Hybrid Uncertainty Quantification Methods for Reactive Transport

Charles Tong, Xiao Chen, Brenda Ng, Yunwei Sun

Lawrence Livermore National Laboratory

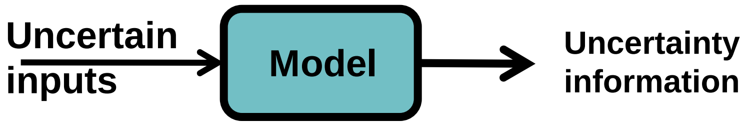
Gianluca Iaccarino, Stanford University

Barry Lee, Pacific Northwest National Laboratory

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344, LLNL-PRES-503294-DRAFT

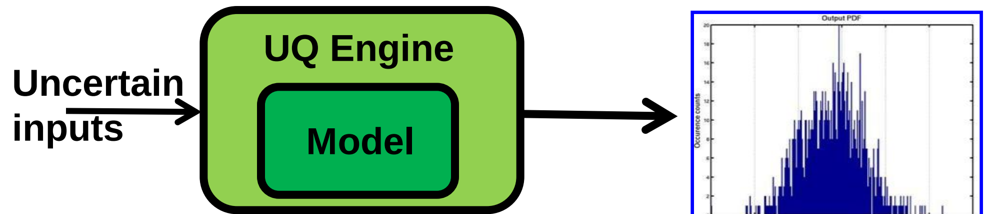
What do we mean by hybrid uncertainty quantification?

Purely Intrusive approach



Stochastic simulation
(UQ e mbedded in the model)

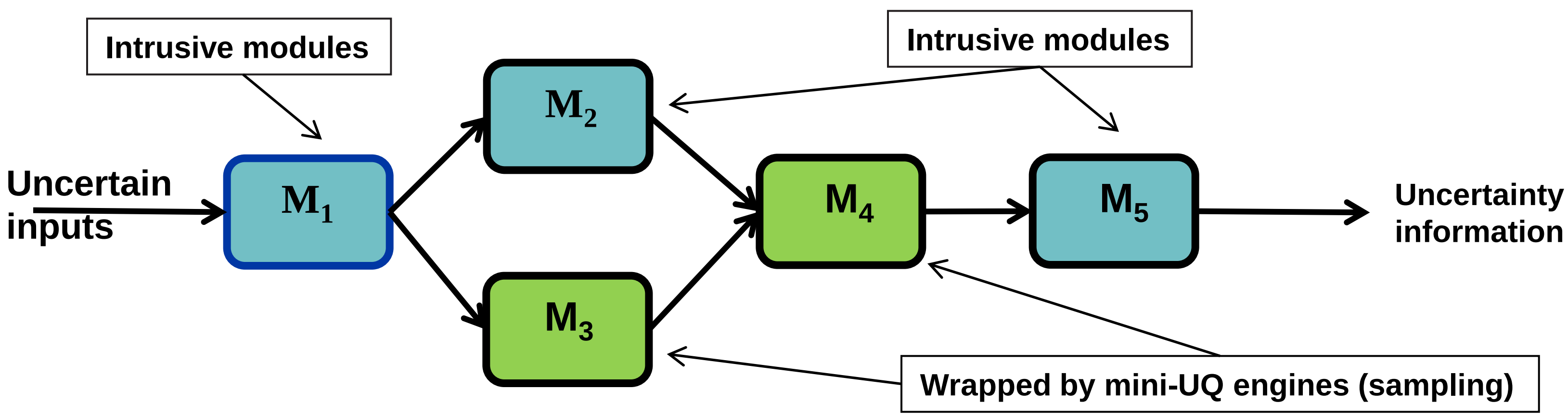
Non-intrusive approach



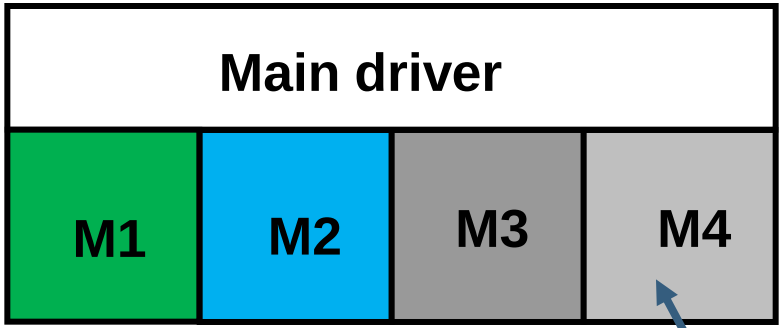
Many deterministic simulations

hybrid approach for multi-physics

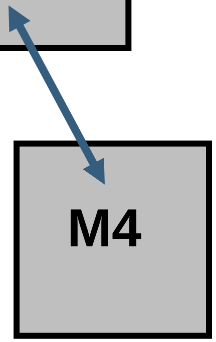
- decompose the system into components with each having its own UQ
- create algorithms to use these modules to propagate global uncertainties
- in this project, we consider decomposition along physics boundaries



Why Hybrid UQ?



- current model development practice
  - “plug-and-play” gives flexibility
  - operator splitting widely used
  - well-defined module interfaces
  - extensive use of open source or commercial codes as modules



- Likely scenario in multi-physics models if UQ is embedded
  - Intrusive UQ may be readily available for some modules
  - Only non-intrusive UQ may be available for other modules
  - With time, advanced UQ may be available for some modules
  - Also, new uncertain parameters may be added from time to time
  - Thus, it is reasonable to consider UQ at module level
  - Yet, global uncertainties/sensitivities must be propagated
  - Is “plug-and-play” feasible with embedded UQ? If so, how?
    - The objective of this project is to explore this.

