Report on BIRS Workshop: Stochastic Multiscale Methods: Bridging the Gap Between Mathematical Analysis and Scientific and Engineering Applications (11w5120)

Roger Ghanem* Guillaume Bal† George Papanicolaou‡
Boris Rozovsky§ Wing Kam Liu ¶

1 Overview

Issues of uncertainty quantification, model validation, and optimization under uncertainty have taken center stage in many areas of science and engineering. Likewise, multiscale modeling and computing capabilities are becoming the standard against which model-based predictions are gauged. It thus behooves the scientific community, at this juncture, to elucidate the mathematical foundation of stochastic multiscale concepts so as to ensure a steady evolution of scientific capabilities as engines of economical growth societal wellbeing.

Exchanging information across scales is one of the most significant challenges in multiscale modeling and simulation. By necessity, and naturally within a multiscale context, information is truncated as it is presented to a coarser scale, and is enriched as it traverses the opposite path. Information is lost and corrupted as it is, respectively, upscaled and downscaled. Mitigating these errors can be set on rigorous ground through a probabilistic description of information, whence finite-dimensional approximations of measures provides an analytical path for describing the coarsening and refining of information. Stochastic analysis, therefore, provides a rational context for the analysis of multiscale methods.

This Workshop was designed to cement a dialog between mathematicians, mechanicians, and computational scientists that will lay the foundation for an accelerated growth in stochastic multiscale methods. Thirty colleagues from academia and US national laboratories and research funding agencies participated in the Workshop, including 6 graduate students, and participants from USA, Canada, France, Belgium, England and Saudi Arabia.

While accommodations at Banff were generously covered by a BIRS grant, travel expenses for the organizers and the graduate students were covered by an NSF grant.

The Workshop featured 18 hour-long talks, spread over four days, with open discussion throughout the meeting and specially in the evenings.

2 Summary of Presentations

Paul Newton talk was entitled "Metastatic progression via biased random walks on a cancer network" in which he described procedures for developing a Markov Chain (MC) model for the progression of cancer in the lung. A 50-dimensional state of the MC is obtained from a network characterization of the lung, and associated stationary transition probabilities are estimated from a dataset containing 3827 data points. The mean first passage time on metastasis is obtained through Monte Carlo sampling and used to identify trajectories of metastasis progression.

*University of Southern California; Email: ghanem@usc.edu
†Columbia University; Email: gb2030@columbia.edu
‡Stanford University; Email: papanico@math.stanford.edu
§Brown University; Email: rozovsky@dam.brown.edu
¶Northwestern University; Email: w-liu@northwestern.edu

In his talk entitled "An elliptic inverse problem arising in groundwater flow," Andrew Stuart described an adaptation of Bayesian parameter estimation to stochastic processes. In the process, he described the extension of Bayes theorem to infinite dimensions. Further, using Polynomial Chaos representation of the posterior, an efficient sampling algorithm is developed that is competitive with MCMC procedures.

Howman Owhadi described optimal uncertainty quantification in a talk entitled "Bridging scales with incomplete information Optimal Uncertainty Quantification." Driven by the need to carry out analytical certification with limited data and poorly understood model, he presented a mathematical construction that provides optimal bounds on tail probabilities. The construction is based on concentration inequalities and McDiarmid's inequality and reduces the problem to the product of convex linear combinations of Dirac masses over a lattice. The fundamental ideas were demonstrated on a hypervelocity impact problem.

Wing Kam Liu, in his talk "Bridging of the Scales, Multiscale Modeling & Simulation of Uncertain Archetype Motion," described a new perspective on material modeling involving building blocks. These so-called "archetypes" are associated with behaviors on a variety of scales, as needed for the problem at hand. In addition to describing the practical relevance and reach of this new "science of material" the presentation described how the physical postulates of material motion can be adapted to these archetypes, resulting in governing equations that are intrinsically multiscale, and the possible sources of uncertainty associated with the new representations.

In her talk entitled "Stochastic Multiscale Analysis and Design," Wei Chen picked-up where Wing Kam Liu left and explained how the adaptation of the archetype theory to the concurrent optimization of hierarchical materials and product designs across multiple scales. She emphasized the significance of design specification to the delineation of uncertainty, and the importance of accounting for modeling errors.

Sergey Lototsky described "Mean-preserving stochastic renormalization of differential equations." By introducing Wick products into stochastic differential equations, the form of these equations is simplified, permitting the use of usual calculus instead of Ito calculus.

Habib Najm described the "Analysis of uncertain dynamical network models." Motivated by the network structure of many physical problems, the critical need for model reduction, and the significant uncertainties in both the original and reduced models, he discussed a coupling between polynomial chaos expansions and singular perturbation-like analysis. A spectral analysis of the discrete system of equations resulting from the stochastic Galerkin projection is thus analyzed, and stability properties of the underlying stochastic physical system are investigated.

