE, we can make statements about the confidence level at characterizing the dynamical properties of the materials. We show numerical results from MC and PCE of the hole conductance in a polymeric composite material used in a neuromorphic device.

Zhongquan Chen
Eindhoven University of Technology
Eindhoven University of Technology
z.chen3@tue.nl

Bjoern Baumeier
Department of Mathematics and Computer Science, TU/Eindhoven
Institute for Complex Molecular Systems, TU/Eindhoven
b.baumeier@tue.nl

PP1

A Variance Deconvolution Approach to Sampling Uncertainty Quantification for Monte Carlo Radiation Transport Solvers

Radiation transport computations in realistic systems are affected by the presence of uncertainty sources, e.g. nuclear cross section data and variability in geometric arrangement. Therefore, it is of paramount importance to statistically characterize the response of quantities of interest by performing accurate uncertainty quantification (UQ). Traditionally, UQ is focused on evaluating how the statistics of a numerical code response are affected by sources of uncertainty, which can be propagated through several runs of the numerical code. However, for radiation transport problems solved using Monte Carlo particle transport methods in which one sample (in the parameter space) of the quantity of interest is obtained by averaging various particles random walks, the non-deterministic nature of the solver introduces an additional source of variance. In this contribution, we describe how we can obtain more efficient and accurate sample variance estimators by taking into account, and removing, the additional variability introduced by the Monte Carlo radiation transport solvers. We will provide a rigorous mathematical treatment of these variance contributions and their sampling estimator coun-

terparts. In particular, we will present several numerical test problems in which this variance deconvolution strategy is deployed with and without scattering, in 1D slabs, and with different sources of uncertainty including uncertain material properties and stochastic mixing.

Kayla B. Clements
Sandia National Laboratory
Oregon State University
clemekay@oregonstate.edu

Aaron Olson
Sandia National Laboratories
aolson@sandia.gov

Gianluca Geraci
Sandia National Laboratories, NM
ggeraci@sandia.gov

PP1

Towards a Data Assimilation Earthquake Estimator: Using An Ensemble Kalman Filter on 2D Perfect Model Experiments

Earthquakes are among the most dangerous natural hazards. In the last decades, there has been a lot of progress in developing more accurate Early Earthquake Warning Systems (EEWS) and Probabilistic Seismic Hazard Assessments (PSHAs) to prepare the population for possible future earthquakes events. However, our ability to produce forecasts is hampered by very limited information on the current state of stress, strength, and governing parameters of the existing faults generating the earthquakes. Ensemble data assimilations (EDAs) provide a means to estimate the time evolution of the state of stress, strength, and governing parameters of existing faults and therefore the earthquakes they produce. The EDAs have the advantage of combining the estimates of physics-based models and observation while quantifying their uncertainties. Perfect model experiments with an Ensemble Kalman Filter (EnKF), connected with 2D earthquake cycle models, demonstrate the ability to estimate the state variables of shear stresses, slip velocities, and state (?) of a straight fault governed by rate-and-state friction surrounded by a homogeneous elastic medium. Furthermore, despite the inherent assumption on the Gaussianity of these distributions, the EnKF still provides a reasonable estimate of the time of occurrence of earthquakes in the 2D experiment.

Hamed A. Diab-Montero
Delft University of Technology (TU Delft)
Department of Geoscience and Engineering
h.a.diabmontero@tudelft.nl

PP1

Bayesian Inversion of a Coupled Acoustic-Gravity Model for Predictive Tsunami Simulation

To improve tsunami preparedness, early-alert systems and real-time monitoring are essential. We propose a novel approach for predictive tsunami modeling within the Bayesian inversion framework. This effort focuses on informing the immediate response to an occurring tsunami event using near-field data observation. Our forward model is based on a coupled acoustic-gravity model (e.g., Lotto and Dunham, Comput Geosci (2015) 19:327340). Similar to other tsunami models, our forward model relies on transient boundary data describing the location and magnitude of the seafloor deformation. In a real-time scenario, these parameter fields must be inferred from a variety of measurements, including observations from pressure gauges mounted on the seafloor. One particular difficulty of this inference problem lies in the accurate inversion from sparse pressure data recorded in the near-field where strong hydroacoustic waves propagate in the compressible ocean; these acoustic waves complicate the task of estimating the hydrostatic pressure changes related to the forming surface gravity wave. Furthermore, the forward model incurs a high computational complexity, since the pressure waves must be resolved in the 3D compressible ocean over a sufficiently long time span. Due to the infeasibility of rapidly solving the corresponding inverse problem for the fully discretized space-time operator, we explore options for using surrogate operators of the parameter-to-observable map.