An Example: hybrid UQ for Reaction Diffusion Equation

- **Consider a 1D reaction-diffusion equation (D and K are second order random variables (rv))**

$$\frac{\partial C}{\partial t} = D\frac{\partial^2 C}{\partial x^2} - KC$$

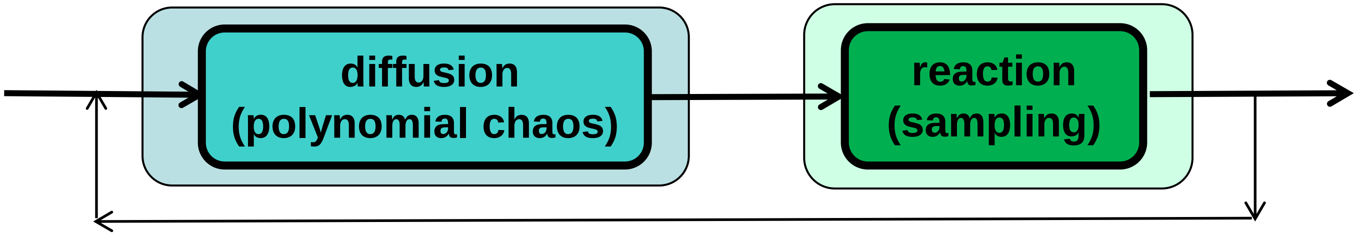
$$0 \le x \le 1; 0 \le t \le T; u(0,t) = u(1,t) = 0; u(x,0) = f(x)$$

- **Operator splitting**

$$C^{n+1/2} = F_T(C^n, D)$$
$$C^{n+1} = F_R(C^{n+1/2}, K)$$

- **Suppose we have the following scenario:**
  - **intrusive polynomial chaos is available for the diffusion equation**
  - **only non-intrusive methods are available for the reaction equation**
  - **this scenario may be realistic, especially for multi-species systems**
    - **the reaction system may have an analytic solution**
    - **the reaction equation may benefit from a fast ODE solver**
    - **in some cases, one module may be open source or commercial code**

Goal: build a global uncertainty propagation algorithm via "gluing" an intrusive and a sampling module



Objective: design a forward "global" uncertainty propagation algorithm

- which permits 2 different UQ methods in the two modules, and
- which permits each module to be developed independently of the other
- given that the base uncertainty format is polynomial chaos coefficients
- Thus, at each time time, the following are performed by each module (which can be captured by a generic software layer):

Intrusive module:
- convert from base to local format
- run local module
- analyze results
- convert from local to base format

Non-intrusive module:
- convert from base to local format
- create appropriate sampling schemes
- run samples
- analyze results
- convert from local to base format

Polynomial Chaos for the Diffusion Equation

- **Expand independent and dependent variables in PCE**

$$D = \sum_{i=0}^{P} D_i \Psi_i(\xi); \quad C = \sum_{i=0}^{P} D_i \Psi_i(\xi)$$

- **Substitute into the transport equation**

$$\frac{\partial}{\partial t} \sum_{i=0}^{P} C_i \Psi_i(\xi) = \sum_{i=0}^{P} D_i \Psi_i(\xi) \frac{\partial^2}{\partial x^2} \sum_{j=0}^{P} C_j \Psi_j(\xi)$$

- **Project onto each of the P+1 polynomial bases**

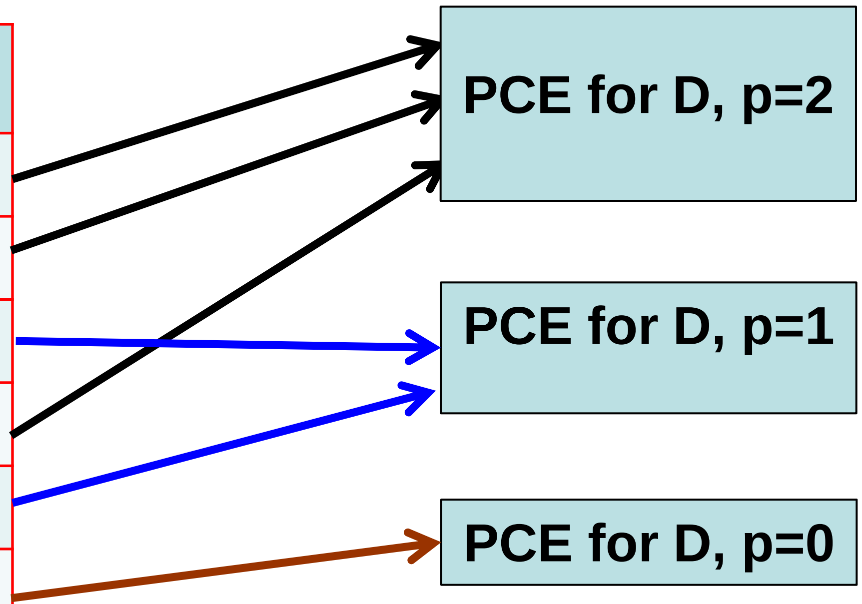
$$\frac{\partial C_k}{\partial t} = \frac{1}{<\Psi_k>^2} \sum_{j=0}^{P} \left[\left(\sum_{i=0}^{P} D_i e_{ijk} \right) \frac{\partial^2 C_j}{\partial x^2} \right]; \quad e_{ijk} = <\Psi_i \Psi_j \Psi_k>, k = 0, \cdots, P$$

- **apply initial/boundary conditions and then solve the equations**
- **Q: how to propagate global uncertainties locally?**

Propagation of Uncertainties through the Diffusion Module

- **Goal: plug the stand-alone PCE-based dffusion module into the overall global uncertainty propagation algorithm**
- **Idea: for linear diffusion equations, the global uncertainty stream can be broken up into pieces that can be handled naturally**
- **An example, to break up the 2nd order PCE for 2 variables (D, K)**