Jim Nolen's presentation on "Importance sampling for random elliptic equations," was about the use of importance sampling to compute statistics of the solutions to a (elliptic) PDE with random coefficients. In the context of a Monte Carlo simulation to compute the statistics, one wants to minimize the number of samples needed for accurate statistical estimation, since generating each sample is an expensive computation (solving a PDE). An importance sampling scheme was described based on linearization and asymptotic analysis of the map from coefficients to solution.

In his talk "Kinetic limits and imaging models for waves in random media," Olivier Pinaud described methods for imaging inclusions buried in a random medium from detector measurements. The focus was on a regime where the interaction between the wave and the medium is strong. After reviewing relevant scales and some standard mathematical results associated with this problem, a new approach based on a transport problem was proposed. Numerical examples were used to provide an interpretation and shed light on the mathematical statements.

In his talk "Variance reduction and multiscale FEM for random media", Frédéric Legoll presents efficient yet accurate numerical approaches for characterizing random media that capitalize on an assumed random perturbation around a periodic mean. For strongly stochastic materials, a variance reduction approach is proposed, while for weakly random materials, specialized approaches are also described that take advantage of multiscale and homogenization approaches.

In her talk "Noise sensitivities in systems with delays and multiple time scales," Rachel Kuske described the effect of delays and multiple time scales on the dynamics of nonlinear systems. Dynamical systems with delayed feedback often exhibit complex oscillations that are significantly affected by stochastic fluctuations, particularly if multiple time scales are present. Then transients ignored in the deterministic system can dominate the long range behavior. The approaches described capture the effects of noise and delay in the contexts of piecewise smooth systems, nonlinearities, and discontinuities.

Wenjia Jing talked about "Fluctuation in random homogenization: motivations, corrector theory, and algorithm test." He introduced a corrector theory that provides an estimate for the fluctuations around the homogenized solution with significant implications to inverse problems. For elliptic equations with random multiscale potentials, a systematic corrector theory was developed that clearly depend on the regularity of Green's function and correlation range of the underlying random field.

Maarten Arnst presented a talk entitled "Dimension reduction and measure transformation in stochastic multiphysics modeling," described the possibility and significance of affecting a transformation of probability measures as information is exchanged between models in a multiphysics problem. Issues of convergence and embedded quadrature were also discussed.

Youssef Marzouk discussed "Multiscale and map-based methods for statistical inference in inverse problems." Specifically, methods for characterizing the mapping between the prior and posterior random variables are presented.

In his talk, "Adaptive Multi-Level Monte Carlo Simulation," Raul Tempone generalizes a multilevel Forward Euler Monte Carlo method introduced in [1] for the approximation of expected values depending on the solution to an Ito stochastic differential equation. Previous work proposed and analyzed a Forward Euler Multilevel Monte Carlo method based on a hierarchy of uniform time discretizations and control variates to reduce the computational effort required by a standard, single level, Forward Euler Monte Carlo method. The present work introduces and analyzes an adaptive hierarchy of non-uniform time discretizations, generated by adaptive algorithms. These adaptive algorithms apply either deterministic time steps or stochastic time steps and are based on adjoint weighted a posteriori error expansions. Under sufficient regularity conditions, both in the present analysis and numerical results, which include one case with singular drift and one with stopped diffusion, significant savings in the computational cost were observed.

Alireza Doostan's talk, "On sparse approximation techniques for uncertainty propagation", described the adaptation of LASSO techniques to the characterization of the solution of partial differential equations with stochastic coefficients. Significant savings both in terms of representation effort and computational effort are described.

In his talk entitled "tP-CKM: Bayesian continuum closure evaluation from stochastic microdynamics," Sorin Mitran considers a generic problem of great significance in polymer flows and several biological applications. The problem exhibits dependence of the continuum scale on microconfigurations that are themselves affected by boundary conditions imposed on the continuum. The physics is specified at the microscale which is stochastic and does not equilibrate quickly. The approach involves predictions at the continuum scale (assumed to be true), through a hierarchical Bayesian formalism across scales (microscale and kinetic scale). Adaptation of the formalism to advanced computational platforms, including GPUs, is also discussed.

Finally, Sonjoy Das talked about an "An energy based stochastic mapping between high and low fidelity models." In this talk a framework and algorithms were described to develop stochastic coarse scale models that encapsulate errors stemming from upscaling from a deterministic fine scale.

3 Summary of the Breakout Sessions

Discussion sessions were held each evening, kicking off for a few hours at one of the classrooms. This section presents a synthesis of those discussions.

3.1 Synergies between UQ and Multiscale

There are clear synergies between multiscale physical behavior and mathematical analysis on one hand, and uncertainty quantification on the other hand. Indeed crossing scales entails shedding or interpolating for information, while uncertainty in behavior is often due to omissions in coupling to subscale information (physical processes or measurements). This tight linkage can be described as follows:

1. multiscale provides possible physical context for uncertainty.

2. scales cannot be all well characterized, even with access to higher performance computing resources. Uncertainty analysis is thus needed to represent irrecoverable information.

