SAND2012-2575C

Quantifying Sampling Noise and Parametric Uncertainty in Coupled Atomistic-Continuum Simulations

Maher Salloum, Khachik Sargsyan, Reese Jones,
Bert Debusschere, Habib N. Najm
Sandia National Laboratories, Livermore, CA
{mnsallo, ksargsy, rjones, bjdebus, hnnajm}@sandia.gov

Funded by the DOE Office of Science,
Advanced Scientific Computing Research,
Applied Mathematics Program

Overview

- Introduction and Motivation
- Couette Flow Test Case
- Coupling Atomistic Sampling Noise and Parametric Uncertainty
  - Intrusive Spectral Projection (ISP)
  - Non-Intrusive Spectral Projection (NISP)
- Conclusions

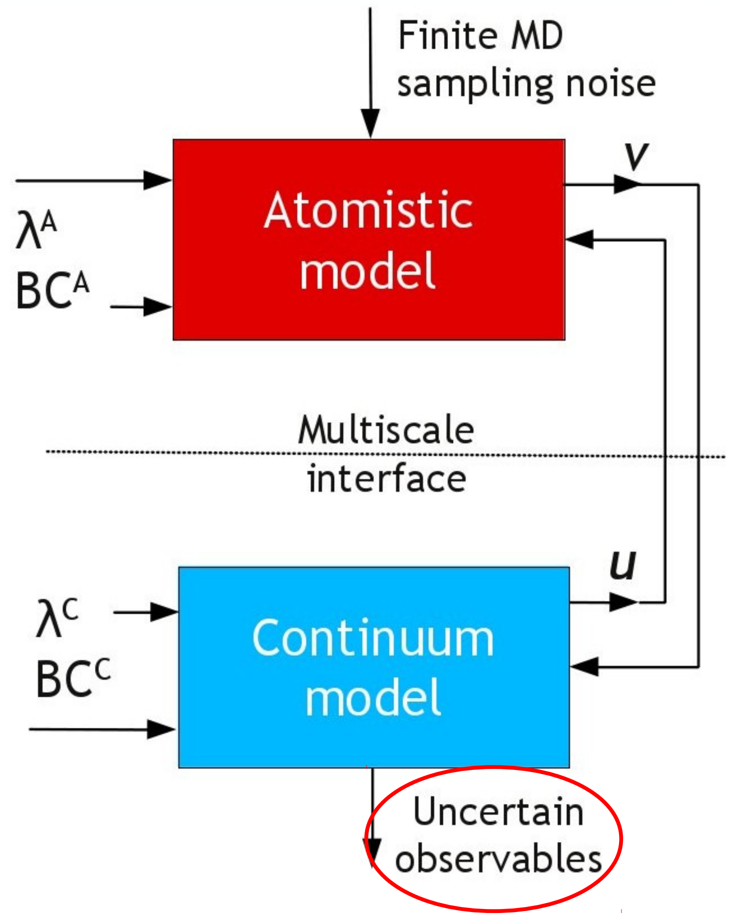
Motivation

- Physical systems often require multiscale simulations to capture phenomena occurring at both bulk and interface (*e.g.* ionic flux through nanopores in water desalination).

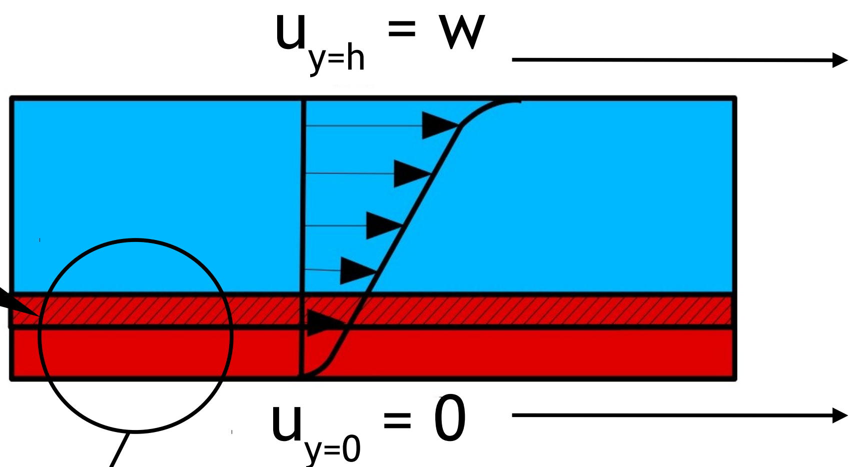
- Modeling such systems requires exchange of information between the different scales.

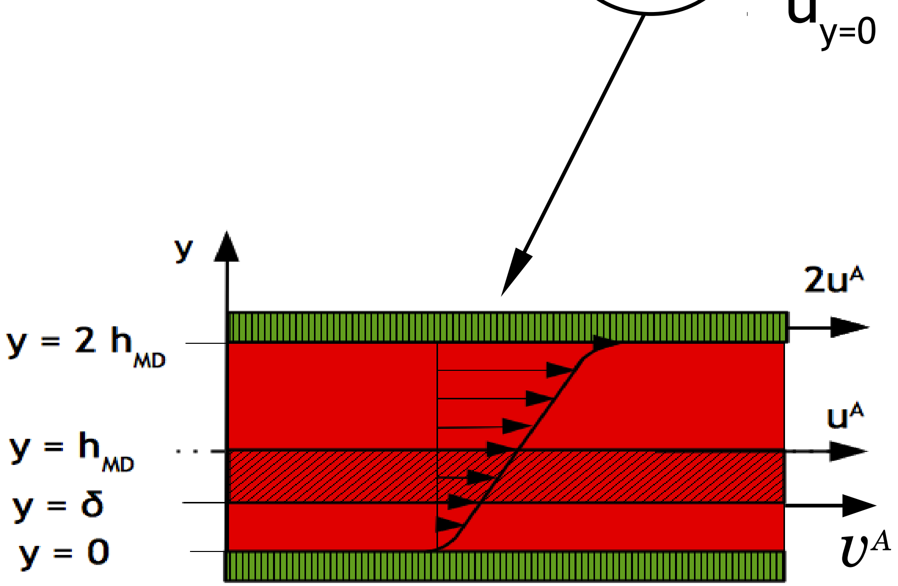
- Uncertainty quantification is needed in order to get a predictive fidelity of the multiscale simulation.

