# *Chemical Model Reduction u*

SAND2015-11096C

H. Najm[1], R.C. Malpica Galassi[2], and M. Valorani[2]

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

SRI-UQ Workshop
KAUST, KSA
Jan 5 – 10, 2016

*Support:*
*DOE Office of Basic Energy Sciences, Division of Chemical Sciences, Geosciences, & Biosciences*
*Clean Combustion Research Center, King Abdullah University of Science and Technology (KAUST)*

*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**

      $k \leftarrow k + 1$

$$\mathcal{R}_k \leftarrow \{R_j \,|\, \exists i : (\mathcal{S}_i \in \mathcal{S}_{k-1} \wedge I^i_{j,\mathsf{s}} > \tau) \vee (\mathcal{S}_i \in \mathcal{S}^{\mathsf{rad}}_{k-1} \wedge I^i_{j,\mathsf{f}} > \tau)\}$$

      $\mathcal{S}_k \leftarrow \{S_i \,|\, \exists j : (R_j \in \mathcal{R}_k \wedge \nu_{ij} \neq 0)\}$

    **while** $\mathcal{S}_k \neq \mathcal{S}_{k-1}$

    $\mathcal{R}_n \leftarrow \mathcal{R}_k$

  **end for**

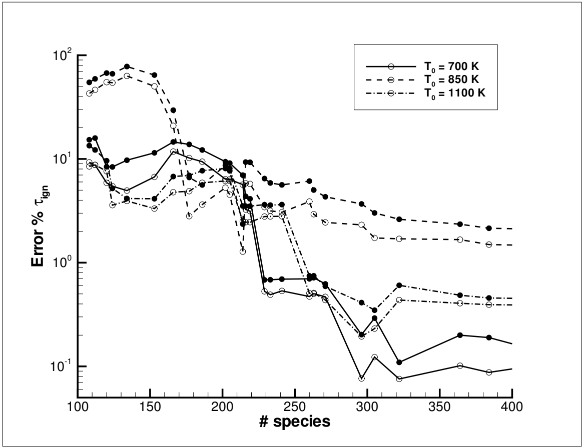
  $\mathcal{R} \leftarrow \bigcup_n \mathcal{R}_n$ ▷ Active Reactions

  $\mathcal{S} \leftarrow \{S_i \,|\, \exists j : (R_j \in \mathcal{R} \wedge \nu_{ij} \neq 0)\}$ ▷ 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 $\lambda$
    - Solution database $D$ of state vectors generated with $(M^*, \lambda)$
    - Quantities of interest $I$
    - Specified error thresholds $\tau$ on $I$
  - discover a simplified model $M(M^*, \lambda, D, I, \tau) := M(\lambda)$
- Probabilistic strategy:
  - Given uncertainty in $\lambda$, 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: $\lambda \to \boldsymbol{\alpha}(\lambda)$

# Uncertain Simplified Model Specification

- For uncertain $\lambda$: $p(\lambda) \rightarrow P(\boldsymbol{\alpha}) \equiv P_{\boldsymbol{\alpha}}$
- Clearly, $P_{\boldsymbol{\alpha}} \geq 0$, and $\sum_{\boldsymbol{\alpha}} P_{\boldsymbol{\alpha}} = 1$
- Illustrative example: $M^*$: A $\underset{2}{\overset{1}{\rightleftarrows}}$ 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 $\lambda$ from $p(\lambda)$
- For each $\lambda^i$, $i = 1, \ldots, N$
  - Analyze resulting $M^*(\lambda^i)$ for ignition – range of $(T, P, \Phi)$ 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

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

- Include reaction $k$ iff:

$$P_k > \theta$$

- Resulting model $M_{\tau,\theta}(\lambda)$ is the CSP-simplified model given
  - the starting detailed model $M^*(\lambda)$
  - the database of solution state vectors
  - the CSP Importance Index tolerance $\tau$
  - 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 $\lambda^i$ 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 evaluation of source term and analytical Jacobian
  - http://www.sandia.gov/tchem
  - Contact: C. Safta: csafta@sandia.gov
- Fast `cvode` based stiff time integration
  - http://computation.llnl.gov/casc/sundials
- Efficient `CSPTk` analysis and reduction
  - Minimal I/O
  - On-demand/as-needed evaluation of Importance Indices
  - Contact: M. Valorani: mauro.valorani@uniroma1.it
- In-memory statistics of random trajectories and associated analysis

