SAND2011-4127C

Uncertainty Quantification in Multiscale Atomistic-Continuum Models

M. Salloum, K. Sargsyan, R. Jones, B. Debusschere, H. Najm, and H. Adalsteinsson

Sandia National Laboratories, Livermore, CA

Uncertainty Quantification and Multiscale Materials Modeling Workshop
Santa Fe, NM; June 13-15, 2011

Acknowledgement

This work was supported by:

The US Department of Energy (DOE), Office of Advanced Scientific Computing Research (ASCR), Applied Mathematics program.

Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94-AL85000.

Outline

1 Motivation

2 Basics

3 UQ in Multiscale AtC Systems

4 Closure

The Case for Uncertainty Quantification

UQ is needed in:

- Assessment of confidence in computational predictions
- Validation and comparison of scientific/engineering models
- Design optimization
- Use of computational predictions for decision-support
- Assimilation of observational data and model construction
- Multiscale and multiphysics model coupling

Sources of Uncertainty

- model structure
  - participating physical processes
  - governing equations
  - constitutive relations
- model parameters
  - transport properties
  - thermodynamic properties
  - constitutive relations
  - rate coefficients
- initial and boundary conditions
- geometry
- truncation errors

UQ in Multiscale Atomistic-to-Continuum Models

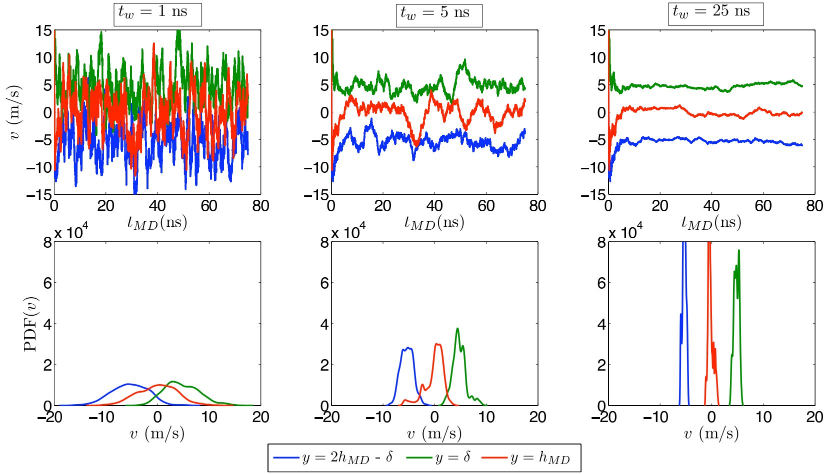
Consider a system comprised of:

- a molecular dynamics (MD) box coupled to
- a continuum simulation

Consider uncertainty in both:

- model parameters *e.g.*
  - force field pair-potential parameters
  - continuum constitutive law parameters
- Coupling terms
  - finite-time averaging of MD statistics

Effect of Averaging time window in AtC Coupling



Overview of UQ Methods

Estimation of model/parametric uncertainty

- Expert opinion, data collection
- Regression analysis, fitting, parameter estimation
- Bayesian inference of uncertain models/parameters

Forward propagation of uncertainty in models

- Local sensitivity analysis (SA) and error propagation
- Fuzzy logic; Evidence theory — interval math
- Probabilistic framework — Global SA / stochastic UQ
  - Random sampling, statistical methods
  - Polynomial Chaos (PC) methods
    - Collocation methods — sampling — non-intrusive
    - Galerkin methods — direct — intrusive

Bayes formula for Parameter Inference

- Data Model (fit model + noise): $y = f(\lambda) + \epsilon$
- Bayes Formula:

$$p(\lambda, y) = p(\lambda|y)p(y) = p(y|\lambda)p(\lambda)$$

$$\underbrace{p(\lambda|y)}_{\text{Posterior}} = \frac{\overbrace{p(y|\lambda)}^{\text{Likelihood}}\ \overbrace{p(\lambda)}^{\text{Prior}}}{\underbrace{p(y)}_{\text{Evidence}}}$$

- Prior: knowledge of λ prior to data
- Likelihood: forward model and measurement noise
- Posterior: combines information from prior and data
- Evidence: normalizing constant for present context