D order	K order
0	0
1	0
0	1
2	0
1	1
0	2

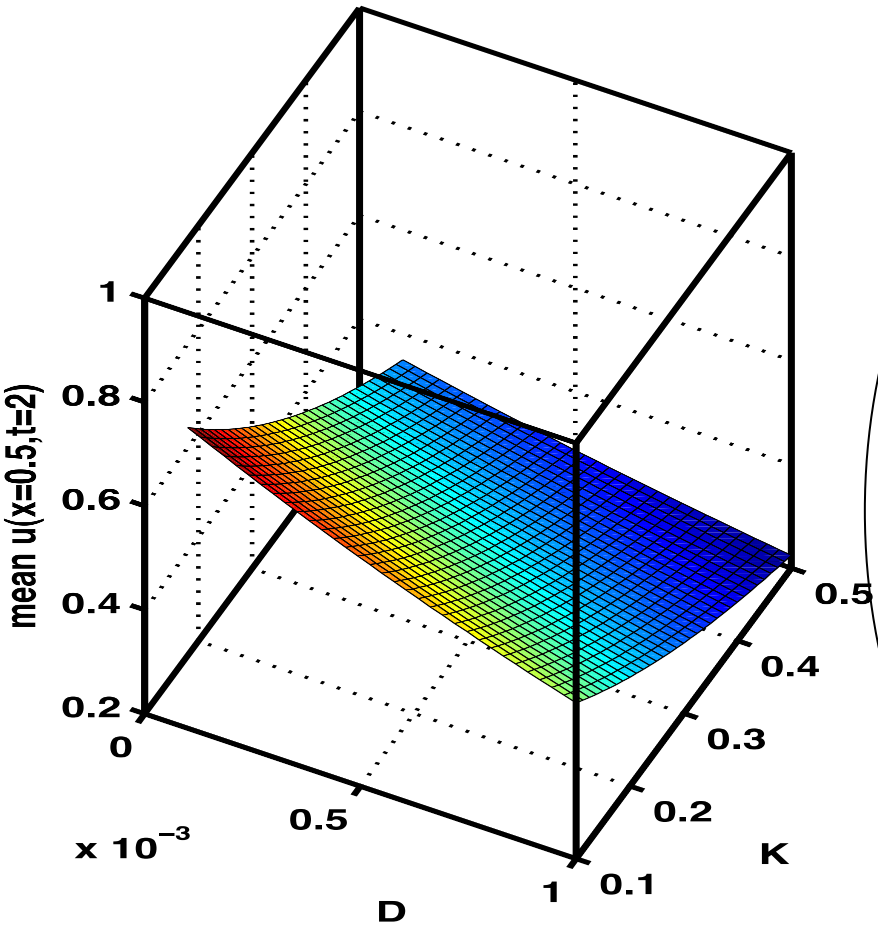


- **Intuitively, the global matrix can be decomposed into subproblems.**

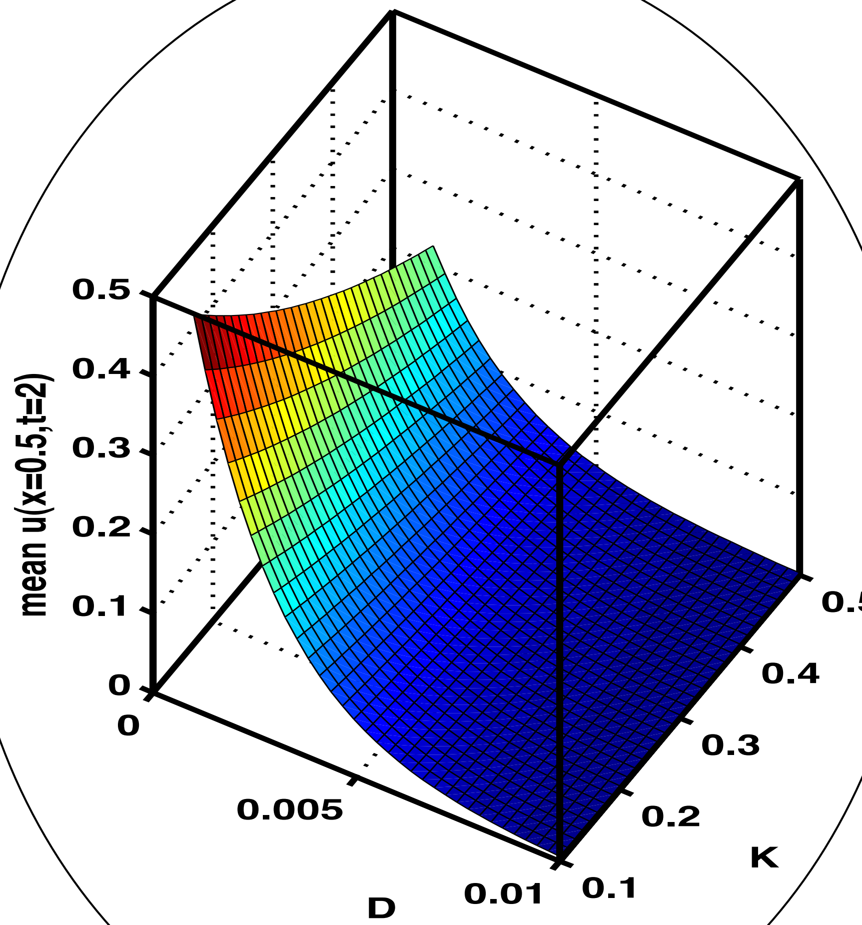
Non-intrusive Method for the Reaction Equation

- $C^{n+1/2}$ coming into the reaction module is a 2-rv PC expansion
- This needs to be converted to initial conditions for reaction
- How is this transformation done?
  - reconstruct the Legendre polynomials (2-parameter)
  - create a sample (2-parameter) for the reaction equation (size N)
  - evaluate the Legendre polynomials at the sample points
- Now we have N initial conditions for the N sample points
- Next, run the reaction solver on each sample point
- How to convert from the sample outputs back to the PC format?
  - the outputs will be fed into a Legendre regression analyzer
  - use either projection or least-squares approach
  - the results will be the PC coefficients
  - beware of scaling issues
- Sampling and transformation can be handled by a software API