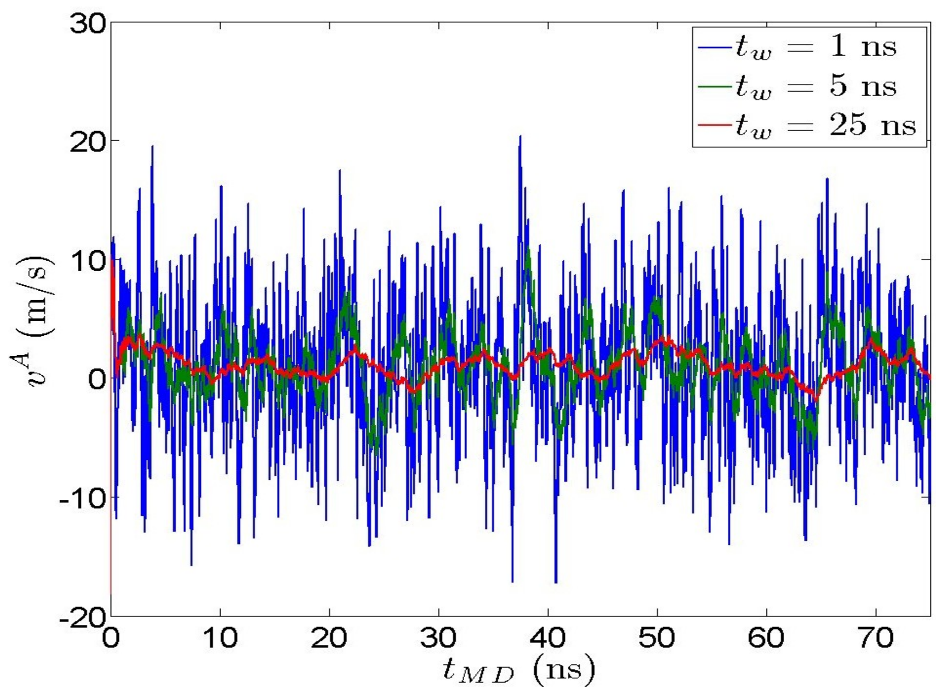
Based on all inputs into the simulation, what is the resulting uncertainty in the predicted value of the coupling variables?



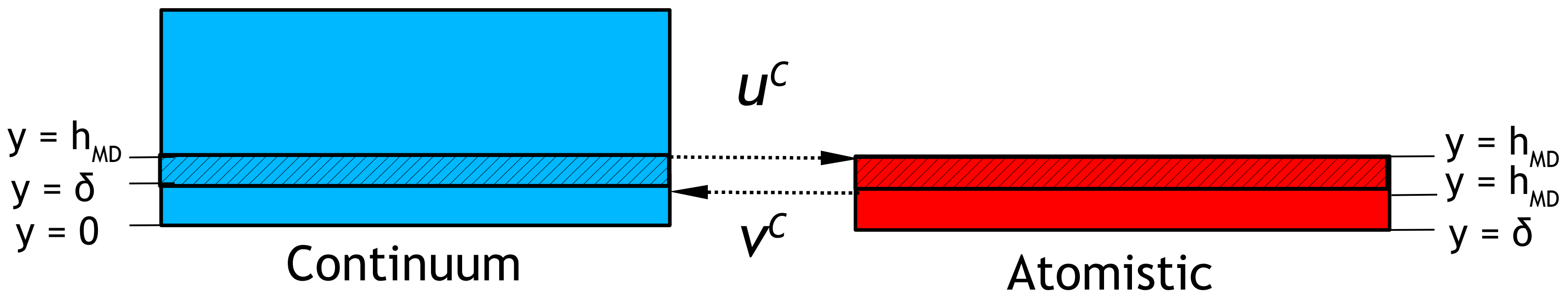
Canonical plane Couette flow is used as model problem for algorithm development







Exchange of variables



• We assume strong separation in the time scales such that we can extract one macroscale uncertain variable v^c from the fluctuating output of the atomistic simulation.

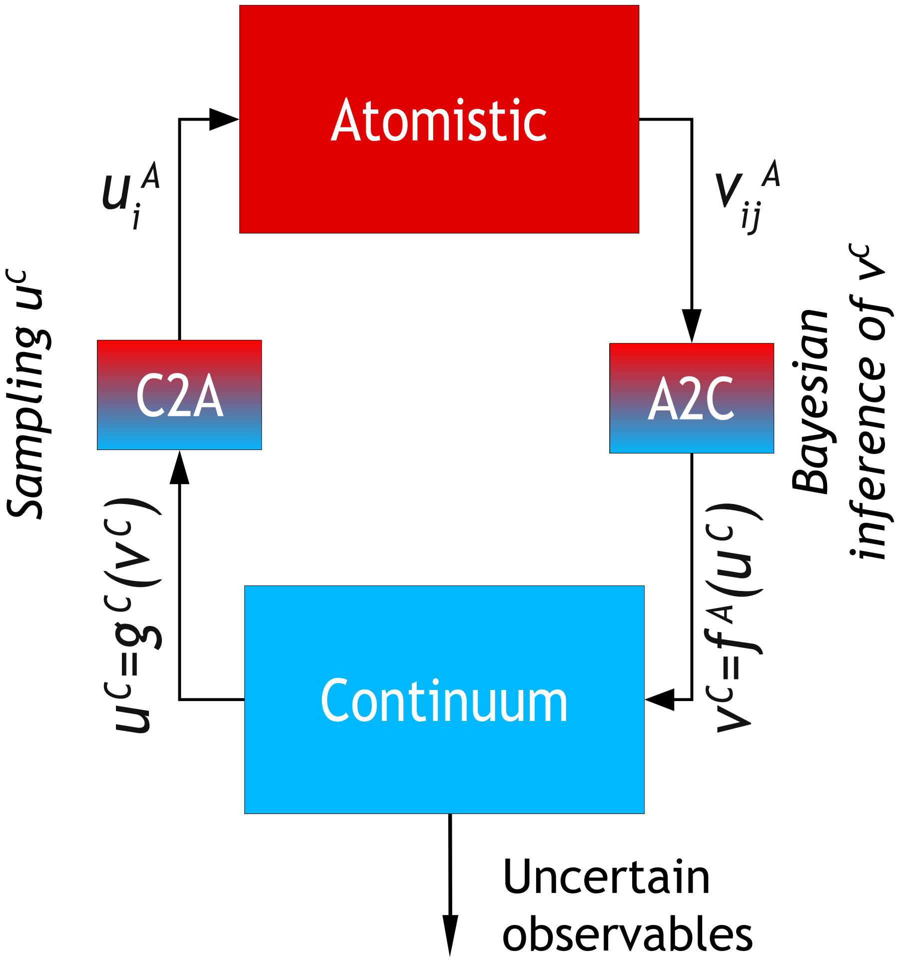
• This macroscale variable is imposed on the continuum model through stochastic coupling.

• We do not propagate the small scale fluctuations to the macroscale. We rather determine the uncertainty in the deterministic macroscale coupling variables due to sampling noise, and propagate this uncertainty.