Stefan Henneking
Oden Institute, UT Austin
stefan@oden.utexas.edu

Omar Ghattas
The University of Texas at Austin
The Odon Institute for Computational Sciences and Engr.
omar@oden.utexas.edu

PP1

Efficient UQ and Global Time-Varying Sensitivity Analysis Using the Spatially Adaptive Combination Technique

Due to their construction, conceptual hydrological models typically exhibit many parameter uncertainties with considerable variance, leading to significant uncertainty in model predictions (e.g., predictions of flood or drought events). The propagation of the resulting high-dimensional joint parameter uncertainty is a challenging task both mathematically and in terms of the computational demands. In this work, we employ a non-intrusive polynomial chaos expansion to model uncertainty in five or more input parameters that characterize high-flow conditions. We rely on sparse grid (SG) strategies to compute the expansion's coefficients while keeping the necessary model runs small. To keep the black-box property of the combination technique while focusing on regions of interest adaptively, we rely on a recently proposed spatially adaptive SG combination technique with a dimension-wise refinement algorithm. Due to the runtime of the model, parallel execution and parallel post-processing of the UQ analysis are crucial in our solution. On the example of specific hydrological models used for flood forecasting, we show that our results can give an insight into the parameters' stochastic importance and provide authorities with a reliable uncertainty band over the flow predictions in a reasonable time. Our work bridges the gap between earlier theoretical work on UQ and more complex real-world problems.

Ivana Jovanovic Buha
Technical University Munich
ivana.jovanovic@tum.de

Michael Obersteiner, Tobias Neckel
Technical University of Munich
oberstei@in.tum.de, neckel@in.tum.de

Hans-Joachim Bungartz
Technical University of Munich, Department of Informatics
Chair of Scientific Computing in Computer Science

bungartz@in.tum.de

PP1

Active Learning Within the Kennedy and O'Hagan Framework for Rare Earth Element Solvent Extraction Equilibria

The Kennedy and O'Hagan (KOH) calibration framework implemented with Gaussian Process (GP) regression can be used to assimilate real and simulated data sources for improved real world prediction. While it is common to use active learning criteria with GPs to sequentially acquire data for improved prediction, active learning for acquiring simulator data for improved real world prediction within the KOH framework is sparsely found within the statistical literature. First, we propose an integrated mean squared prediction error criteria which can be utilized with the KOH calibration framework and computed in closed form. Then, we provide an empirical comparison to other methods using results from a Monte Carlo experiment on a toy data generating mechanism. Lastly, we overview results from the motivating example: predicting equilibrium conditions of rare earth elements for a liquid-liquid extraction chemical process.

Scott Koermer, Robert Gramacy, Aaron Noble
Virginia Tech
skoermer@vt.edu, rbg@vt.edu, aaron.noble@vt.edu

PP1

Maximin Distance Designs with Mixed Continuous, Ordinal and Binary Variables

We propose a new method to construct maximin distance designs with arbitrary numbers of dimensions and points. The proposed designs hold interleaved-layer structures and handel the mixed continuous, ordinal and binary variables. Applicable to distance measures with equal or unequal weights, our method is useful for emulating computer experiments when a relatively accurate a priori guess on variable importance is available.