Response Surfaces for different diffusion ranges

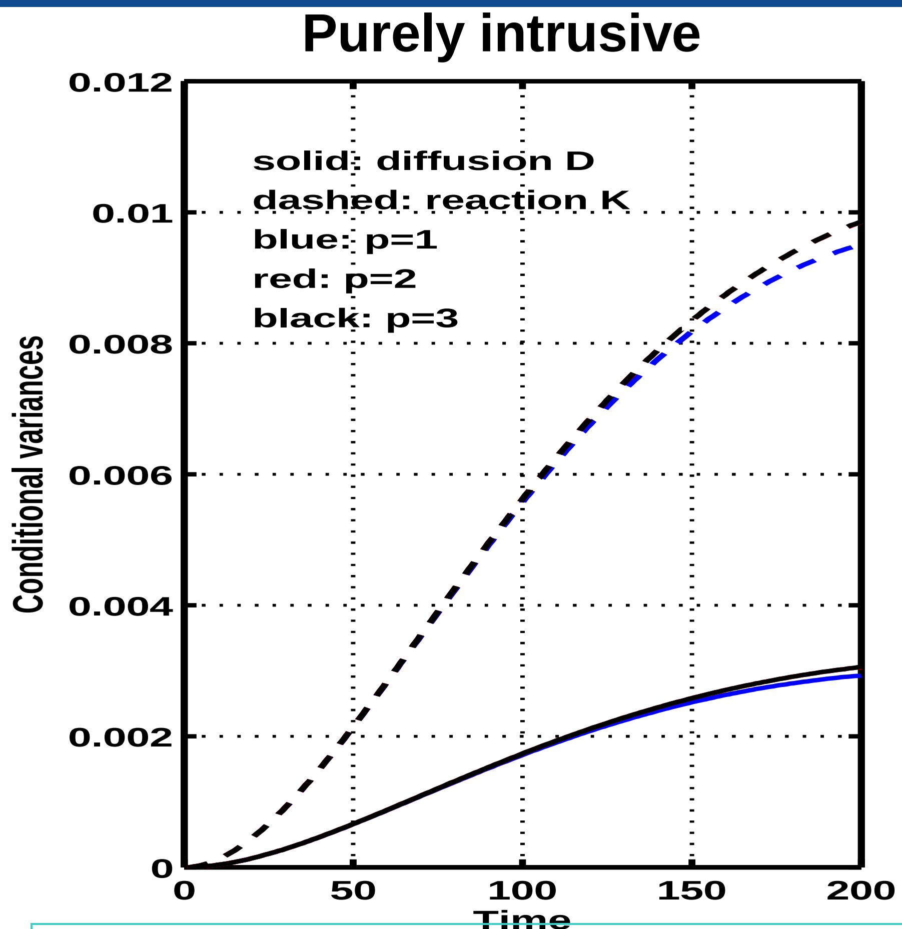


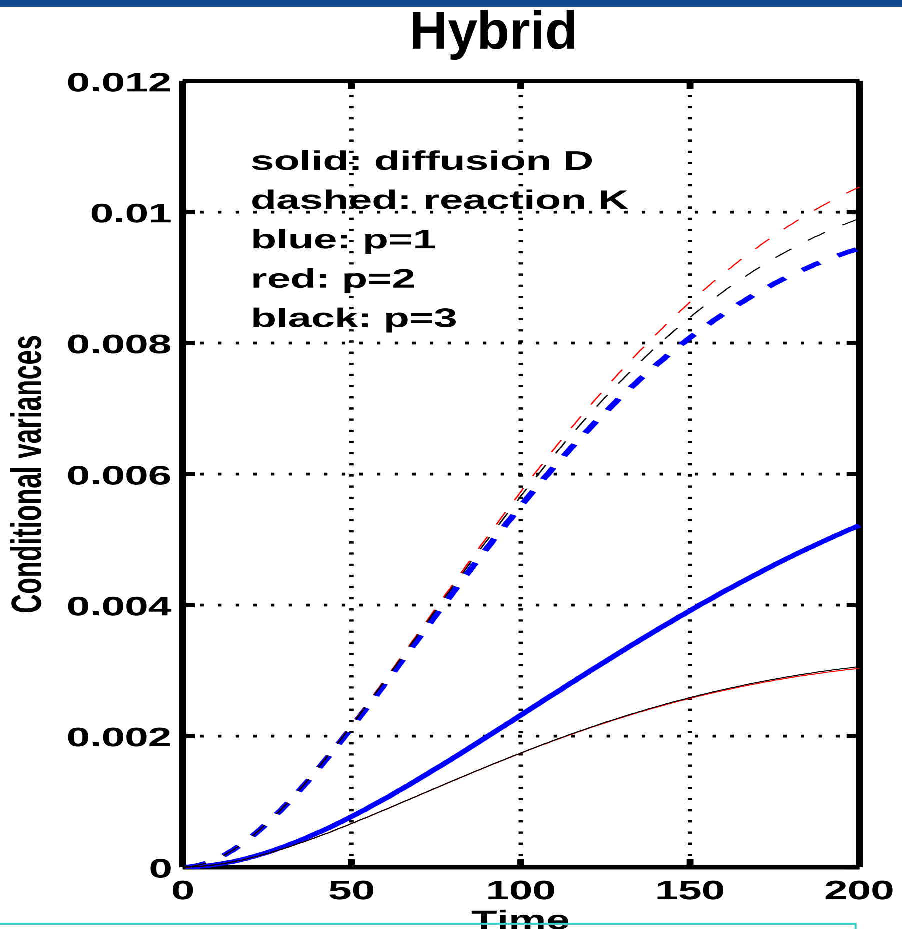
D=[0.0001,0.001], K=[0.1, 0.5]



