Chemical Model Reduction under Uncertainty

H. Najm[1], M. Valorani[2], C. Safta[1], M. Khalil[1],
P.P. Ciottoli[2] and R.C. Galassi[2]

[1]Sandia National Laboratories
Livermore, CA
[2]Sapienza University of Rome
Rome, Italy

5th International Workshop on Model Reduction in Reacting Flows
Spreewald, Germany
Jun 28 – Jul 1, 2015

Support: DOE Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, & Biosciences

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

Uncertainty in Reacting Flow Modeling

- Chemical models involve much empiricism
- Model uncertainties: choice of species and reactions
- Parametric uncertainties:
  - Chemical rate constants
  - Thermodynamic parameters
  - turbulence/subgrid models
  - mass/energy transport and fluid constitutive laws
  - geometry and initial/boundary conditions
- Present focus on parametric uncertainty
  - kinetic rate coefficients

Uncertainty and Chemical Model Reduction

- Typical ingredients in chemical model reduction
  - A detailed starting chemical kinetic mechanism M_0
  - Operating conditions of interest
  - Quantities of interest (QoIs) desired with specified accuracy

$$\mathcal{E} \equiv \|\Phi - \Phi_0\| < \alpha$$

- Consequences of uncertainty in the detailed model?
  - Errors in QoIs: acceptable over range of uncertainty
  - QoIs are uncertain – error measure definition
  - Probabilistic measures of model fidelity

$$\begin{aligned} \mathcal{E} \equiv \|\Phi - \Phi_0\| \quad &\Rightarrow \quad P(\mathcal{E} < \alpha) < \epsilon \\ \mathcal{E} \equiv \mathcal{D}[p(\Phi), p(\Phi_0)] \quad &\Rightarrow \quad \mathcal{E} < \alpha \end{aligned}$$

Model Robustness, Error and Uncertainty

- Robustness: A reduced model developed based on a given database should not learn "too much" from the data
  - Reduced models based on different training/test data subsets should not vary "much"
- Optimally, requirements on reduced model error should be made in light of uncertainty in detailed model predictions
  - There is little point in insisting on error bounds much smaller than uncertainty in the reference data
- Measures of reduced model fidelity can include accurate prediction of
  - the nominal reference solution
  - uncertainty in specific observables