Polynomial Chaos Methods for UQ

- Model uncertain quantities as random variables (RVs)
- Any RV with finite variance can be represented as a Polynomial Chaos expansion (PCE)

$$u(\mathbf{x}, t, \omega) \simeq \sum_{k=0}^{P} u_k(\mathbf{x}, t) \Psi_k(\boldsymbol{\xi}(\omega))$$

- $u_k(\mathbf{x}, t)$ are mode strengths
- $\boldsymbol{\xi}(\omega) = \{\xi_1, \cdots, \xi_n\}$ is a vector of standard RVs
- $\Psi_k()$ are functions orthogonal w.r.t. the density of $\boldsymbol{\xi}$

- with dimension n and order p:

$$P + 1 = \frac{(n+p)!}{n!p!}$$

Orthogonality

By construction, the functions $\Psi_k()$ are orthogonal with respect to the density of the basis/*germ* ξ

$$u_k(\boldsymbol{x},t) \;=\; \frac{\langle u\Psi_k\rangle}{\langle \Psi_k^2\rangle} = \frac{1}{\langle \Psi_k^2\rangle}\int u(\boldsymbol{x},t;\lambda(\boldsymbol{\xi}))\,\Psi_k(\boldsymbol{\xi})p_{\boldsymbol{\xi}}(\boldsymbol{\xi})d\boldsymbol{\xi}$$

Examples:

- Hermite polynomials with Gaussian basis
- Legendre polynomials with Uniform basis, ...
- Global versus Local PC methods
  - Adaptive domain decomposition of the stochastic support of u

Non-intrusive Spectral Projection (NISP) PC UQ

- *Sampling*-based
- Relies on black-box utilization of the computational model
- Evaluate projection integrals *numerically*
- For any model output of interest $\phi(\boldsymbol{x}, t; \lambda)$:

$$\phi_k(\boldsymbol{x}, t) = \frac{1}{\langle \Psi_k^2 \rangle} \int \phi(\boldsymbol{x}, t; \lambda(\boldsymbol{\xi}))\, \Psi_k(\boldsymbol{\xi}) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi}, \quad k = 0, \ldots, P$$

- Integrals can be evaluated using
  - A variety of (Quasi) Monte Carlo methods
  - Quadrature/Sparse-Quadrature methods

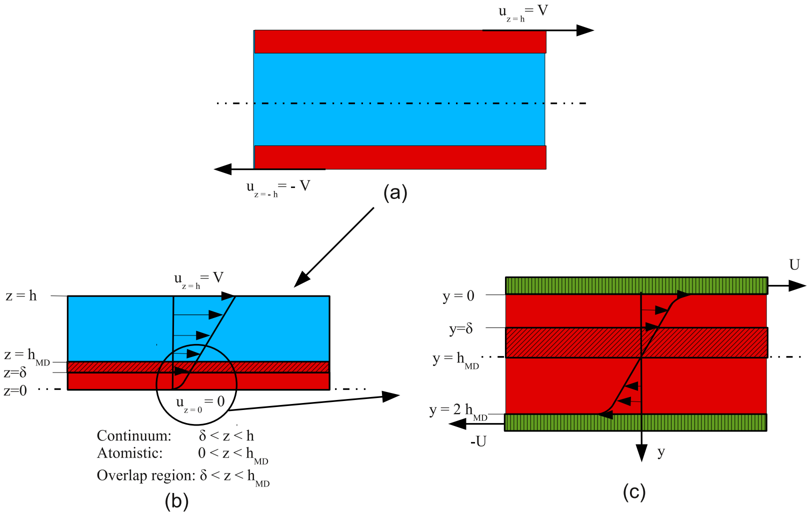
PCE Construction for Noisy Functions

- Quadrature formulae presume a degree of smoothness
  - No convergence for a noisy function

$$u_k = \frac{1}{\langle \Psi_k^2 \rangle} \int u(\lambda(\xi))\, \Psi_k(\xi) p_\xi(\xi) d\xi, \quad k = 0, \ldots, P$$

- Sparse-Quadrature formulae are *ill-conditioned* and highly-sensitive to noise
  - No convergence with order
  - Error grows with increased dimensionality
- Options in the presence of noise:
  - RMS fitting for PC coefficients
  - Bayesian inference of PC coefficients