D=[0.001,0.01], K=[0.1, 0.5]

Numerical Results (D = [0.001, 0.01]; k = [0.1, 0.5])





- Purpose: investigate whether the hybrid method gives accurate answers
- Observations: convergence of hybrid at p=3, pure intrusive at p=2
- Hybrid gives more accurate answers due to the use of analytic reaction
- In practical cases, hybrid may be needed out of necessity

Another Test Problem: multi-species reactive transport

Sequential Reactions

$$
\begin{cases}
A_1 + C_1 \overset{k_1'}{\Rightarrow} y_1 C_2 \\
A_2 + C_2 \overset{k_2'}{\Rightarrow} y_2 C_3 \\
\cdots \\
A_i + C_i \overset{k_i'}{\Rightarrow} y_i C_{i+1} \\
\cdots \\
A_n + C_n \overset{k_n'}{\Rightarrow} y_n C_{n+1}
\end{cases}
$$

Reactive Transport

$$
\begin{cases}
\dfrac{\partial c_1}{\partial t} = L(c_i) - k_1 c_1 \\
\dfrac{\partial c_2}{\partial t} = L(c_2) - k_2 c_2 + y_1 k_1 c_1 \\
\cdots \\
\dfrac{\partial c_i}{\partial t} = L(c_3) - k_i c_i + y_{i-1} k_{i-1} c_{i-1} \\
\cdots \\
\dfrac{\partial c_n}{\partial t} = L(c_n) - k_n c_n + y_{n-1} k_{n-1} c_{n-1}
\end{cases}
$$

A_i* is a reactant participating in reaction *i
C_i* is product of reaction *i-1
y_i* is the yield coefficient of reaction *i
k_i* is the reaction rate constant in reaction *i

<u>Uncertain Paramters</u>

-diffusion, velocity (2 rv's in *L*)

-reaction rates (4 rv's, n=4)

Example reactions in subsurface flow: TCE → DCE → VC → ETH

Stochastic Transport System

Let c(x,t) be species concentration defined on a bounded domain Ω x [0,T].
Let $\mathscr{V} = \left\{ c \in H_0^1(\Omega) : \; c = c_d \text{ on } \Gamma_d \right\}$ ***and nonhomogeneous boundary*** Γ_d

Consider now α, v^x , and v^y are functions of random event θ of an abstract probability space (Θ, Σ, P):

$$\alpha = \alpha(\theta), \;\; v^x = v^x(\theta), \;\; v^y = v^y(\theta)$$

Find

$$c(\mathbf{x}, \theta, t) \in \mathscr{V} \otimes L_2(\Theta, P)$$

such that it satisfies almost surely the stochastic problem:

$$\begin{cases} \frac{\partial c(\mathbf{x},\theta,t)}{\partial t} = \frac{\alpha(\theta) v^x}{R} \frac{\partial^2 c(\mathbf{x},\theta,t)}{\partial x^2} + \frac{\alpha(\theta) v^y}{R} \frac{\partial^2 c(\mathbf{x},\theta,t)}{\partial y^2} \\ - \frac{v^x(\theta)}{R} \frac{\partial c(\mathbf{x},\theta,t)}{\partial x} - \frac{v^y(\theta)}{R} \frac{\partial c(\mathbf{x},\theta,t)}{\partial y} & \mathbf{x} \in \Omega \\ c(\mathbf{x}, \theta, t) = c_d(\mathbf{x}, \theta) & \mathbf{x} \in \Gamma_d \end{cases}$$

α : dispersivity; v^x, v^y : velocity in x, y direction; R : retardation factor

Stochastic Variational Form for theTransport System

Find $c(\mathrm{x}, \theta, t) \in \mathscr{V} \otimes L_2(\Theta, P)$ such that

$$A(c, w) = \mathrm{E}[a(c, w)] \;\; \forall w(\mathrm{x}, \theta) \in \mathscr{V} \otimes L_2(\Theta, P)$$

where

$$\begin{aligned}
A(c, w) \;=\;& \int_\Theta w(\mathrm{x}, \theta) \left[\int_\Omega \frac{\partial c(\mathrm{x}, \theta, t)}{\partial t} \mathrm{dx} \right] \mathrm{d}P(\theta) \\
+\;& \frac{1}{R} \int_\Theta \alpha(\theta)\, v^x(\theta) \left(\int_\Omega \frac{\partial w(\mathrm{x}, \theta)}{\partial x} \frac{\partial c(\mathrm{x}, \theta, t)}{\partial x} \mathrm{dx} \right) \mathrm{d}P(\theta) \\
+\;& \frac{1}{R} \int_\Theta \alpha(\theta)\, v^y(\theta) \left(\int_\Omega \frac{\partial w(\mathrm{x}, \theta)}{\partial y} \frac{\partial c(\mathrm{x}, \theta, t)}{\partial y} \mathrm{dx} \right) \mathrm{d}P(\theta) \\
-\;& \frac{1}{R} \int_\Theta v^x(\theta) \left(\int_\Omega w(\mathrm{x}, \theta) \frac{\partial c(\mathrm{x}, \theta, t)}{\partial x} \mathrm{dx} \right) \mathrm{d}P(\theta) \\
-\;& \frac{1}{R} \int_\Theta v^y(\theta) \left(\int_\Omega w(\mathrm{x}, \theta) \frac{\partial c(\mathrm{x}, \theta, t)}{\partial y} \mathrm{dx} \right) \mathrm{d}P(\theta)
\end{aligned}$$

Stochastic Discretization

Let $\mathcal{N}$ be the set of nodes of the finite-element mesh not lying on Γ_d and $\Phi_i(\mathrm{x})$ be the corresponding shape function