3. the implications of multiscale uncertainties have to be described in the context of specific predictions (on specific quantities of interest) and specific decisions, with well defined objectives functions.

3.2 Typical Applications

While it was recognized that multiscale and uncertainty analyses are ubiquitous in science, engineering and the physical world, an attempt was made to delineate the scope of relevance of multiscale processes as captured by the Workshop participants. The following (incomplete) list was produced, and thought to reflect many current areas in science where synergetic advances in uncertainty quantification and multiscale analysis is likely to lead to significant mathematical, scientific, and technological breakthroughs.

interacting multiscale systems (social/infrastructures)	subsurface flows including carbon sequestration and nuclear repository (QoI are fracture, failure, extremes) and environmental remediation (percolation, homogenized limits).
power grid: fault analysis on huge systems	materials science: homogenized systems, fracture, failure, design, meta materials, damage.
complex fluids	regulatory finance
turbulence	interfaces
weather and geophysical processes (eg. Ice)	climate
biological processes (cellular, biomechanics, etc)	

3.3 Components of a Multiscale System and Associated b Challenges

Some features of a coherent stochastic multiscale system emerged following a number of discussions to include **(1)** Characterization, **(2)** Computing, **(3)** Probabilistic Modeling, **(4)** Validation, **(5)** Scientific Discovery, and **(6)** Decision & Design.

A description of any of these features seems to be strongly dependent on an identification of the quantity (or quantities) of interest. Under, **Characterization**, measurements of the same phenomenon at different scales was deemed significant in order to understand the structure of statistical dependence between the different scales. This in turn requires developing models of sensors. It is clear that a combination of statistical and physics-based understanding would be critical at this stage. Hierarchical models, including hierarchical Bayesian models were mentioned as significant for addressing some of these challenges. **Computing** is clearly a critical driver of research for both multiscale and uncertainty quantification. It plays a role in uncertainty quantification, and in integration of analyses across scales. Mathematical challenges related to computing pertain essentially to convergence, error estimation, and numerical efficiency. It was felt that **Probabilistic Modeling** is a key component for both formulating a well-posed problem (that does not violate physical premises or mathematical hypotheses), a useful problem (that acknowledges the quantity of interest) and an optimal problem (that packages available evidence appropriately). The issue of **Validation** was deemed critical for the multiscale problems. In particular, the validation of single-scale simulations, cross-scale exchange, and coupled multiscale simulations seem to each be important and challenging in its own right. There are conceptual, statistical, and physical aspects to each of these pieces. The issue of **Scientific Discovery** was identified as an important contribution of both multiscale and uncertainty analyses. The discovery of emergent behaviors from multiscale couplings, and from interaction of fluctuating parameters is a clear path for discovering behaviors and issues not measurable in a deterministic or single-scale setting. This highlights the value of numerical explorations without the imperative of a clear quantity of interest. Alternatively, one could try to formulate quantities of interest that are consistent with the purpose of scientific discovery. Finally, it was also felt that in most cases, understanding the ultimate purpose of a numerical simulation, whether **Decision or Design**, is paramount for developing acceptable probabilistic models and ascertaining the value of multiscale enrichment.

3.4 Some Mathematical Challenges

We had a long discussion on reduced order models (ROM), including reductions within a single-scale and within a multiple scale context.

Some issues that were discussed:

1. Reduced order models (ROM) must be tested against mathematically provable limits.

2. ROM must be adapted to achievable data modalities.

3. Characterize the combined uncertainty in cross-scaling/model reduction.

4. statistical/logical/mathematical/computational relations between data from across scales.

5. does prediction of individual specimen behavior make sense ? if not, what are limits on prediction ? how can multiscale push those limits ?

6. Bayesian paradigm revisited: what is our prior information, and what is the mathematical structure available to compute the likelihood in a multiscale context, with data and uncertainty at every scale and junction.

7. curse of dimensionality.
   - minimnum-maximum spanning sets in high dimensions.
   - low rank tensor approximation and sparse approximation.
   - connection to reality (are observable really high-dimensional ?)

8. mathematical formulations and analysis of uncertainty propagation.
   - analysis on tensor spaces.
   - verification and error analysis.

9. inference from limited data.

10. validation of key concepts including upscaling.

11. sufficiency of data for subscale reconstruction.

12. upscaling as a physical process and as an information exchange process.

13. to what extent can subscale be constrained by coarse scale ?

14. constrain mathematical models with multiscale features.

15. algorithms/ computation: model-based simulations and statistical computing.

16. stochastic programming for multiscale problems.

17. constraining inference with multiscale information and practical considerations (too little data, incomplete models).

18. design-informed separation of scales.

3.5 Grand Challenges

A small number of grand challenges were identified as generically relevant across all applications and which are likely to spawn new physics and mathematics. These are stated as follows:

- Define multiscale data needed to be sufficiently predictive for a given QoI
- Quantify credibility of multiscale coupling and information passing
- Understand implications of information fusion from multiple scales with varying levels of confidence
- Develop benchmark problems to drive research.