Hui Lan
Beijing University of Technology
huilan@bjut.edu.cn

Xu He
Academy of Mathematics and Systems Science, CAS
xuhe@amss.ac.cn

PP1

A Bayesian Framework for Parameter Estimation from Sparse and Noisy Measurement Data in Systems Biology

We propose a Bayesian parameter estimation framework to facilitate model calibration and uncertainty analysis for predictive models of biological systems. Many of these models are nonlinear differential equations that characterize the dynamics of the relevant biochemical species. System biologists need to estimate many free parameters from sparse and noisy experimental data to calibrate these models. Despite the uncertainties associated with this data, systems biologists directly fit the free parameters, typically ignoring any uncertainty altogether. Our work moves past this traditional model-fitting paradigm by adapting an approximate marginal Markov chain Monte Carlo (MCMC) method for Bayesian parameter estimation. We find that this method can recover the posterior parameter distribution from experimental data for a well-known biological system, the mitogen-activated protein kinase signaling pathway. Further, we use this method to explore several questions related to this system: *What are the effects of the experimental conditions on the recovered model parameters?* and *Can we assume that models calibrated to these varied data show equivalent behavior?* Our work introduces a new modeling paradigm that bridges the gap between standard systems biology modeling practices and rigorous uncertainty quantification.

Nate Linden
University of California, San Diego
nlinden@ucsd.edu

PP1

Spatio-Temporal Kriging from CloudSat Observations

Spatiotemporal statistical learning has received increased attention in the past decade, due to spatially and temporally indexed data proliferation, especially collected from satellite remote sensing. Observational studies of clouds are recognized as an important step to improve cloud representation in weather and climate models. Since 2006, the satellite CloudSat of NASA carries a 94 GHz cloud profiling radar and is able to retrieve, from radar reflectivity, microphysical parameter distribution such as water or ice content. The collected data is piled up with the successive satellite orbits of nearly 2 hours, leading to a large database of 2 TO. To go further than cloud analysis, it is interesting to be able to interpolate, in space, and to predict, in time, the cloud microphysics in medium and long term. Since an accurate estimation is obviously unattainable, it is an issue of uncertainty quantification. Starting from a data exploratory analysis, we have recently initiated a statistical kriging-based approach that is able to interpolate/predict from the dataset and provide uncertainties. Beforehand, it requires in particular estimating the parameters of the spatio-temporal covariance model; it is performed in a Bayesian setting, which allows for estimation and uncertainties quantification. The approach is then applied to a subset of the CloudSat dataset.

Pierre Minvielle
CEA
pierre.minvielle@cea.fr

Jean-Marie Lalande
Météo France
jean-marie.lalande@meteofrance.fr

Guillaume Bourmaud, Jean-François Giovannelli
IMS
guillaume.bourmaud@ims-bordeaux.fr, giova@ims-bordeaux.fr

PP1

Take-Away Impartial Combinatorial Game on Different Geometric and Discrete Structures

Following from the winning strategy for a Take-Away Impartial Combinatorial Game on only Oddly Uniform or only Evenly Uniform Hypergraphs in the Ph.D. Dissertation of Dr. Kristen Barnard (an Assistant Professor of Mathematics at Berea College), Molena Nguyen found the new winning strategy for a Take-Away Game on neither Oddly nor Evenly Uniform Hypergraphs during her

Undergraduate Independent Research opportunity. However, those neither Oddly nor Evenly Uniform Hypergraphs must meet the given special requirements. In a Take-Away Game on hypergraphs, two players take turns to remove the vertices and the hyperedges of a hypergraph. In each turn, a player must remove either only one vertex or only one hyperedge. When a player chooses to remove one vertex, all of the hyperedges that contain the chosen vertex are also removed. When a player chooses to remove one hyperedge, only that one chosen hyperedge is removed. Whoever removes the last vertex wins the game. All of the new theorems in this research paper are in agreement with the previous theorems in Dr. Kristen Barnards Ph.D. Dissertation.

T. H. Molena
North Carolina State University
thmolena@gmail.com

PP1

An Expected Log-Likelihood Based Acquisition Function for Bayesian Optimization.

We present a novel acquisition function that is well-suited for Multi-Objective Bayesian Optimization (MOBO) problems. Instead of sequentially sampling points that reduce the uncertainty around Quantities of Interest (QOI), we acquire new points in a region where the objective function decreases. We achieve this by exploiting an expected log-likelihood based acquisition function that efficiently balances the exploration/exploitation trade-off. The developed method incorporates a rigorous Uncertainty Quantification (UQ) on the input and output parameters and is well-adapted for design problems where a targeted output needs to be produced. We validate the presented strategy on a set of problems that arise from manufacturing experiments where our approach is particularly relevant, and compare its performance to Bayesian Optimization algorithms with common acquisition functions such as the Expected Improvement (EI) or the Upper Confidence Bound (UCB).