Deterministic Chemical ODE System Analysis

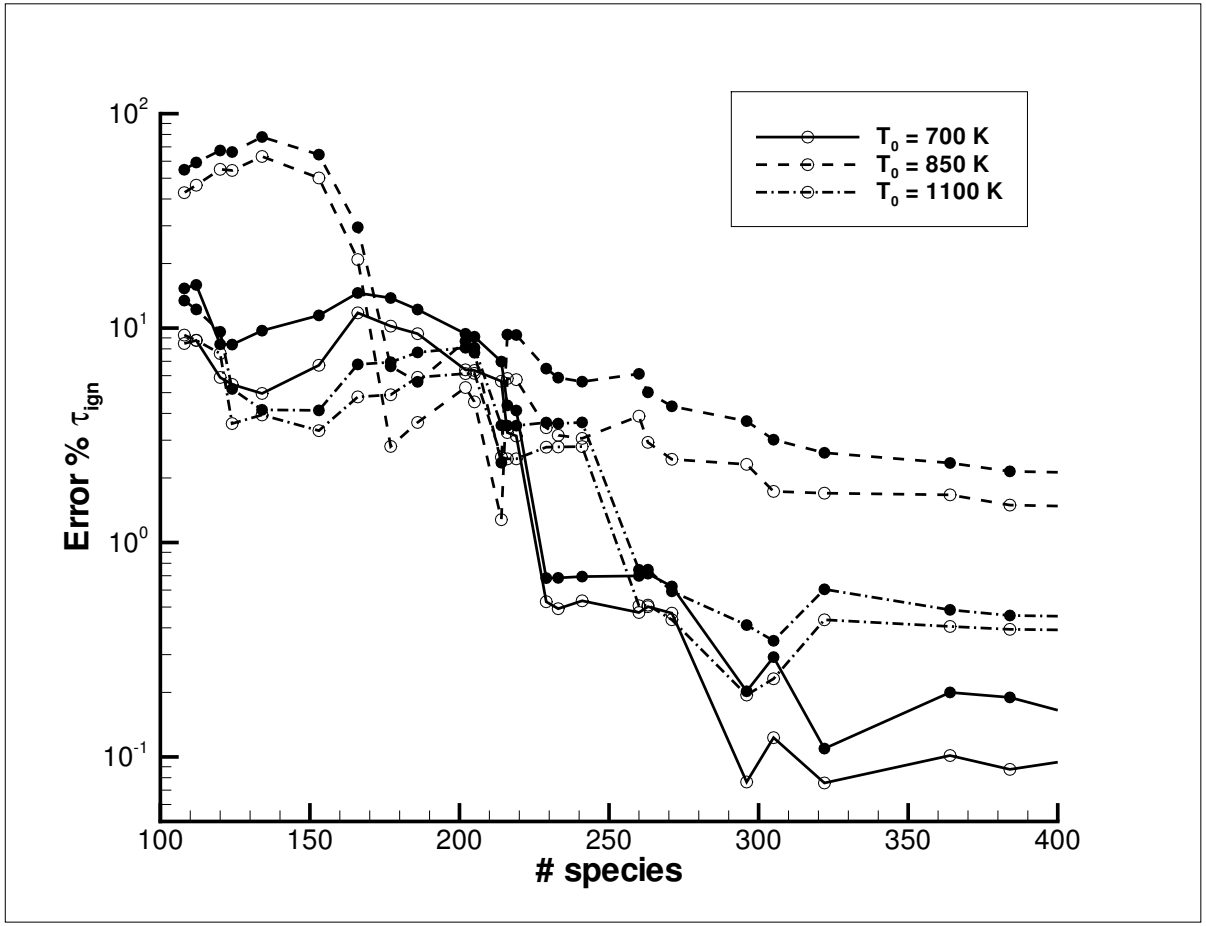
- Computational Singular Perturbation (CSP) analysis
- Jacobian eigenvalues provide first-order estimates of the time-scales of system dynamics: $\tau_i \sim 1/\lambda_i$
- Jacobian eigenvectors provide first-order estimates of the vectors that span the fast/slow tangent spaces
- With chosen thresholds, have M “fast" modes
  - M algebraic constraints define a slow manifold
  - Fast processes constrain the system to the manifold
  - System evolves with slow processes along the manifold
- CSP Importance indices provide estimates of “importance" of a given reaction to a given species in each of the fast/slow subspaces

A CSP-based Mechanism Simplification Algorithm

procedure CSPSIMP($D = \{\boldsymbol{y}_1, \cdots \boldsymbol{y}_N\}, \tau, \mathcal{S}_0, \mathcal{S}^*, \mathcal{R}^*$)
  for each $\boldsymbol{y}_n \in D$ **do** ▷ Loop over database
    $k \leftarrow 0$
    do
      $k \leftarrow k + 1$
      $\mathcal{R}_k \leftarrow \{R_j \,|\, \exists i : (\mathcal{S}_i \in \mathcal{S}_{k-1} \wedge I^i_{j,\mathrm{s}} > \tau) \vee (\mathcal{S}_i \in \mathcal{S}^{\mathrm{rad}}_{k-1} \wedge I^i_{j,\mathrm{f}} > \tau)\}$
      $\mathcal{S}_k \leftarrow \{S_i \,|\, \exists j : (R_j \in \mathcal{R}_k \wedge \nu_{ij} \neq 0)\}$
    while $\mathcal{S}_k \neq \mathcal{S}_{k-1}$
    $\mathcal{R}_n \leftarrow \mathcal{R}_k$
  end for
  $\mathcal{R} \leftarrow \bigcup_n \mathcal{R}_n$ ▷ Active Reactions
  $\mathcal{S} \leftarrow \{S_i \,|\, \exists j : (R_j \in \mathcal{R} \wedge \nu_{ij} \neq 0)\}$ ▷ Active Species
end procedure