Building blocks in atomistic to continuum coupling

MD sampling noise only

Fixed point iterations on the atomistic level (Salloum *et.al.*, 2012)



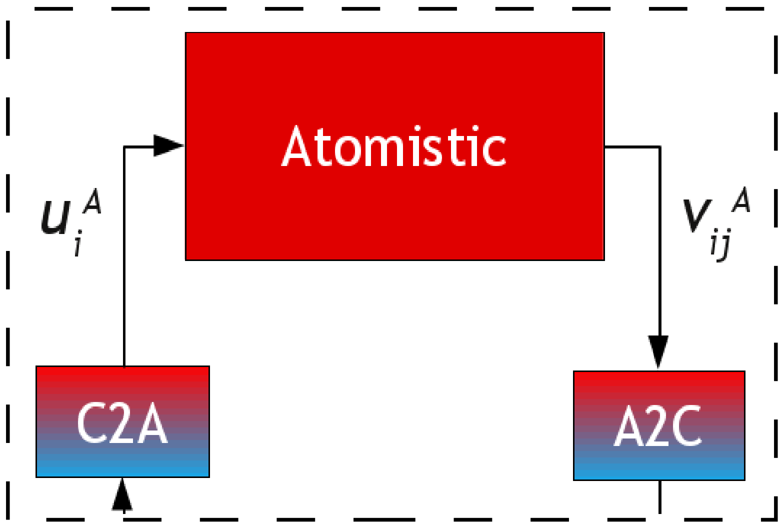
- This approach requires many atomistic simulations at nearby inputs, until convergence.
- Thus, the implementation is expensive unless a surrogate to the atomistic simulations is used.

Salloum, M., Sargsyan, K., Najm, H.N., Debusschere, B., Jones, R., Adalsteinsson, H. "A Stochastic Multiscale Coupling Scheme to account for Sampling Noise in Atomistic-to-Continuum Simulations" *SIAM Multiscale Modeling and Simulation*, 2012 (in press).

Building blocks in atomistic to continuum coupling

MD sampling noise and parametric uncertainty

Infer a response surface of the atomistic blocks



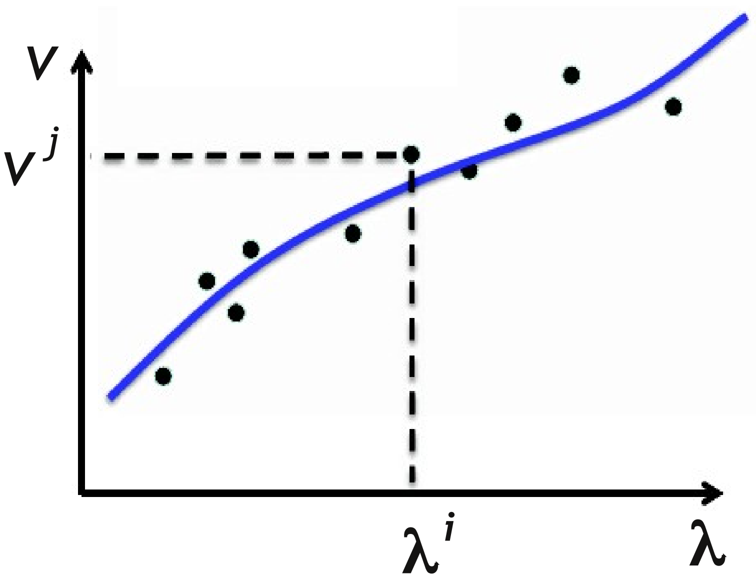
Group these blocks into one response surface that accounts for the uncertain parameters and the sampling noise:

$$v^C = f^A(u^C, p^A, \zeta)$$

$$u^C = g^C(v^C, p^C)$$

Solve these equations for the polynomial chaos expansions (PCE) u^c and v^c.

Bayesian Inference of Response Surfaces



$$v = \boldsymbol{p}^T \cdot \boldsymbol{a}$$

$$\boldsymbol{p} = \{p_q(\lambda_m)\}_{q=1}^{Q}$$

$$\boldsymbol{a} = \{a_q\}_{q=1}^{Q}$$

$$\boldsymbol{\lambda} = \{\lambda_m\}_{m=1}^{M}$$

- p_n are polynomials.
- a_q are the polynomials coefficients.
- λ_m can be either input variables or uncertain parameters.

$$\underbrace{\mathrm{P}(a_q, s^2 | \boldsymbol{d})}_{\text{Posterior}} \propto \underbrace{\mathrm{P}(\boldsymbol{d} | a_q, s^2)}_{\text{Likelihood}} \underbrace{\mathrm{P}(a_q, s^2)}_{\text{Prior}}$$

$$\boldsymbol{d} = \{v^j\}_{j=1}^{N}$$

- Gaussian likelihood

  $v^j = v(\boldsymbol{\lambda}^i) + s\,\eta^{ij}$

  $\eta^{ij} \sim \mathcal{N}(0,1)$
- v is linear in λ^i
- Infinite (improper) prior for a_q
- Jeffrey's prior for s^2

Analytical solution

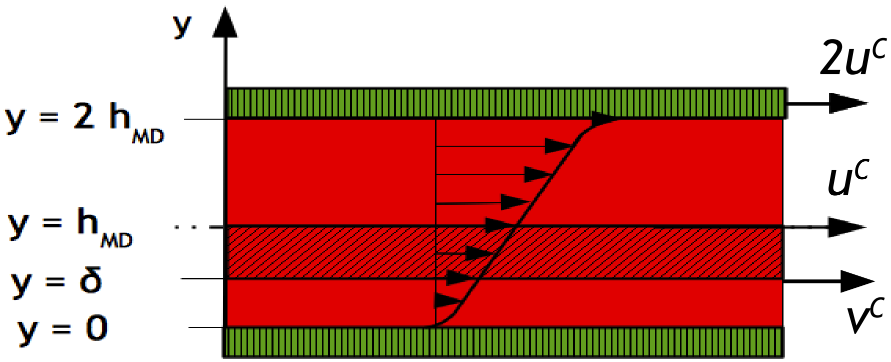
$$\boldsymbol{a} \sim St(\bar{\boldsymbol{a}}, S, \gamma)$$

$$v = \boldsymbol{p}(\lambda_m)^T \cdot \bar{\boldsymbol{a}} + \zeta \sqrt{\boldsymbol{p}(\lambda_m)^T \cdot S \cdot \boldsymbol{p}(\lambda_m)}$$

$$\zeta \sim St(0, 1, \gamma)$$

ζ accounts for the sampling noise

Response Surface of an Atomistic Couette Flow



Lennard-Jones interaction potential:

$$\phi = 4\,\phi_0 \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right]$$

