Statistical Aspects of a Multiscale Computational Scheme for Complex Fluids

Peter Kramer[1,2],
Sorin Mitran[1,3], Sunli Tang[2], M. Susie Bayarri[1,4],
James Berger[1,5], Murali Haran[1,6], Hans Rudolf Künsch[1,7]

[1]Statistics and Mathematical Sciences Institute
[2]Rensselaer Polytechnic Institute
[3]University of North Carolina at Chapel Hill
[4]Universitat de Valencia
[5]Duke University
[6]Pennsylvania State University
[7]ETH Zurich

May 3, 2012

Statistical Investigations in Multiscale Computing

Time-Parallel Continuous-Kinetic-Molecular (tp-CKM):

- Offline assessment of quality of mesoscale representation frameworks
- Online assessment of consistency between microscale variables and mesoscale PDF
- Estimation of secular errors in microscale-mesoscale comparisons

Note that statistical model validation and verification is done on-the-fly!

Elements of tp-CKM (Mitran & Young 2011)

Molecular: ensemble of detailed trajectories of microscale variables $\{\mathbf{Y}^{(j)}(t)\}_{j=1}^{N}$; $N \sim 10^4$ realizations

Kinetic: approximate propagation of mesoscale PDF $\tilde{\psi}(\mathbf{y}, t)$ of microscale variables

Continuum: evolution of macroscale field variables

- closure requires some local statistic of microscale variables $\tau(t) = \mathbf{F}(\tilde{\psi}(\cdot, t))$ (dependence on macroscale spatial variable suppressed)

Motivation for method is when statistics of microscale can be much slower than their nominal time scale.

- rupture, cellular "blebbing,"...

My particular interest is primarily on the mesoscale and its interaction with the microscale.

Mesoscale Representation of PDF: Challenges

Microscale variable $\mathbf{Y}$ is high-dimensional ($\sim 3M - 5$ for polymer with M monomers)

- require some simplification to permit efficient propagation and sampling
- but rich enough to capture important statistical structure (correlations, ...)
- parametric approaches (fitting Gaussian or Gibbs-Boltzmann distribution) may or may not be appropriate for desired non-equilibrium applications

But microscale data is relatively abundant ($N \sim 10^4$ realizations).

Mesoscale Representation of PDF: One Implementation

Expansion into radial basis functions

$$\tilde{\psi}_{\mathbf{Y}}(\mathbf{y},t) \equiv \sum_{\ell=1}^{L} c_\ell(t) B\left(\|\mathbf{y}-\mathbf{y}^{(\ell)}\|\right)$$

where

- $\{\mathbf{y}^{(\ell)}\}_{\ell=1}^{L}$ are points interpolating between mean of $\tilde{\psi}(\mathbf{y},t_n)$ at beginning of macro time step and mean of estimated $\tilde{\psi}(\mathbf{y},t_{n+1})$ at end of macro time step ($L \sim 10^2$)
  - uses Jordan-Kinderlehrer-Otto (JKO) gradient flow formalism for Fokker-Planck equation
- $B(y)$ is a radial basis function, taken as Gaussian

$$B(y) = \frac{\exp(-y^2/(2h^2))}{\sqrt{2\pi h^2}}$$

  with h a fixed width parameter
- $c_\ell(t)$ are coefficients computed through JKO and spline interpolation

Mesoscale Representation of PDF: Research Directions

Systematic offline assessment of quality of mesoscale representation framework

- direct confrontation with microscale simulation data via statistical hypothesis testing
- statistical challenges:
  - high-dimensional
  - hypothesis for PDF framework does not have specific parameters
- explore and develop multivariate nonparametric hypothesis tests
- formulate test based on metric involving statistic $\tau(t) = \mathbf{F}(\tilde{\psi}(\cdot, t))$ of interest to macroscale?
- ascertain adequate level of statistical resolution via tests for statistically significant improvement of goodness-of-fit of data in richer frameworks

Compare against alternative mesoscale representation frameworks

- Markov chain models
- Normal modes

Statistical Comparison: Microscale $\to$ Mesoscale

From simulated ensemble $\{\mathbf{Y}^{(j)}(t)\}_{j=1}^{N}\}$, want to compare its fidelity with predicted mesoscale PDF $\tilde{\psi}(\mathbf{y}, t)$.

Currently, one way this is done is by projecting the microscale data onto the vector space spanned by the same radial basis functions as used to represent the mesoscale, to obtain PDF estimated by microscale data:

$$\hat{\psi}_{\mathbf{Y}}(\mathbf{y}, t) \equiv \sum_{\ell=1}^{L} \hat{c}_\ell(t) B\left(\|\mathbf{y} - \mathbf{y}^{(\ell)}\|\right)$$

- only $\{\hat{c}_\ell\}_{\ell=1}^{L}(t)$ depend on the data

Comparison of microscale and mesoscale data is done by appropriate vector space norm $\|\mathbf{c} - \hat{\mathbf{c}}\|$ with

$$\mathbf{c} = \{c_\ell\}_{\ell=1}^{L}, \qquad \hat{\mathbf{c}} = \{\hat{c}_\ell\}_{\ell=1}^{L}.$$

Statistical Comparison: Microscale $\rightarrow$ Mesoscale

From simulated ensemble $\{\mathbf{Y}^{(j)}(t)\}_{j=1}^{N}\}$, want to compare its fidelity with predicted mesoscale PDF $\tilde{\psi}(\mathbf{y}, t)$.

More recent developments (Mitran):

- Bayesian hypothesis testing for Gaussian parametric model,
- lattice-Boltzmann type representation

Research direction:

- Develop new online parametric hypothesis tests for how well microscale data is actually consistent with proposed mesoscale representation when not Gaussian or comparably simple

Systematic Biases in Microscale-Mesoscale Comparisons

The microscale-mesoscale comparisons may not detect a small systematic error which accumulates.

Ideas:

- Statistical test for correlations in errors
- Estimate accumulated error through "vectorial" addition of errors at each time step
- Use mesoscale theory to estimate amplification and accumulation of error

Work in Progress

Time-Parallel Continuous-Kinetic-Molecular (tp-CKM):

- Offline assessment of quality of mesoscale representation frameworks
- Online assessment of consistency between microscale variables and mesoscale PDF
- Estimation of secular errors in microscale-mesoscale comparisons

General goals:

- Stimulate new directions of statistical research motivated by multiscale computation
- Incorporate more sophisticated statistical ideas into multiscale computation

References on tp-CKM

S. Mitran, J. Young, "Multiscale Computation of Cytoskeletal Mechanics During Blebbing," pp. 345-372, in *Cellular and Biomolecular Mechanics and Mechanobiology*, Springer, A. Gefen (ed.), 2011.

- See in particular Sec. 3 for description of method
- download from `http://mitran.web.unc.edu/publications/`

S. Mitran, "Time parallel kinetic-molecular interaction algorithm for CPU/GPU computers," *Proc. Comp.Sci.* **1**:745-752, 2010.

- Elements of combining parareal and multiscale computing

Sorin Mitran will be visiting this KITP program June 11-29.