Valorani *et al.*CF 2006

nHeptane Kinetic Model Simplification with CSP



- % Relative error in ignition time vs. simplified model sizes
- Control using error tolerances on CSP importance indices

Valorani *et al.*PCI 2007

Probabilistic Analysis of Uncertain ODE Systems

- Handle uncertainties using probability theory
- Every random instance of the uncertain inputs provides a “sample" ODE system
  - Uncertainties in fast subspace lead to uncertainty in manifold geometry
  - Uncertainties in slow subspace lead to uncertain slow time dynamics
- Probabilistic measures of importance
- Probabilistic comparison of models
- One can analyze/reduce each system realization
  - Statistics of $y(t;\lambda)$ trajectories

Reduction Strategy under Uncertainty

- Deterministic strategy:
  - Given
    - Detailed starting chemical model M^*, with parameters λ
    - Solution database D of state vectors generated with (M^*, λ)
    - Quantities of interest I
    - Specified error thresholds τ on I
  - discover a simplified model $M(M^*, \lambda, D, I, \tau) := M(\lambda)$
- Probabilistic strategy:
  - Given uncertainty in λ, we model this parameter vector as a random vector with a given joint density $p(\lambda)$.
  - As a result, the resulting model structure $M(\lambda)$ is a random object, with a probability for any given M, denoted by $P(M)$.
  - Each $M \in \mathcal{M}$ is defined by a network of species/reactions
  - The set $\mathcal{M}$ is not easy to work with

Convenient coordinates on model space

- Given the starting detailed model M^*, any simplifed model M is uniquely defined by the set of retained reactions
  - Retained species are those involved in retained reactions
- Set of elementary reactions in M^*: $\mathcal{R}_{M^*} = \{R_1, \cdots, R_K\}$
- Define the bit vector $\boldsymbol{\alpha} = (\alpha_1, \cdots, \alpha_K) \in \{0,1\}^K$
- A model M is specified by $\boldsymbol{\alpha}(M)$ where, for $r = 1, \ldots, K$,

$$\alpha_r(M) = \begin{cases} 1 & \text{if } R_r \in \mathcal{R}_M \\ 0 & \text{otherwise} \end{cases}$$

clearly: $\boldsymbol{\alpha}(M^*) = (1, \cdots, 1)$

- Thus, given M^*, we have the mapping: $\lambda \rightarrow \boldsymbol{\alpha}(\lambda)$

Uncertain Simplified Model Specification

- For uncertain λ: $p(\lambda) \rightarrow P(\boldsymbol{\alpha}) \equiv P_{\boldsymbol{\alpha}}$
- Clearly, $P_{\boldsymbol{\alpha}} \geq 0$, and $\sum_{\boldsymbol{\alpha}} P_{\boldsymbol{\alpha}} = 1$
- Illustrative example: M^*: A $\underset{2}{\overset{1}{\rightleftharpoons}}$ B
  - $K = 2$, such that $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$
  - Set of possible values of $\boldsymbol{\alpha}$: $\{(1,1),(1,0),(0,1),(0,0)\}$
  - Set of possible models M: $\{M_{(1,1)}, M_{(1,0)}, M_{(0,1)}, M_{(0,0)}\}$
  - Uncertain simplified model specification:

$$\{P_{(1,1)}, P_{(1,0)}, P_{(0,1)}, P_{(0,0)}\}$$

where $P_{(i,j)} \equiv P(\boldsymbol{\alpha} = (i,j))$

Uncertain Reduction Strategy – 1

- Generate N random samples of λ from $p(\lambda)$
- For each λ^i, $i = 1, \ldots, N$
  - Analyze resulting $M^*(\lambda^i)$ for ignition – range of (T, P, Φ) ICs
  - Get simplified model $M^i(\mathcal{S}^i, \mathcal{R}^i)$
  - Evaluate $\boldsymbol{\alpha}^i = \boldsymbol{\alpha}(M^i)$:

$$\alpha_k^i = \begin{cases} 1 & \text{for } R_k \in \mathcal{R}_{M^i} \\ 0 & \text{otherwise} \end{cases} \qquad k = 1, \ldots, K$$