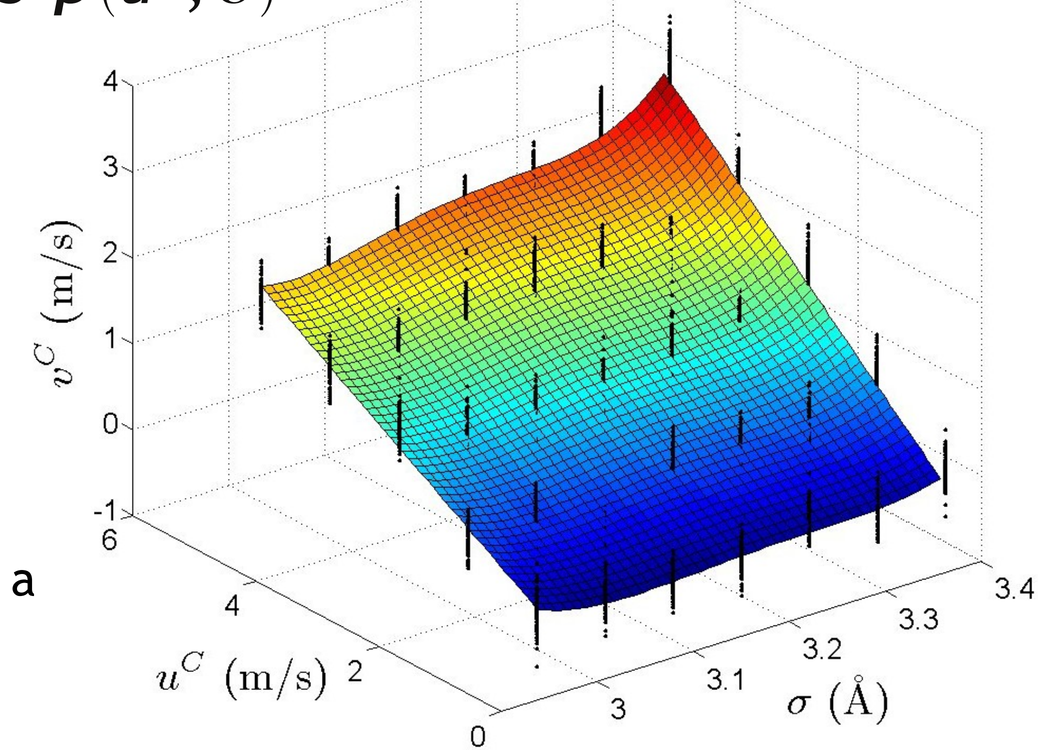
We assume that σ is an uncertain parameter in the atomistic simulation.

$$v^C = \boldsymbol{p}(u^C, \sigma)^T \cdot \bar{\boldsymbol{a}} + \zeta \sqrt{\boldsymbol{p}(u^C, \sigma)^T \cdot \boldsymbol{S} \cdot \boldsymbol{p}(u^C, \sigma)}$$

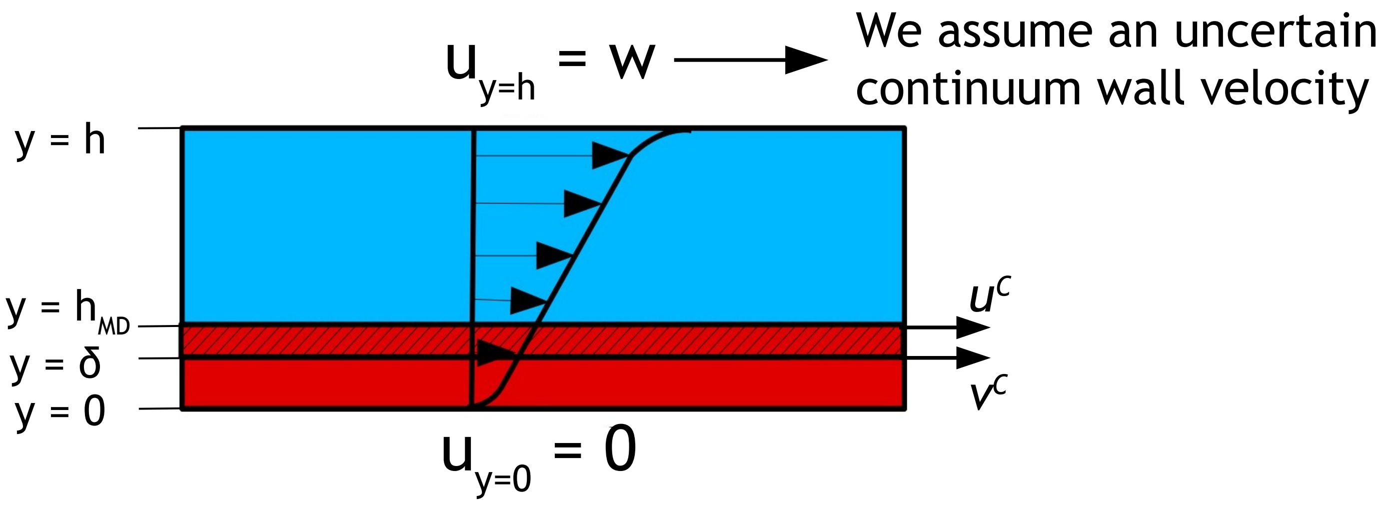
$$\boldsymbol{p} = \mathrm{O}(u^C, \sigma^4)$$

$$v^C = f^A(u^C, \sigma, \zeta)$$

- $\bar{\boldsymbol{a}}$ and $\boldsymbol{S}$ are inferred from MD data at sampled values of u^C and σ.
- The sampling noise is represented as a student-t process.



Response Surface of a Continuum Couette Flow



The steady state linear velocity profile allows analytical propagation of uncertainties.

$$u^C = w + \beta\left(w - v^C\right) \qquad \beta = \frac{h_{MD} - h}{h - \delta}$$

$$u^C = g^C\left(v^C, w\right)$$

Polynomial Chaos Expansions (PCE) are used in uncertainty quantification for an efficient representation of a random variable

Let X be a random variable with finite variance.

$$X : \Omega \rightarrow \mathbb{R}$$

$$X(\omega) = \sum_{k=0}^{\infty} X^k \Psi^k (\xi_{1,} \xi_{2.} .. \xi_{N_d})$$

$\{\xi_i\}_{i=1}^{N_d}$ are *i.i.d.* random variables (e.g., Gaussian)

$\{\Psi^k\}_{k=0}^{\infty}$ are multivariate orthogonal polynomials (e.g., Hermite)

We truncate the expansion at order N_o and dimension N_d such that:

$$X = \sum_{k=0}^{P} X^k \Psi^k (\boldsymbol{\xi})$$

$$P + 1 = \frac{(N_d + N_o)!}{N_o! N_d!}$$

Methods to Operate on PCEs

e.g. Product of two PCEs

$$\left.\begin{aligned} X &= \sum_{k=0}^{P} X^k \Psi^k(\xi) \\ Y &= \sum_{k=0}^{P} Y^k \Psi^k(\xi) \end{aligned}\right\} \longrightarrow \text{Find } Z = X\,.\,Y$$

1. Direct (Intrusive)

Z^k are obtained by Galerkin projection

$$Z = \sum_{k=0}^{P} Z^k \Psi^k(\xi) \qquad Z^k = \frac{1}{\langle \Psi^k \Psi^k \rangle} \sum_{i=0}^{P} \sum_{j=0}^{P} X^k Y^k \langle \Psi^i \Psi^j \Psi^k \rangle$$

2. Sampling (Non-Intrusive)

Get samples X_i and Y_i from quadrature points, compute $Z_i = X_i . Y_i$, then project the Z_i on the PC basis Ψ^k.