Marieme Ngom
University of Illinois at Chicago
mngom@anl.gov

Carlo Graziani
Argonne National Laboratory
cgraziani@anl.gov

PP1

Examining Effect of CG Error on Computational Pipelines with Probabilistic Numerics

Many computational pipelines depend on solving systems of linear equations. The Conjugate Gradient method (CG) is a widely used iterative method that solves systems of linear equations. Early termination of CG sacrifices accuracy to save computational resources. We analyze the propagation of CG error in computational pipelines, specifically the computation of a generalized singular value decomposition. The Bayesian Conjugate Gradient method (BayesCG) is a probabilistic generalization of CG that solves systems of linear equations and produces a probability distribution that models the error. By sampling from the BayesCG distribution and propagating those samples through the computational pipeline, we obtain a distribution that models the effect of CG error on the output of the computational pipeline.

Tim Reid
North Carolina State University, USA
twreid@ncsu.edu

Joseph Hart
Sandia National Laboratories
joshart@sandia.gov

Ilse Ipsen
North Carolina State University
Department of Mathematics
ipsen@ncsu.edu

PP1

Quantifying Aleatoric and Epistemic Uncertainties in RLC Circuits with Data-consistent Inversion

The characterization of uncertainty as aleatoric or epistemic is a fundamental part of solving ill-posed inverse problems. However, in practice, when epistemic uncertainty is modeled with aleatoric random variables, the distinction between aleatoric and epistemic uncertainty can become blurred, especially in Bayesian contexts. When the parameter of interest in a physical model is subject to epistemic uncertainty, the Bayesian formulation treats the parameter as a random variable whose distribution collapses to the true value as more data is observed (i.e., parameter identification). On the other hand, when the parameter of interest is subject to aleatoric uncertainty, the appropriate Bayesian formulation to the inverse problem necessitates a hierarchical approach to modeling the inherent variability of the parameter (i.e., distribution estimation). In this work, we illustrate the importance of these philosophic distinctions by comparing two inverse problems involving RLC circuits. In addition, we present new research into the measure-theoretic approach to quantifying these uncertainties called Data-consistent inversion. Both Bayesian and Data-consistent methods are applied to the two RLC circuit problems. Theoretical properties and numerical results of the solutions are compared and contrasted, along with their benefits and drawbacks.

Tian Yu Yen
University of Colorado Denver
tyen@sandia.gov

PP1

Pagp: A Physics-Assisted Gaussian Process Framework with Active Learning for Forward and Inverse Problems of Partial Differential Equations

In this work, a Gaussian process regression(GPR) model incorporated with given physical information in partial differential equations(PDEs) is developed: physics-assisted Gaussian process(PAGP). The targets of this model can be briefly described by solving forward and inverse problems of given PDEs. We introduce three different models: continuous time, discrete time and hybrid models. The given physical information is integrated into GPR model through our designed Gaussian process(GP) loss functions following the idea of penalized GPR. The first part introduces the continuous time model which treats temporal domain the same as spatial domain. The unknown coefficients in given PDEs can be jointly learned with GP hyper-parameters by minimizing the designed loss function. In the discrete time models, we first choose a time discretization scheme. Then the PAGP model is applied at each time step to-

gether with the scheme to approximate PDE solutions at given test points. To discover unknown coefficients in this setting, observations at two specific times are needed and a mixed mean square error function is constructed to obtain the optimal coefficients. In the last part, a novel hybrid model combining the continuous and discrete time models is presented. It merges the flexibility of continuous time model and the accuracy of the discrete time model. The effectiveness of the proposed PAGP methods is illustrated by five numerical experiments.

Shiqi Zhang, Jiahao Zhang, Guang Lin
Purdue University
zhan2585@purdue.edu, jiahzhang2@outlook.com, guanglin@purdue.edu