- Estimate Model probabilities: $P_{\boldsymbol{\alpha}} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\boldsymbol{\alpha}, \boldsymbol{\alpha}^i}$
- Marginal reaction probabilities:

$$P_{\alpha_k} = \frac{1}{N} \sum_{i=1}^{N} \delta_{\alpha_k, \alpha_k^i}, \qquad k = 1, \ldots, K$$

Uncertain Reduction Strategy – 2

- Marginal reaction inclusion probability

$$P_k := P_{\{\alpha_k=1\}} = \frac{1}{N}\sum_{i=1}^{N} \alpha_k^i, \qquad k = 1, \ldots, K$$

- Include reaction k iff:

$$P_k > \theta$$

- Resulting model is the CSP-simplified model given
  - the starting detailed model $M^*(\lambda)$
  - the database of solution state vectors
  - the CSP Importance Index tolerance τ
  - for $\lambda \sim p(\lambda)$

  with marginal reaction inclusion probability $> \theta$

Estimation of Moments in High Dimensional Space

- Estimation of expectations (e.g. $P_{\boldsymbol{\alpha}}$) relies on integration

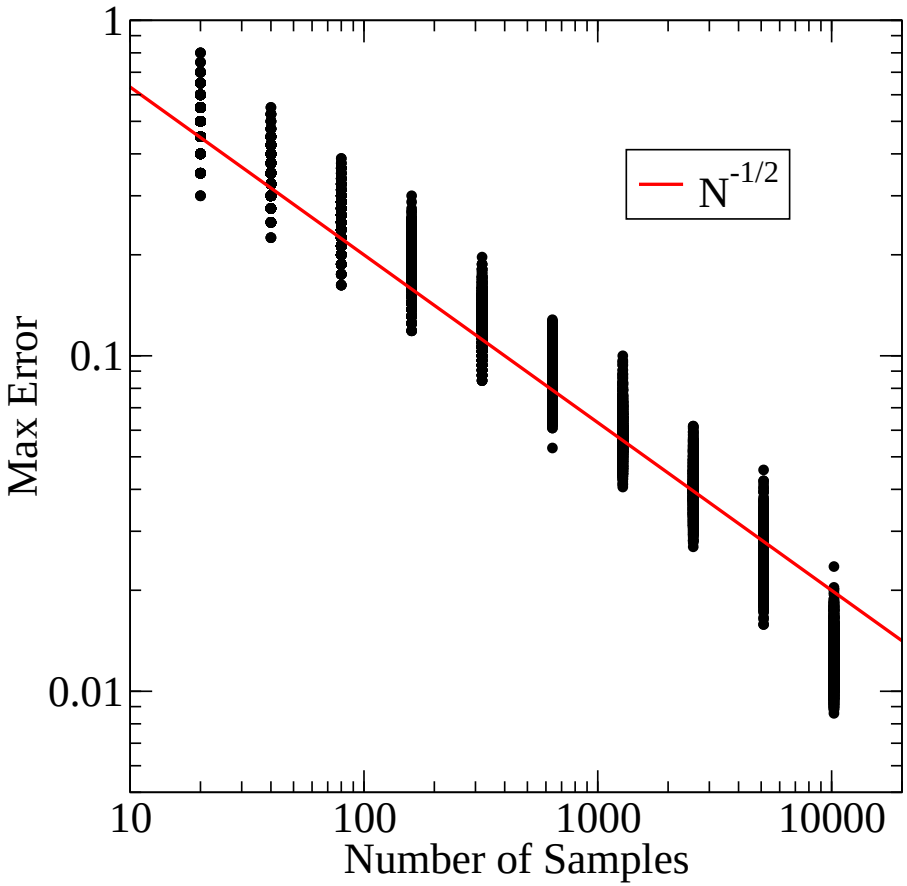
$$P_{\boldsymbol{\alpha}} = \int_{\Lambda} \delta_{\boldsymbol{\alpha},\boldsymbol{\alpha}(\lambda)} p(\lambda) \mathrm{d}\lambda \approx \frac{1}{N} \sum_{i=1}^{N} \delta_{\boldsymbol{\alpha},\boldsymbol{\alpha}(\lambda^i)} \Big|_{\lambda \sim p(\lambda)}$$