Let α, v^x, v^y be independent second-order random variables in

$$\mathscr{W}^P \equiv \text{span}\{\Psi_0, \ldots, \Psi_P\} \subset \mathscr{W} \equiv L_2(\Theta, P)$$

Using $(P+1)$-term polynomial chaos expansion, one obtains

$$\alpha(\boldsymbol{\xi}, \theta) = \sum_{i=0}^{P} \alpha_i \Psi_i(\boldsymbol{\xi}(\theta))$$

$$v^x(\boldsymbol{\xi}, \theta) = \sum_{i=0}^{P} v_i^x \Psi_i(\boldsymbol{\xi}(\theta)) \quad v^y(\boldsymbol{\xi}, \theta) = \sum_{i=0}^{P} v_i^y \Psi_i(\boldsymbol{\xi}(\theta))$$

$$\alpha_i = \frac{< \alpha(\boldsymbol{\xi}), \Psi_i >}{< \Psi_i^2 >} \quad v_i^x = \frac{< v^x(\boldsymbol{\xi}), \Psi_i >}{< \Psi_i^2 >} \quad v_i^y = \frac{< v^y(\boldsymbol{\xi}), \Psi_i >}{< \Psi_i^2 >}$$

$$c^h(\mathrm{x}, \boldsymbol{\xi}, t) = \sum_{i \in \mathcal{N}} \left(\sum_{k=0}^{P} c_{i,k}(t) \Psi_k(\boldsymbol{\xi}) \right) \Phi_i(\mathrm{x}) \in \left(\mathscr{V}^h \otimes \mathscr{W}^P \right)$$

Stochastic Discretization and Galerkin Projection

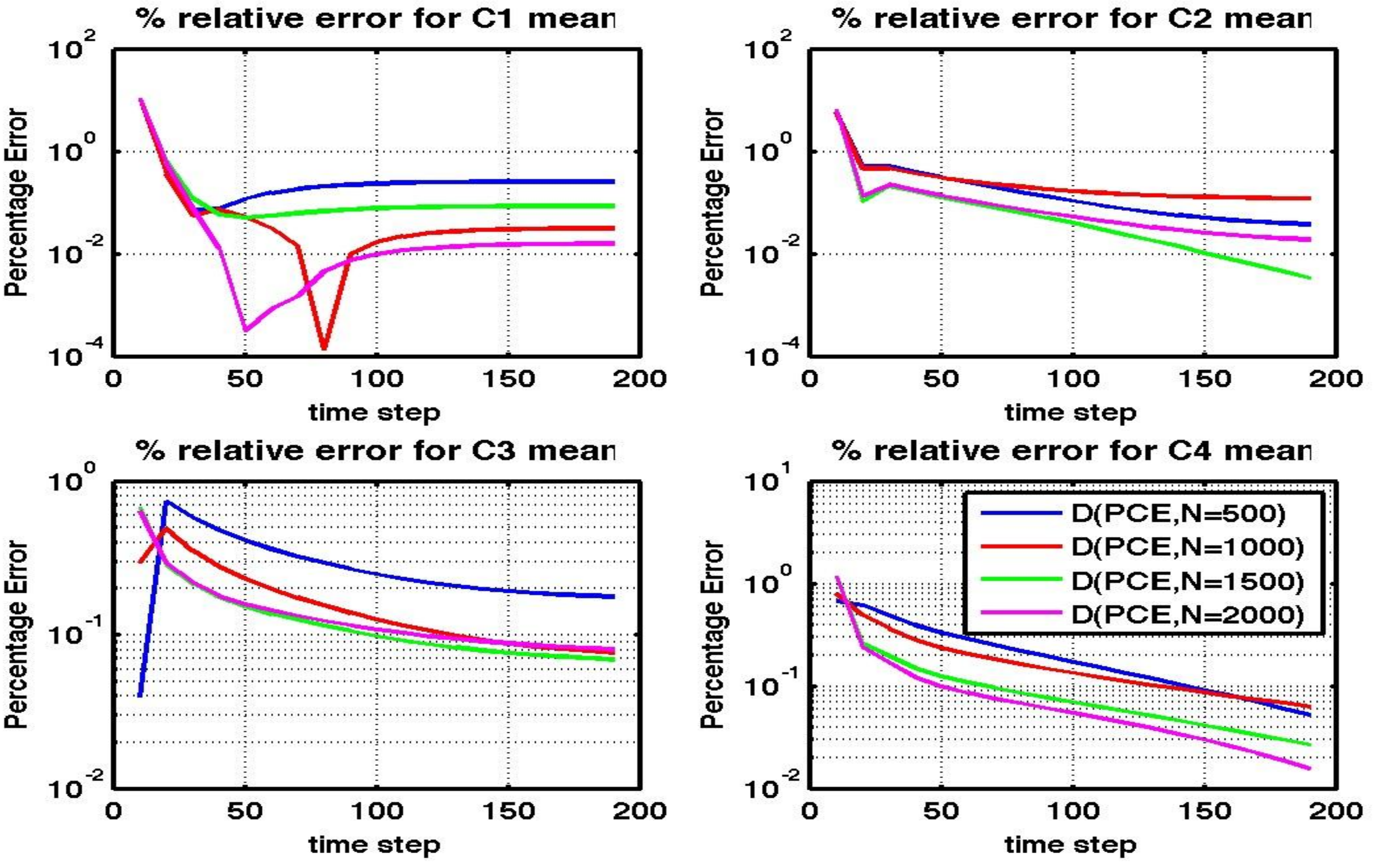
Find $u_{i,k}$, $i \in \mathcal{N}$, $0 \leq k \leq P$ for $\forall w_{j,k}$ $j \in \mathcal{N}$, $0 \leq k \leq P$ such that