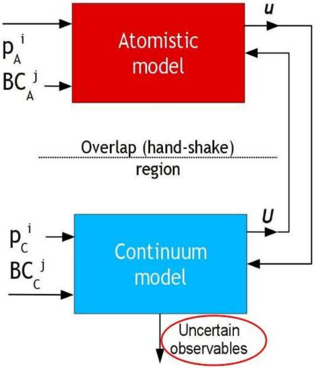
Multiscale System Geometry – Couette Flow

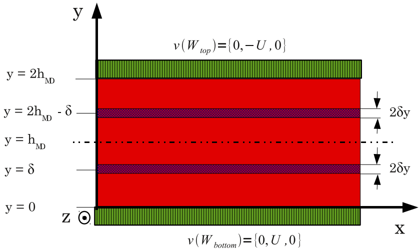


AtC Coupling Strategy

Velocity-velocity coupling

- Atomistic model output velocity u inferred based on **data**: $\{v^j\}_{j=1}^N$
  - $N = N_r \times N_t$
  - N_r replica MD simulations
  - N_t time-window averaged MD velocities per sim.
- Continuum model output velocity U





Uncertainty versus Numerical/Truncation Errors

- In the presence of parametric uncertainty, judicious choice of the
  - MD averaging window
  - length of runs, and
  - number of replicas

  can be made
- Uncertainty due to MD-averaging noise can be
  - controlled to be of same order as parametric uncertainties
  - accounted for in the same framework as that of parametric uncertainty

UQ Procedure

NISP Procedure

- Parametric Uncertainty can be handled using quadrature
  - an outer loop over quadrature points
- For each vector of parameter values, at each quadrature point, we define a UQ problem involving only noise
  - an inner loop employing PC + Bayesian inference

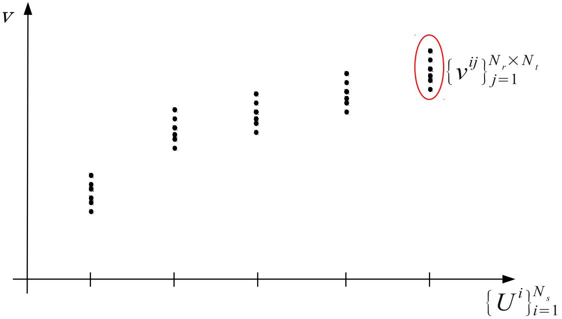
Focus on inner loop

- Only uncertainty due to MD-averaging noise
- Fixed point iteration to converge on constant uncertain $u = \sum_k u_k \Psi_k$ and $U = \sum_k U_k \Psi_k$, or chosen observable

Noise-UQ Inner Loop Procedure

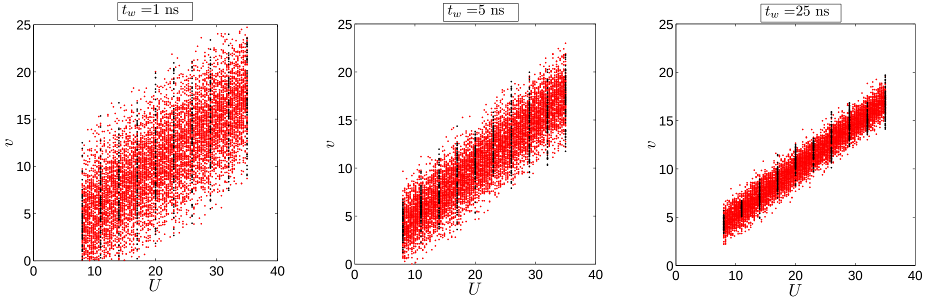
- Introduce a single degree of freedom ξ to represent noise-induced uncertainty
- Given $U = \sum_k U_k \Psi_k(\xi)$, generate samples $\{U^i\}_{i=1}^{N_s}$
- For each U^i sample, evaluate atomistic model output velocity data $\{v^{ij}\}_{j=1}^{N}$
- Use Bayesian inference to estimate $u = \sum_k u_k \Psi_k(\xi)$
- Propagate PCE of u through Continuum model
    - Arrive at new iterate of PCE of U