- High dimensional space: $\lambda \in \Lambda \subset \mathbb{R}^L$, where $L \geq K$
- Monte Carlo sampling useful for evaluating hi-D integrals
  - particularly when the integrand is non-smooth
- MC convergence rate independent of dimensionality
  - However, the level of error for a given number of samples increases with the intrinsic dimensionality of the integrand
- Concentration of measure – $E[\|\lambda\|] \uparrow$ and $V[\|\lambda\|] \downarrow$ as $L \uparrow$
  - $L \uparrow \implies$ every λ^i likely to have one/more extreme elements
  - Use truncated distributions $p(\lambda)$

Computational Considerations

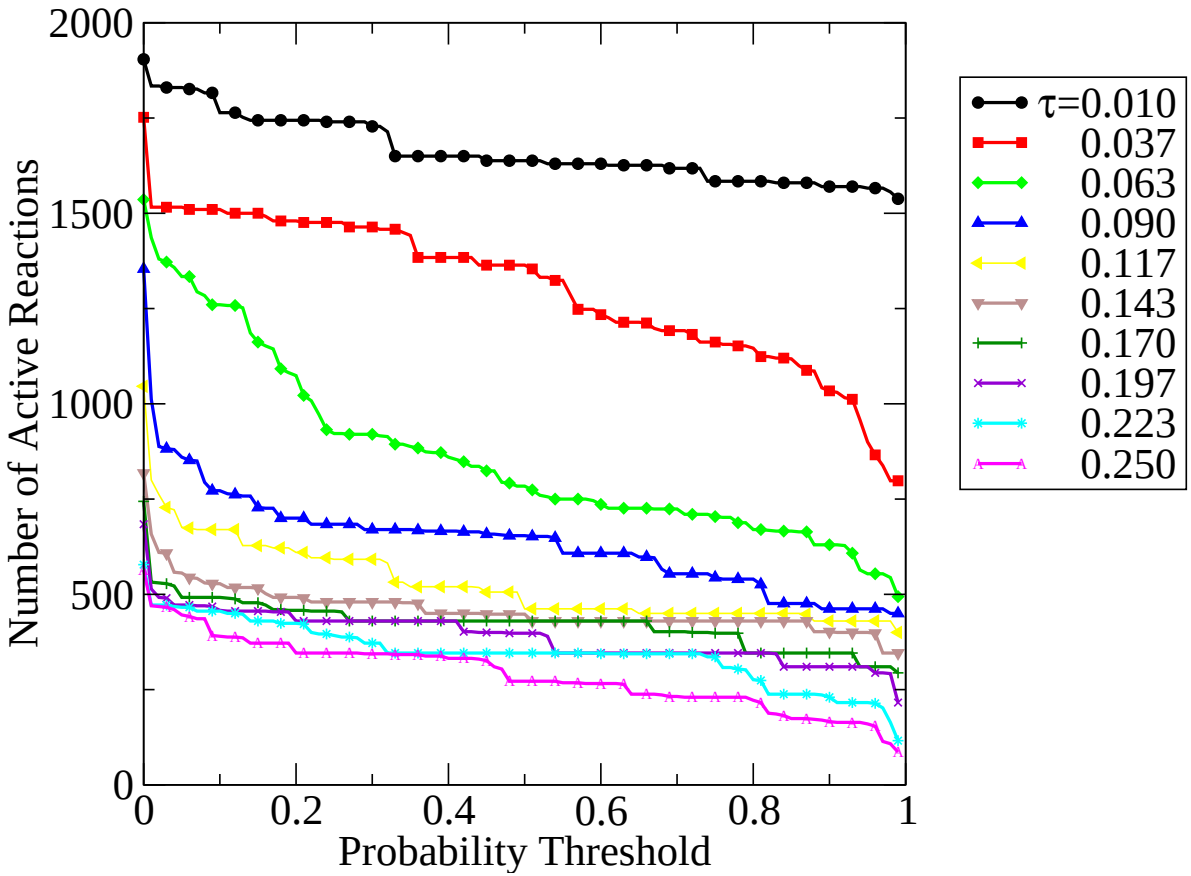
- Efficient `Tchem` based thermochemistry
  - In-memory manipulation of Arrhenius parameters
  - Fast source term and analytical Jacobian evaluation
  - http://www.sandia.gov/tchem
- Fast `cvode` based stiff time integration
  - http://computation.llnl.gov/casc/sundials
- Efficient `CSPTk` analysis and reduction
  - Minimal I/O
  - Adaptive/selective evaluation of Importance Indices
  - Contact: M. Valorani: mauro.valorani@uniroma1.it
- In-memory statistics of random trajectories and associated analysis