$$
\begin{aligned}
0 \quad = \quad & \sum_{i \in \mathcal{N}} \sum_{k=0}^{P} \left[\int_\Omega \frac{\partial \Phi_i(\mathrm{x},\theta,t)}{\partial t} \Phi_j(\mathrm{x},\theta)\,\mathrm{dx} \right] u_{i,k} w_{j,k} + \frac{1}{R} \Bigg\{ \sum_{i,j \in \mathcal{N}} \\
& + \Bigg[\sum_{k,l,m,s=0}^{P} \alpha_s v_k^x \langle \Psi_k \Psi_l \Psi_m \Psi_s \rangle \left(\int_\Omega \frac{\partial \Phi_i(\mathrm{x},\theta,t)}{\partial x} \frac{\partial \Phi_j(\mathrm{x},\theta)}{\partial x} \mathrm{dx} \right) u_{i,l} w_{j,m} \\
& + \sum_{k,l,m,s=0}^{P} \alpha_s v_k^y \langle \Psi_k \Psi_l \Psi_m \Psi_s \rangle \left(\int_\Omega \frac{\partial \Phi_i(\mathrm{x},\theta,t)}{\partial y} \frac{\partial \Phi_j(\mathrm{x},\theta)}{\partial y} \mathrm{dx} \right) u_{i,l} w_{j,m} \\
& - \sum_{k,l,m=0}^{P} v_k^x \langle \Psi_k \Psi_l \Psi_m \rangle \left(\int_\Omega \frac{\partial \Phi_i(\mathrm{x},\theta,t)}{\partial x} \Phi_j(\mathrm{x},\theta)\,\mathrm{dx} \right) u_{i,l} w_{j,m} \\
& - \sum_{k,l,m=0}^{P} v_k^y \langle \Psi_k \Psi_l \Psi_m \rangle \left(\int_\Omega \frac{\partial \Phi_i(\mathrm{x},\theta,t)}{\partial y} \Phi_j(\mathrm{x},\theta)\,\mathrm{dx} \right) \Bigg] u_{i,l} w_{j,m} \Bigg\}
\end{aligned}
$$

Hybrid UQ for multi-species reactive transport

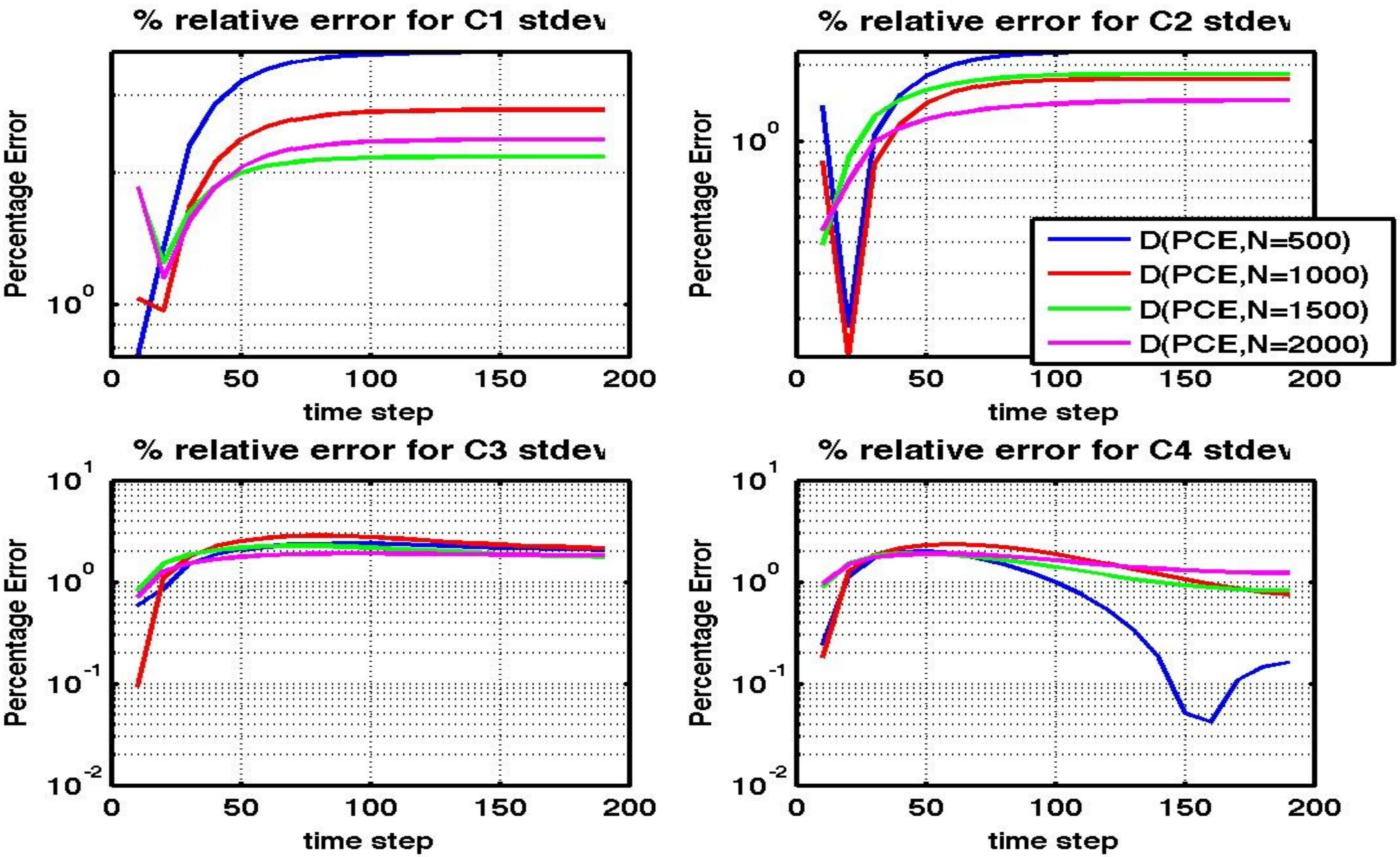
- Operator splitting
  - 2D transport using finite element (4 equations, decoupled)
  - Multi-species reaction (coupled, use analytic solution)
- Test problem
  - Transport (PCE): dispersivity and velocity varies +/- 20%
  - Reaction (non-intrusive): +/-20% in reaction rates
- At each time step
  - Partition the incoming global states into subproblems
  - Perform transport solves
  - Re-package solution into global states
  - Generate initial conditions for each reaction solve
  - Perform reaction solves (sampling)
  - Reconstruct reaction solutions into global states
    - Latin hypercube (least squares)
    - Sparse grid (projection)