Other mathematical operations such as divisions, log, exp, square root... can also be performed on PCEs in a similar fashion.
http://www.sandia.gov/UQToolkit/

Intersecting Response Surfaces with Fixed Point Iterations

$$u^C = \sum_{k=0}^{P} u^{C,k} \Psi^k(\xi_1, \xi_2, \xi_3)$$

$$v^C = \sum_{k=0}^{P} v^{C,k} \Psi^k(\xi_1, \xi_2, \xi_3)$$

$$\sigma \rightarrow \xi_1$$
$$w \rightarrow \xi_2$$
$$\zeta \rightarrow \xi_3$$

$$v^C = f^A(u^C, \sigma, \zeta)$$

$$u^C = g^C(v^C, w)$$

- Assume known uncertainties in *w* and *σ*
- Substitute PC expansions into response surfaces
- Start with an initial guess of either u^c or v^c
- Iterate on u^c and v^c: two well-known approaches:

- Perform Galerkin operations on the expansions of u^c and v^c: *intrusive spectral projection* **(ISP)**

- Sample *ζ*, *w* and *σ* on quadrature points, solve for u^c and v^c then project on the PC basis $\Psi^k(\xi_1,\xi_2,\xi_3)$: *non-intrusive spectral projection* **(NISP)**

Solution method
1. Intrusive Spectral Projection (ISP)

$$v^C = \boldsymbol{p}(u^C, \sigma)^T \cdot \bar{\boldsymbol{a}} + \zeta \sqrt{\boldsymbol{p}(u^C, \sigma)^T \cdot S \cdot \boldsymbol{p}(u^C, \sigma)}$$

$$u^C = w + \beta(w - v^C)$$

$$u^C = \sum_{k=0}^{P} u^{C,k} \Psi^k(\xi_1, \xi_2, \xi_3)$$

$$v^C = \sum_{k=0}^{P} v^{C,k} \Psi^k(\xi_1, \xi_2, \xi_3)$$

$$\sigma = \sigma^0 + \sigma^1 \xi_1$$

$$w = w^0 + w^1 \xi_2$$

$$\zeta = \xi_3$$

- $\bar{\boldsymbol{a}}$, $\boldsymbol{S}$, β and the polynomial coefficients vector $\boldsymbol{p}$ are known entities.
- The PCEs of ζ, w and σ are known.
- Start with an initial guess of the $u^{C,k}$ or $v^{C,k}$
- Different Galerkin operations take place at each iteration:
  - Multiple products of PCEs depending on the order of the polynomial p.
  - Square root of a PCE.

Solution method
2. Non-Intrusive Spectral Projection (NISP)

- Sample the PCEs of *ζ*, *w* and *σ* on quadrature points.
- For each sample $(\zeta, w, \sigma)_i$, solve:

$$v_i^C = \boldsymbol{p}(u_i^C, \sigma_i)^T \cdot \bar{\boldsymbol{a}} + \zeta_i \sqrt{\boldsymbol{p}(u_i^C, \sigma_i)^T \cdot S \cdot \boldsymbol{p}(u_i^C, \sigma_i)}$$

$$u_i^C = w_i + \beta(w_i - v_i^C)$$

- Project the obtained $(u^C, v^C)_i$ on the PC basis $\Psi^k(\xi_1,\xi_2,\xi_3)$

$$(v^{C,k}, u^{C,k}) = \frac{\langle (v^C, u^C)_i \Psi^k \rangle}{\langle \Psi^k \Psi^k \rangle}$$

$$v^C = \sum_{k=0}^{P} v^{C,k} \Psi^k(\xi_1, \xi_2, \xi_3) \qquad u^C = \sum_{k=0}^{P} u^{C,k} \Psi^k(\xi_1, \xi_2, \xi_3)$$