Convergence with number of MC samples–GRImech



- Self-convergence of max error in P_{α} with increasing number of MC samples
- Expected slope of $1/\sqrt{N}$ in ensemble mean error

Number of active reactions vs. Probability threshold



- Number of retained/active reactions goes down with
  - increasing threshold θ on P_k
  - increasing Importance Index threshold τ

A posteriori error estimation

- Given mass fractions $Y_i(t), i = 1, \ldots, N$ for all species in a simplified mechanism
- Parametrize Y_i in terms of fuel mass fraction: $Y_f \in [0, Y_{f0}]$
- Let $t_{Y_f} = t(Y_f) := \{t \mid Y_{i_f}(t) = Y_f\}$

Define the following trajectory error

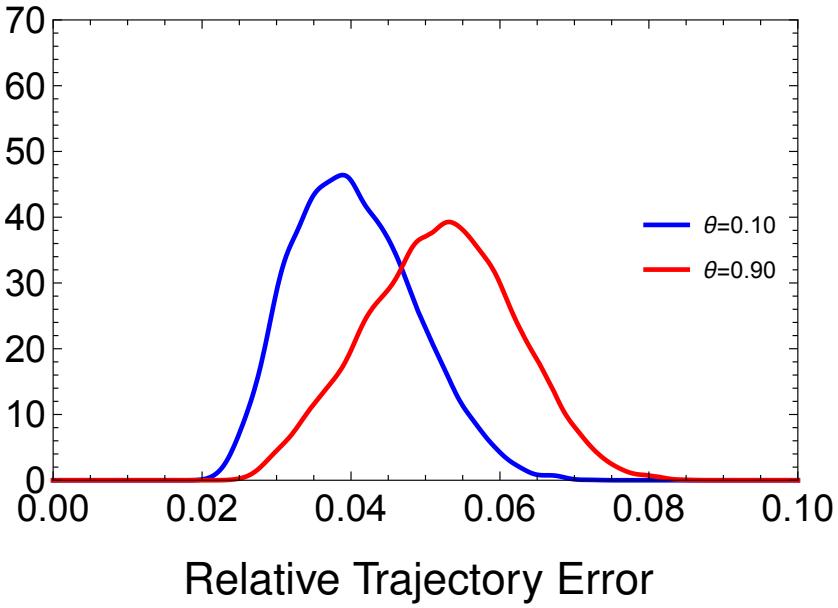
$$E_i = \frac{1}{Y_{f0}} \int_{Y_{f0}}^{0} \left| \frac{Y_i(t_{Y_f}) - Y_i^*(t^*_{Y_f})}{Y_i^*(t^*_{Y_f})} \right| \mathrm{d}Y_f, \qquad i = 1, \ldots, N$$