MD Data Noise for a Range of U



- Data for Bayesian inference
  $D = \{v^{ij}\} \quad i = 1, \dots, N_s \quad j = 1, \dots, N$
- Direct use of MD computations in repeated iterations is prohibitive
- Construct a PC surrogate for atomistic model output

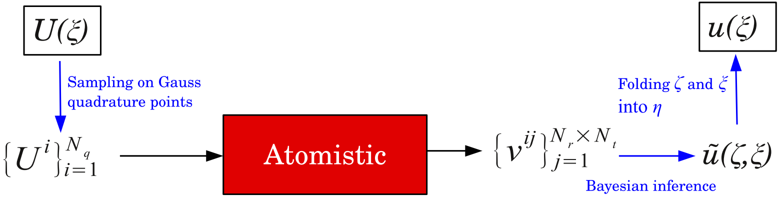
MD Data Surrogate



Surrogate: $v = v_0 + v_1 U + \sigma\eta, \qquad \eta \overset{i.i.d.}{\sim} N(0,1) \quad \text{(CLT)}$

- Bayesian inference for $p(v_0, v_1, \sigma | D)$
- $D =$ data from MD simulations
- Posterior predictive replicates noisy MD data for any U

Uncertainty Propagation in MD Model



- Input velocity: $U(\xi) = \sum_{k=0}^{P} U_k \Psi_k(\xi)$
- Output velocity: $\tilde{u} = \sum_{k=0}^{P} \tilde{u}_k \Psi_k(\xi)$
- Data from Surrogate Posterior Predictive: $\{v^{ij}\}_{i=1j=1}^{N_q\ \ N}$
- Bayesian inference $\Rightarrow$ posterior: $p(\tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_P | D)$
- With improper uniform priors on $\tilde{u}_k$ and Jeffreys prior on σ :

$$[\tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_P]^T \sim \mathcal{S}(\gamma, \boldsymbol{\lambda}, \boldsymbol{\Sigma})$$

Multivariate Student-t distribution

Uncertainty Propagation in MD Model

With $\tilde{\boldsymbol{u}} \equiv [\tilde{u}_0, \tilde{u}_1, \ldots, \tilde{u}_P]^T$, $\boldsymbol{\Psi} \equiv [\Psi_0, \Psi_1, \ldots, \Psi_P]^T$

$$\begin{aligned}
\tilde{u}(\xi) &= \sum_{k=0}^{P} \tilde{u}_k \Psi_k(\xi) = \tilde{\boldsymbol{u}}^T \boldsymbol{\Psi}(\xi) \\
&\sim \mathcal{S}\left(\gamma, \boldsymbol{\Psi}(\xi)^T \cdot \boldsymbol{\lambda}, [\boldsymbol{\Psi}(\xi)^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\Psi}(\xi)]^{1/2}\right)
\end{aligned}$$

Student-t process

or, with $\zeta \sim \mathcal{S}(\gamma, 0, 1)$,

$$\tilde{u}(\xi, \zeta) = \boldsymbol{\Psi}(\xi)^T \cdot \boldsymbol{\lambda} + \zeta \, [\boldsymbol{\Psi}(\xi)^T \cdot \boldsymbol{\Sigma} \cdot \boldsymbol{\Psi}(\xi)]^{1/2}$$

Using the inverse CDF transform, construct the 1D PCE

$$u = \sum_{k=0}^{P} u_k \Psi_k(\xi) \stackrel{\text{dist}}{=} \tilde{u}(\xi, \zeta)$$

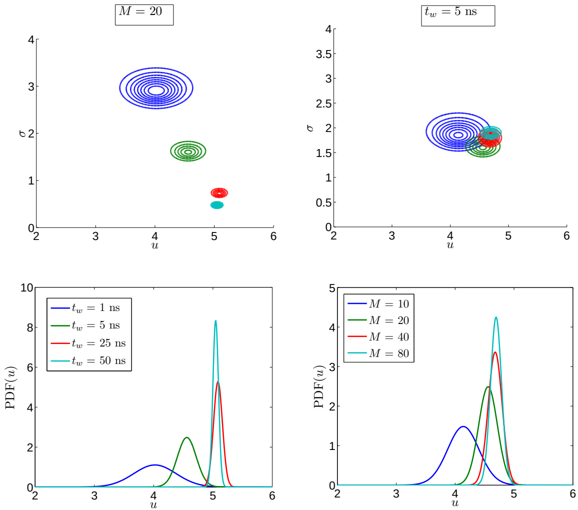
Coupled System

- Atomistic model $u(\xi) = f_a[U(\xi)]$
- Analytical Couette flow model $U(\xi) = f_c[u(\xi)]$
- Fixed point iteration: for $i = 1, 2, \ldots$