Numerical Results on 4-species reactive transport (means)



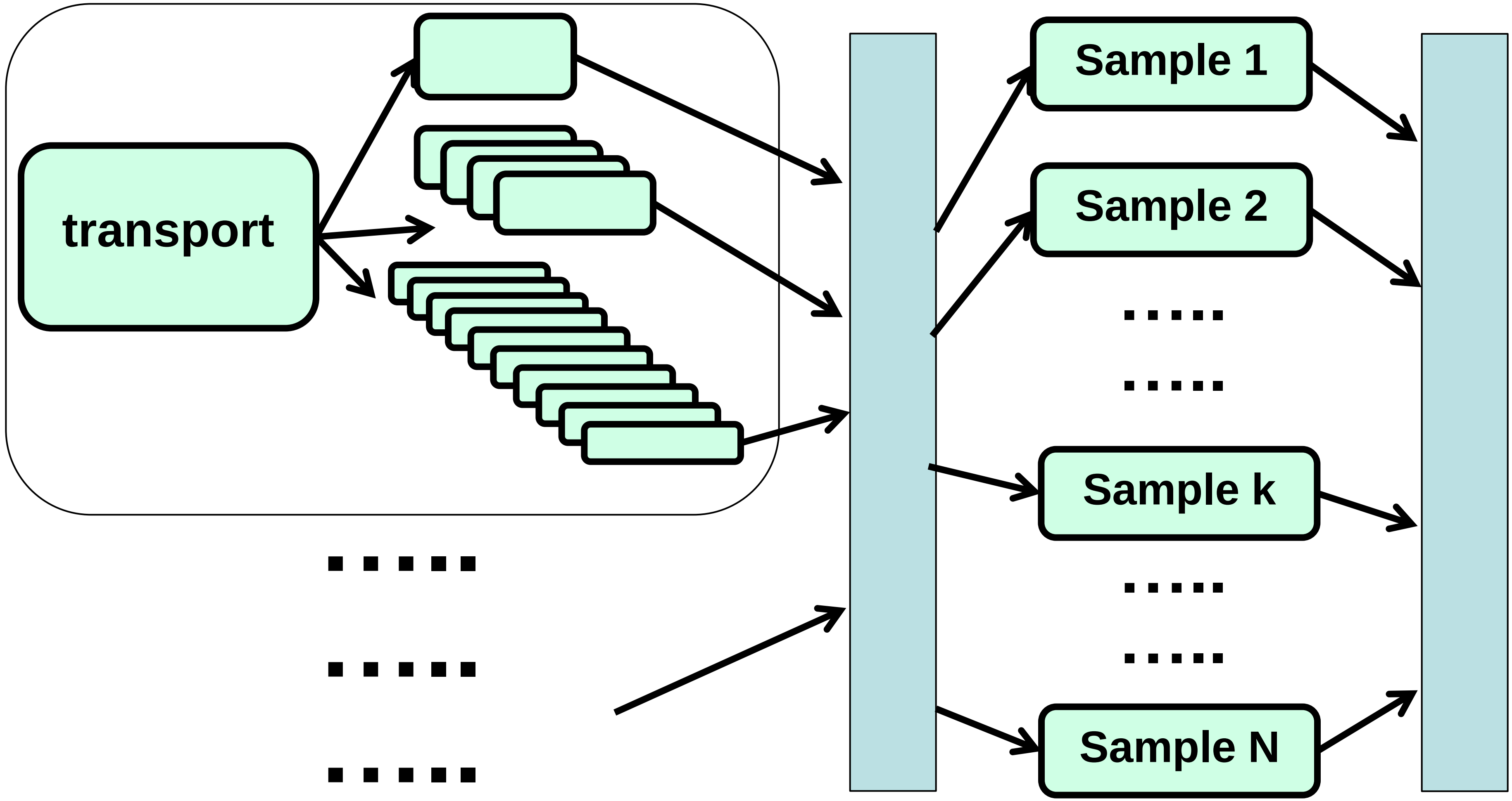
- PCE order = 2 (same for p=3)   Note: absolute differences are small (~5e-4 relative to 1.5e-2)

Numerical Results on 4-species reactive transport (std dev)



- PCE order = 2 (same for p=3)   Note: absolute differences are small (~1e-4 relative to 5e-2)

Scalability analysis of Hybrid UQ methods



- Transport: species doupled, PCE decoupled, multiple rhs systems
- Reaction: each sample instantiation independent from the others
- However, there are all-to-all communications

Summary and future work

- **Hybrid uncertainty quantification is appealing**
  - **accommodate embedded (intrusive) UQ methods**
  - **compatible with modern-day "plug-and-play" philosophy**
  - **facilitate progressive integration of advance UQ methods**
  - **sometimes out of necessity, e.g.**
    - **different time step requirement for each module**
    - **accommodate commercial/open source modules**
  - **may increase parallelism**
- **Still there are many challenges**
  - **high dimensional uncertain parameter space**
  - **highly nonlinear parameter to output mapping**
  - **errors in propagating uncertainties**
  - **load balancing/fault tolerance on HPC**