where * corresponds to the detailed model

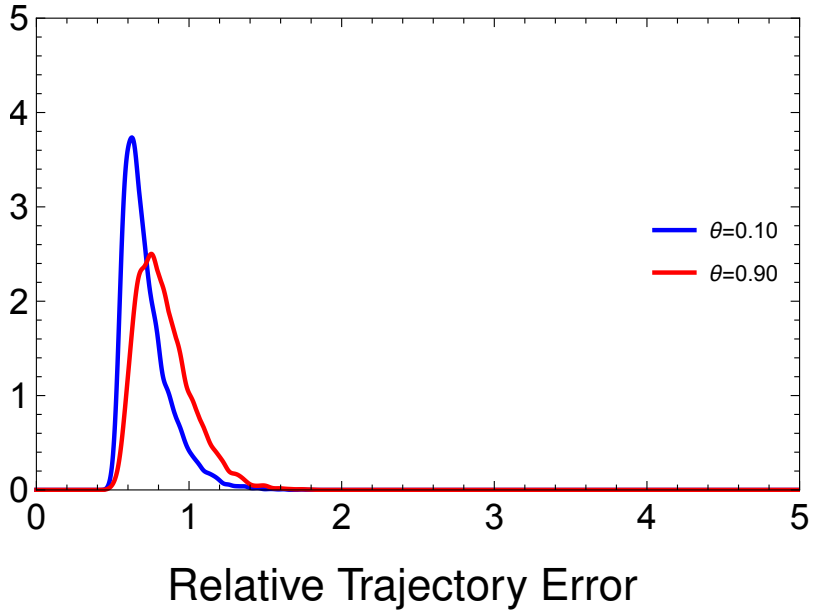
Trajectory error PDF

- PDF of trajectory error for $\tau = 0.22$
- Error averaged over set of species

Target species

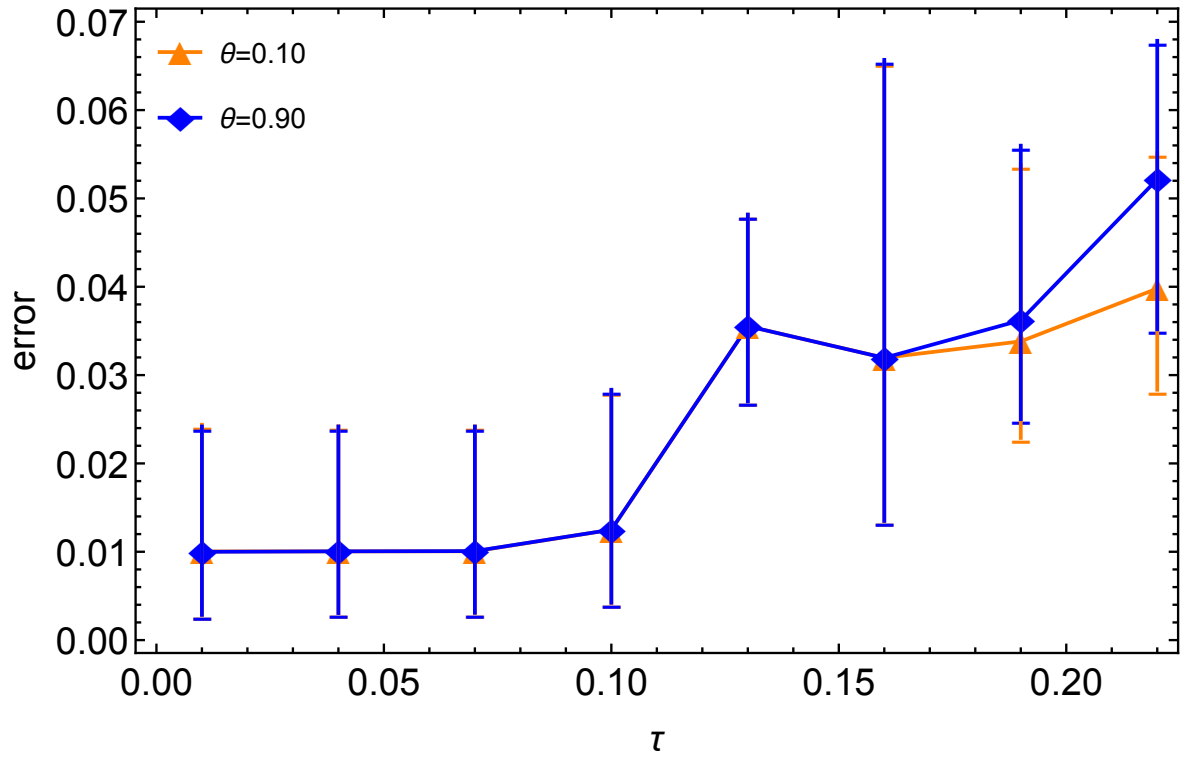


All species



- Error increases with θ
  - less reactions included with higher inclusion threshold