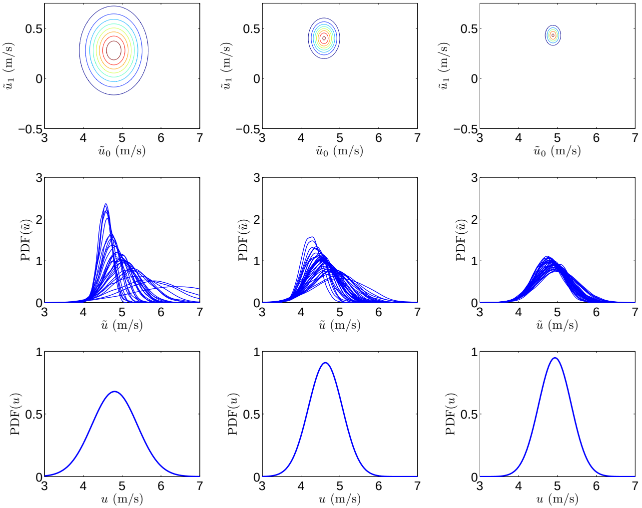
$$\begin{aligned} u^i(\xi) &= f_a^i[U^{i-1}(\xi)] \\ U^i(\xi) &= f_c[u^i(\xi)] \end{aligned}$$

  iterate till convergence
- Options: with/without Sequential Bayesian Update (SBU)
  - Without: Same prior in each iteration
  - With: Prior$_i \equiv$ Posterior$_{i-1}$

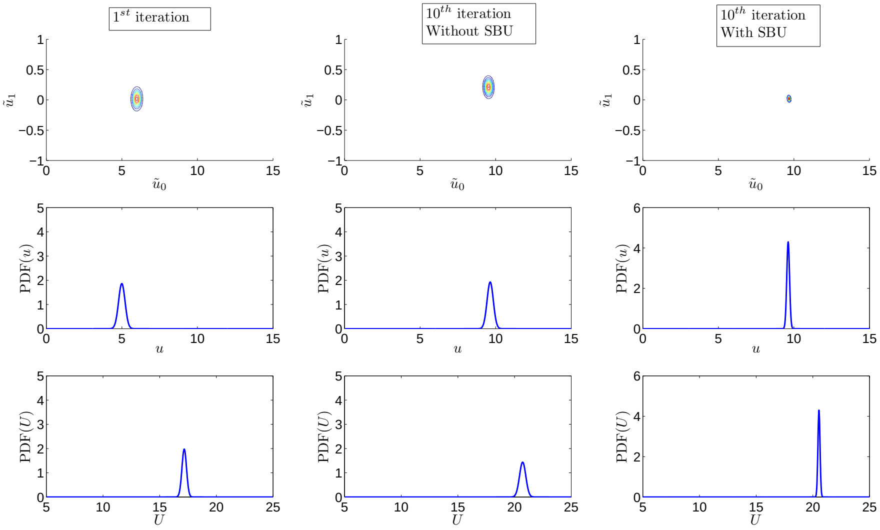
Given U^i – Effect of Avg. Window & # MD samples



Uncertain Input $U(\xi)$ – Effect of Avg. Window



Coupled System – With or Without SBU



Closure

- Outlined a parametric+noise UQ strategy in coupled atomistic-continuum multiscale models
  - Quadrature for parametric uncertainty
  - Bayesian Inference for handling noise uncertainty
    - then quadrature
  - PC Surrogate for MD box
- Results highlight trade-off between averaging window and prevailing uncertainty
- Faster convergence with SBU
- Stopping criterion when
  - change in one iteration is less than specified tolerance
  - observable uncertainty is comparable to that in the surrogate model parameters
- Other alternatives beside fixed point iteration