Method Validation: Case of no Parametric Uncertainty

$$u^C = u^{C,0} + u^{C,1}\xi_3$$

$$u^C = u^{C,0} + u^{C,1}\xi_3$$

$$\zeta \rightarrow \xi_3$$

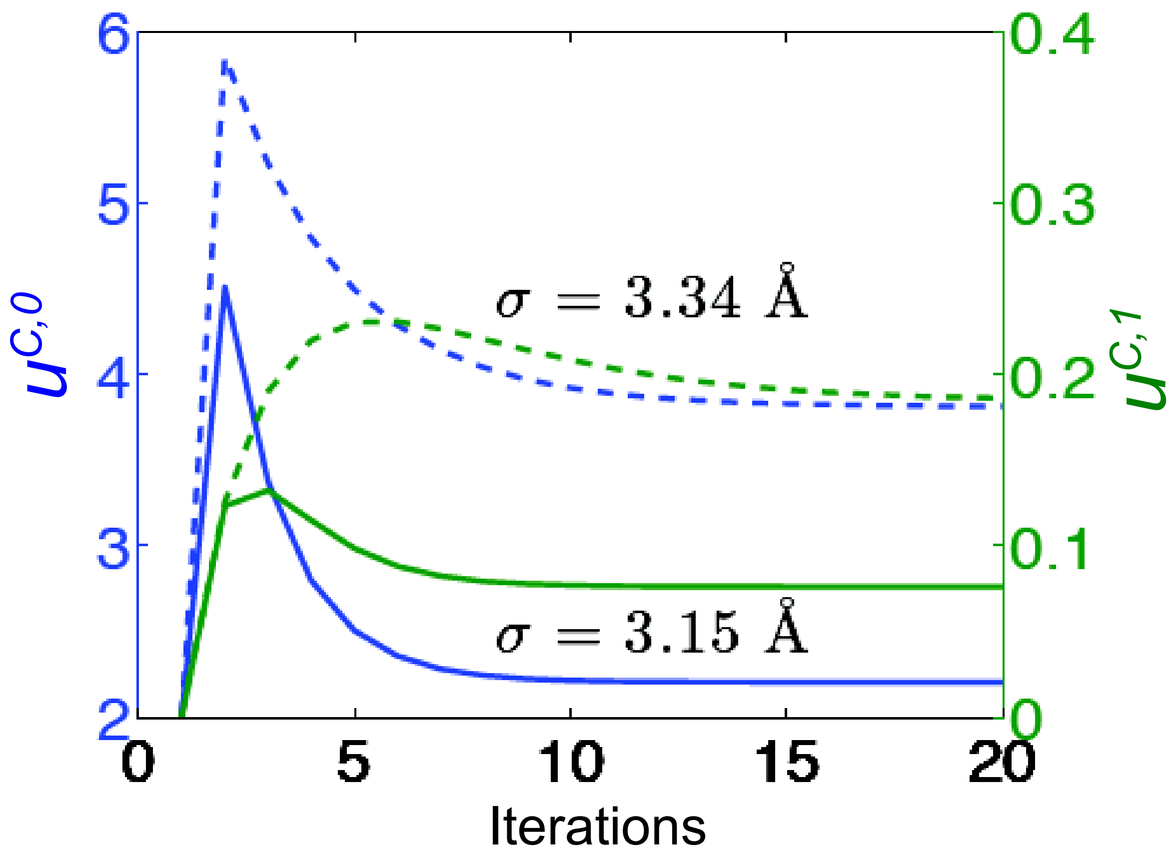
$$v^C = f^A(u^C, \zeta)$$

$$u^C = g^C(v^C)$$

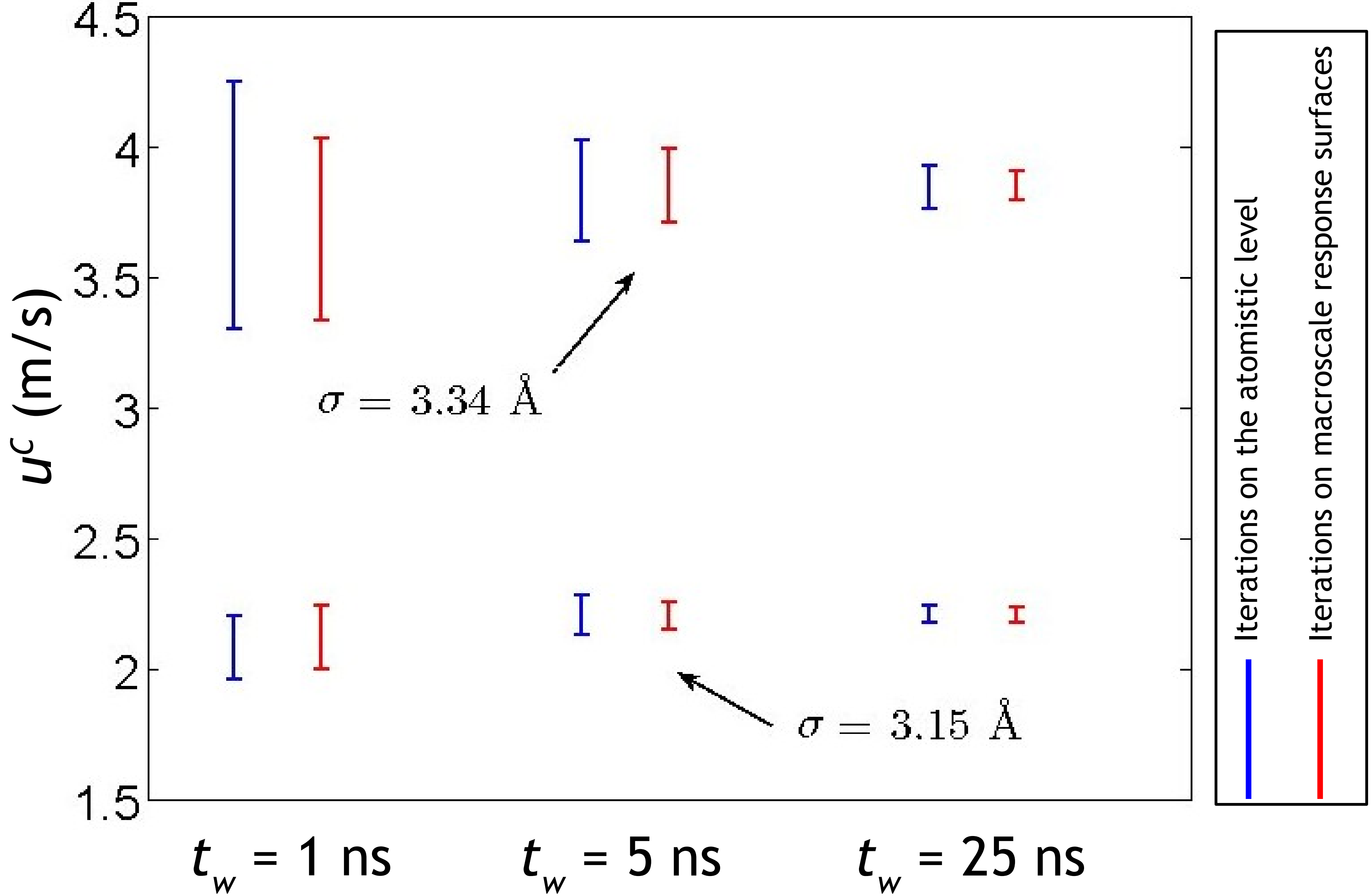
$$\zeta \sim St(0,1,\gamma) \underbrace{\sim N(0,1)}_{\text{for } \gamma > 30}$$

γ is related to the amount of data used to infer f^A.

In this study we have $\gamma = 190$



Method Validation: Case of no Parametric Uncertainty



Results: Case of Parametric Uncertainty

$$\sigma = 3.15 + 0.074\,\xi_1$$
$$w = 20 + \xi_2$$
$$\zeta = \xi_3 \left(\gamma = 190, t_w = 5ns\right)$$