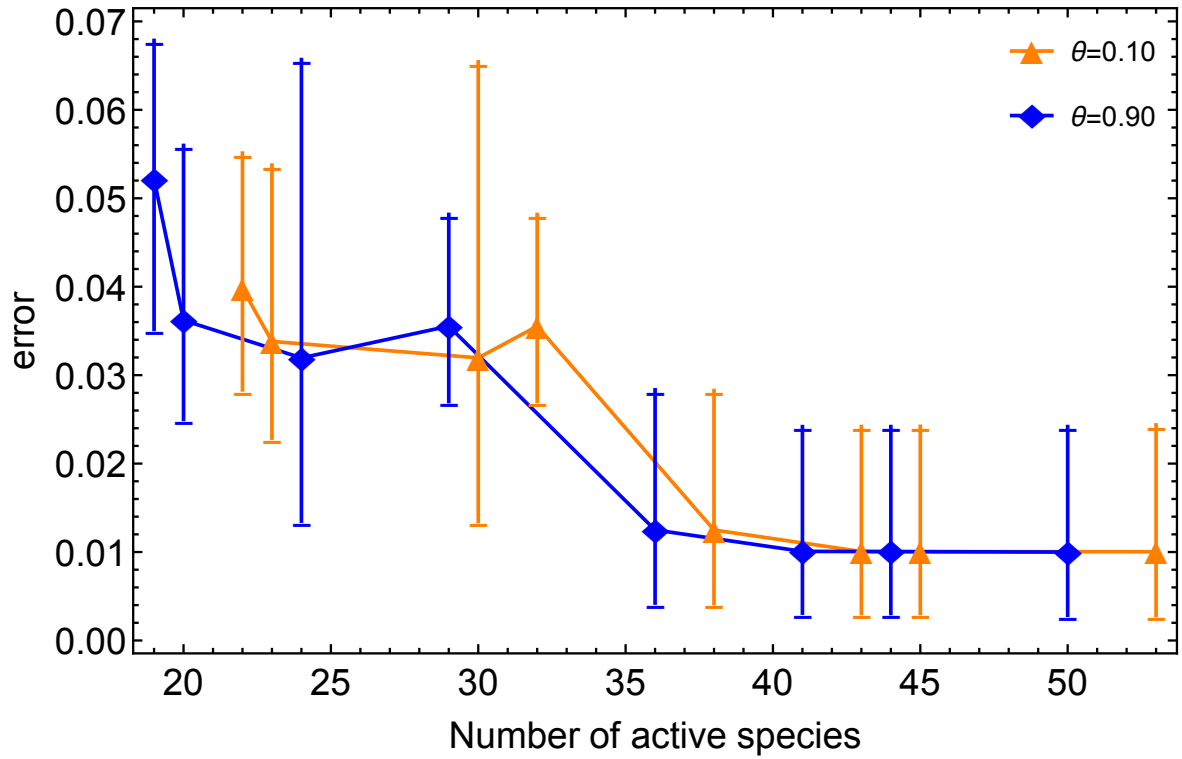
Trajectory error statistics – target species



Median, 5% and 95% quantiles of trajectory error versus τ

- Trajectory error increase with τ and θ

Trajectory error statistics – target species



Median, 5% and 95% quantiles of trajectory error versus N_{spec}

- Trajectory error decrease with N_{spec}

Closure

- We presented a probabilistic framework for analysis and reduction of chemical models under uncertainty
- The construction employs the target problem and the deterministic analysis/reduction strategy as a black box, which is queried to compile relevant statistics
- We employ a convenient indexing of models using a binary vector, which facilitates the analysis
- We demonstrated the construction, and illustrated its convergence behavior given relevant tolerances
- More work ahead on *a posteriori* error analysis and convergence analysis in more complex cases