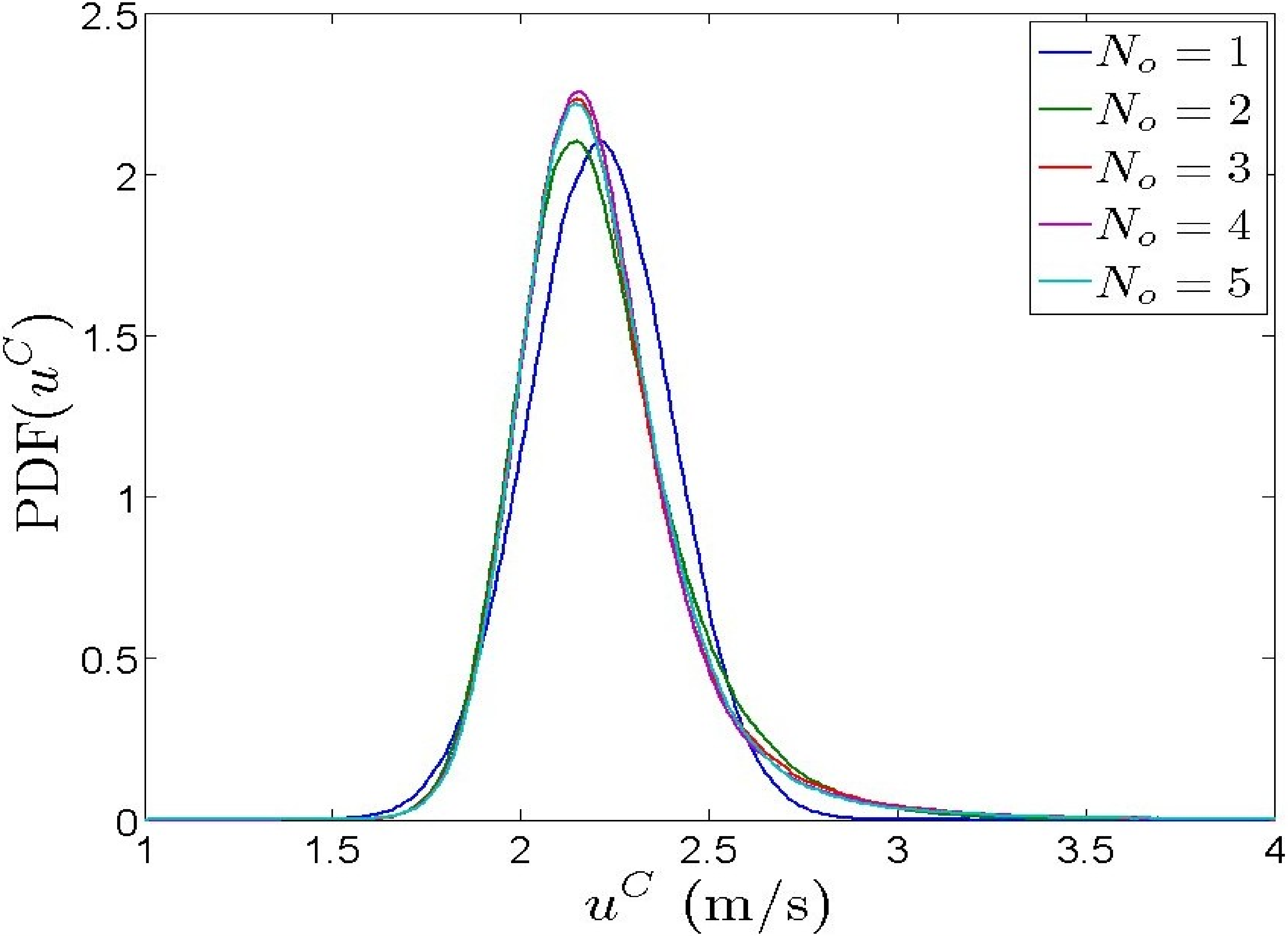
$$v^C = \sum_{k=0}^{P} v^{C,k} \Psi^k(\xi_{1,} \xi_{2,} \xi_3) \qquad u^C = \sum_{k=0}^{P} u^{C,k} \Psi^k(\xi_{1,} \xi_{2,} \xi_3)$$

We assume that all ξ_1,ξ_2 and ξ_3 follow Gaussian distributions i.e., $\Psi^k(\xi_1,\xi_2,\xi_3)$ are Hermite polynomials truncated at order N_o

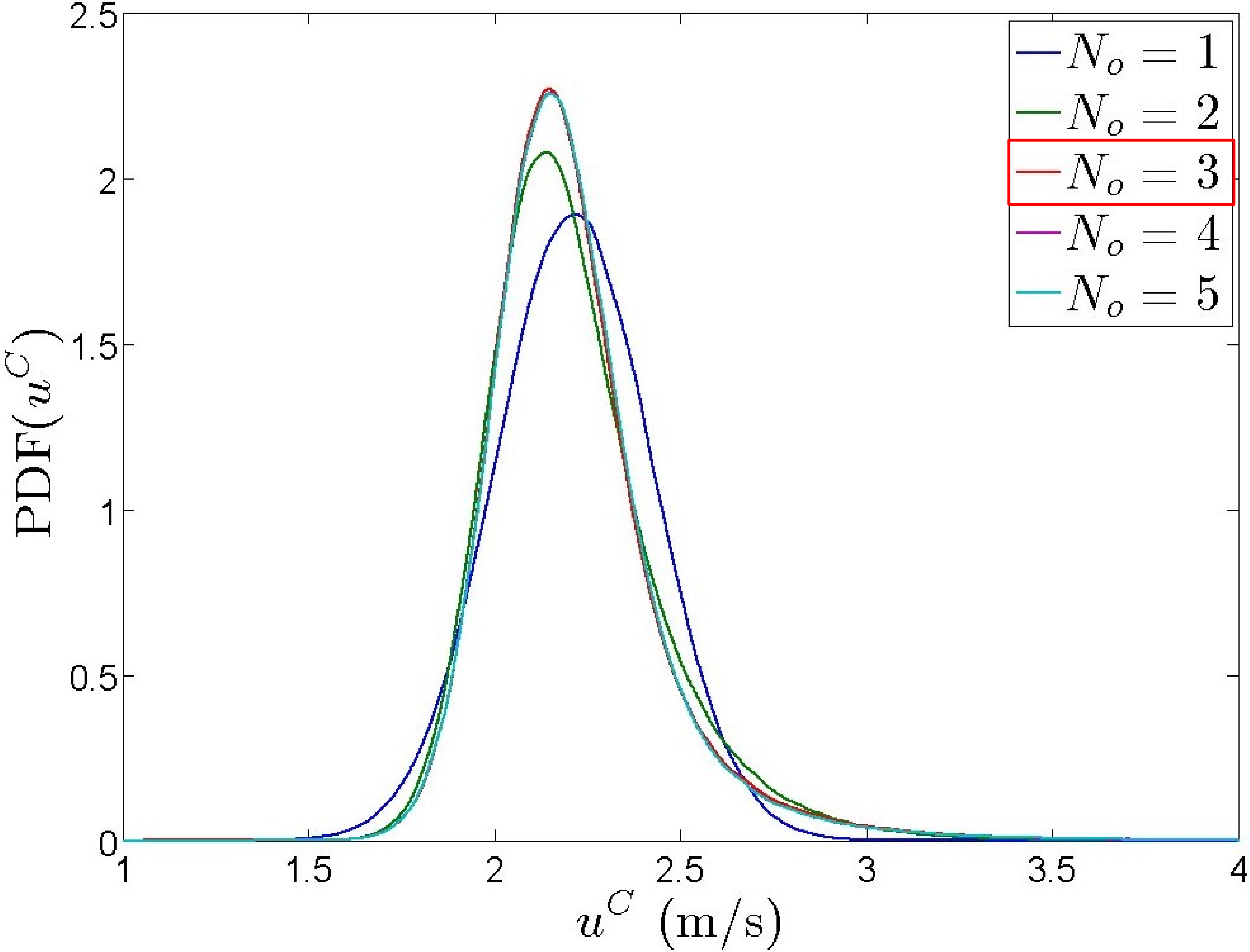
$$v^C = \boldsymbol{p}(u^C, \sigma)^T \cdot \bar{\boldsymbol{a}} + \zeta \sqrt{\boldsymbol{p}(u^C, \sigma)^T \cdot S \cdot \boldsymbol{p}(u^C, \sigma)}$$
$$u^C = w + \beta(w - v^C)$$

p is linear in u^c and 4th order in σ. Thus, we expect u^c *and* v^c to have a linear dependence on w and a 4th order dependence on σ.

Results: ISP



Results: NISP



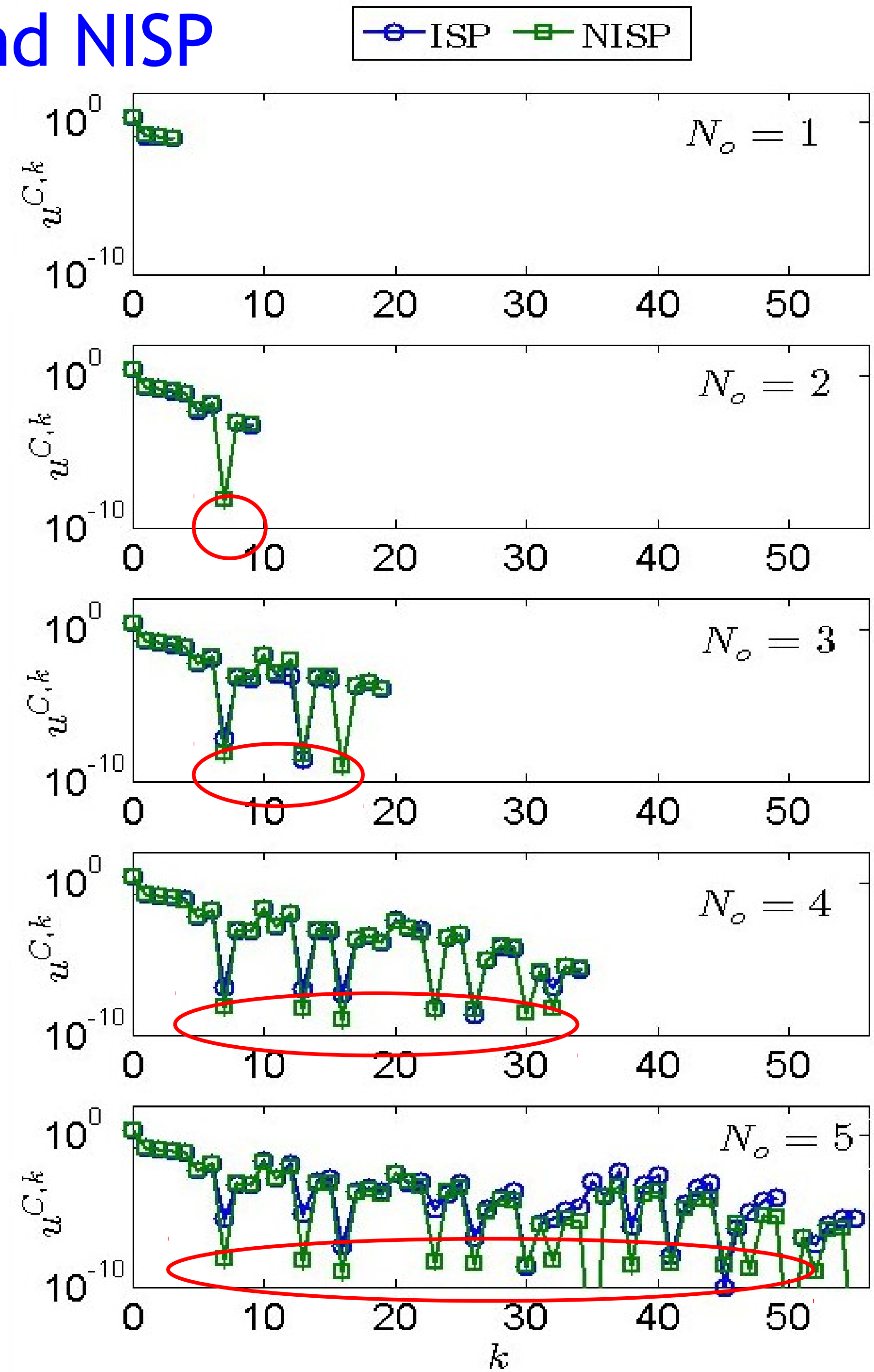
Comparison between ISP and NISP

• The ISP and NISP approaches are in a very good agreement.

• The modes with orders higher than one in *w* have a negligible amplitude and can be reduced from the PCE.

• The ISP approach is much more expensive due to the computational overhead caused by the Galerkin operations.

• Thanks to the response surface representation of the atomistic model, deterministic solutions of the obtained system are cheap, making NISP much more attractive in terms of performance.



Contribution of different sources of uncertainty

• We perform a global sensitivity analysis to quantify the contribution of the uncertain parameters (w and σ) and the sampling noise (ζ) to the uncertainty in the macroscale variables u^c and v^c.

• We compute the total sensitivity indices* of u^c and v^c from their PC representations (N_o=3):

	σ	w	ζ
u^c	0.606	0.254	0.140
v^c	0.745	0.083	0.172

$$\sigma = 3.15 + 0.074\,\xi_1$$
$$w = 20 + \xi_2$$
$$\zeta = \xi_3(\gamma = 190, t_w = 5ns)$$

• The uncertain Lennard-Jones parameter σ of the atomistic simulation contributes the most to the uncertainty in u^c and v^c.

* Le Maitre, O., Knio, O. Spectral Methods for Uncertainty Quantification with Applications to Computational Fluid Dynamics. Springer, Berlin (2010)

Conclusions and Ongoing Work

- We showed a systematic approach to infer response surfaces that account for both finite sampling noise and parametric uncertainty in atomistic simulations.
- This response surfaces approach allows coupling on the macroscale level and uncertainty quantification using either intrusive or non-intrusive methods.
- We found that for the given range of uncertainty in the parameters and the sampling noise, the uncertain Lennard-Jones parameter of the atomistic simulation is the dominant source of uncertainty.
- The Couette flow used here allows for analytical solutions, but the formulation is generally applicable.

- *The application to nanopore ionic fluxes is in progress.*

THANK YOU FOR YOUR ATTENTION

Questions???