# Demo on n-butane ignition

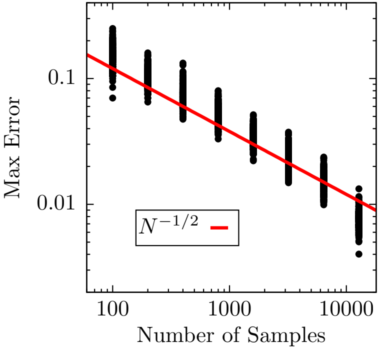
- Detailed chemical mechanism for n-butane/air combustion, with specified uncertainty factors in the pre-exponentials
  E. Hebrard, A.S. Tomlin, R. Bounaceur, F. Battin-Leclerc, Proc. Comb. Inst. **35**(1):607-616, 2015.
- $N = 1111$ reactions
- Temperature-dependent uncertainty factors
  - Mechanism specificies $(f_r, g_r)$ for each reaction $r$
  - Uncertainty factor: $\ln A = \ln A_{\text{nom}} \pm \ln F$

$$F_r(T) = f_r \exp\left(\left| g_r \left(\frac{1}{T} - \frac{1}{300}\right)\right|\right)$$

- For now, we employ a temperature-independent $F_r$:

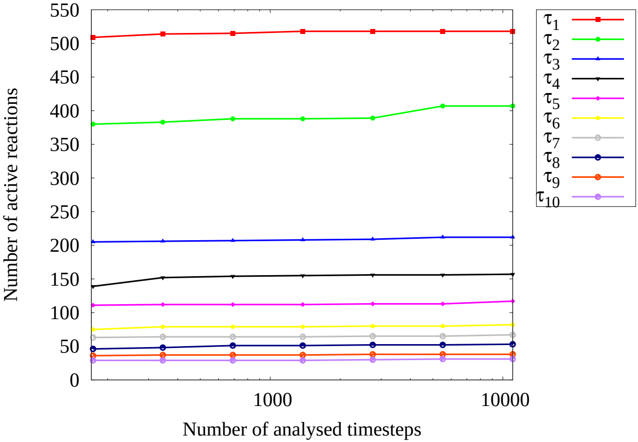
$$F_r := F_r(T)|_{T=1500\text{ K}}$$

# Convergence with number of MC samples




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

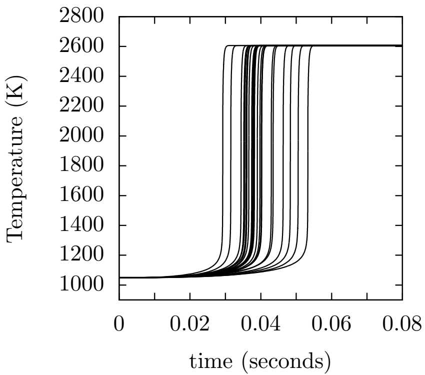
# Convergence with number of sampled trajectory steps




- Somewhat weak dependence on the number of sampled points in each trajectory

# Sampled ignition trajectories – detailed mechanism

- Significant uncertainty in ignition time
- Large range of state-variable uncertainty vs time
  - fast ignition transient
- Examine trajectory errors and uncertainty in an alternate progress-variable phase space
  - Entropic phase space



# Normalized entropy progress variable

- Define the entropy progress variable:

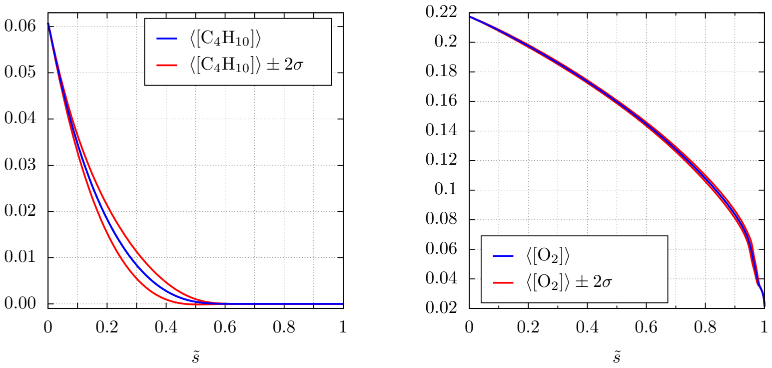
$$\hat{s}(t) = \int_0^t |\mathrm{d}s|$$

- Given any quantity of interest $\phi(t, \cdot)$
- And, time trajectory $(t_k, \phi(t_k, \cdot))$, $k = 1, \ldots, K$
- Re-parametrize the trajectory using a normalized entropy:

$$\tilde{s}(t) = \frac{\hat{s}(t)}{\hat{s}(t_{\text{final}})}$$

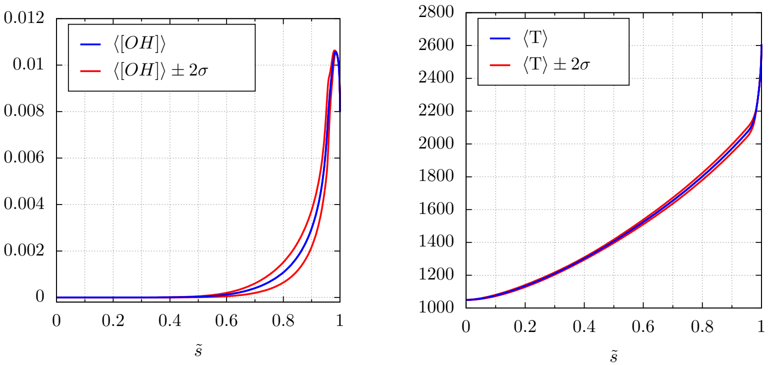
- Thus, $\phi(\mathfrak{s}, \cdot) := \phi(t_{\mathfrak{s}}, \cdot)$ where $t_{\mathfrak{s}} = t(\mathfrak{s}) := \{t \mid \tilde{s}(t) = \mathfrak{s}\}$
- Compare solutions in the $(\mathfrak{s}, \phi)$ phase space
  - Trajectories $(\mathfrak{s}_k, \phi(\mathfrak{s}_k, \cdot))$, $k = 1, \ldots, K$

# Uncertain trajectories in the entropy phase space



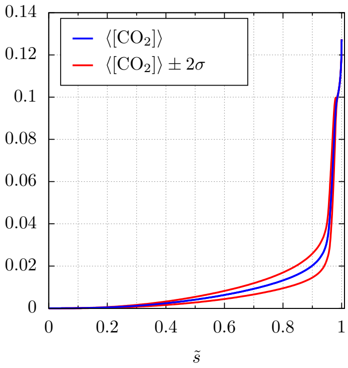

- Mean and Mean$\pm 2\sigma$ trajectories for select state variables
- Entropic phase-space trajectories are explicit functions $y = f(x)$ – by construction
- Coefficient of variation can be large when the mean is low

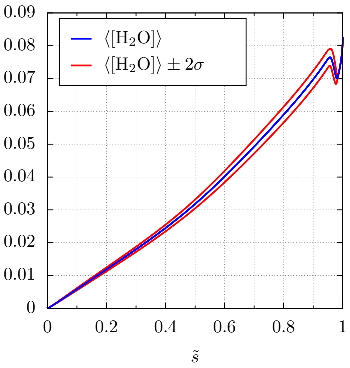
# Uncertain trajectories in the entropy phase space




- Mean and Mean$\pm 2\sigma$ trajectories for select state variables
- Entropic phase-space trajectories are explicit functions $y = f(x)$ – by construction
- Coefficient of variation can be large when the mean is low

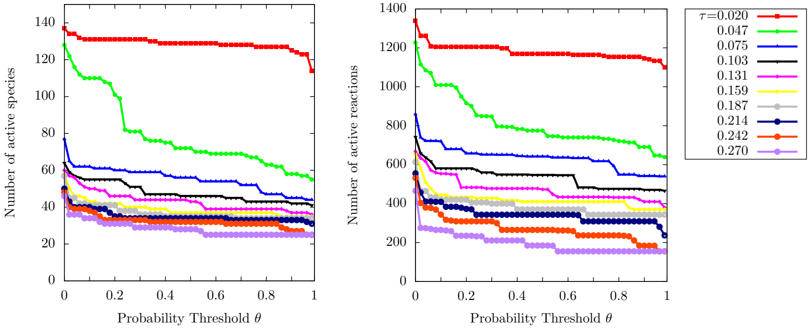
# Uncertain trajectories in the entropy phase space




- Mean and Mean$\pm 2\sigma$ trajectories for select state variables
- Entropic phase-space trajectories are explicit functions $y = f(x)$ – by construction
- Coefficient of variation can be large when the mean is low

# # active reactions/species varies inversely with $(\tau, \theta)$

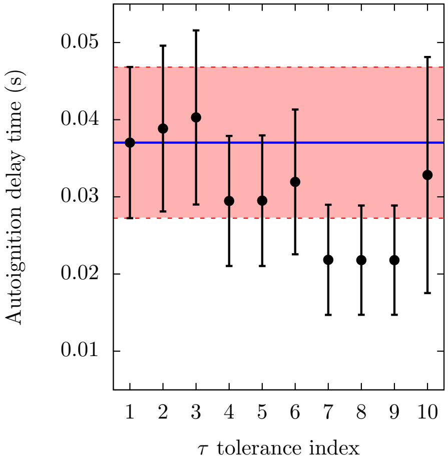



- Number of retained/active reactions/species goes down w/:
  - increasing threshold $\theta$ on $P_k$
  - increasing Importance Index threshold $\tau$

# Uncertain Ignition time prediction

- Detailed model: $t^d_{\text{ign}}(\lambda)$
- Simplified model: $t_{\text{ign}}(\lambda, \tau, \theta = 0.3)$
- Global trend towards
  - lower $t_{\text{ign}}$
  - larger $t_{\text{ign}}$-error

  with increasing $\tau$
- Non-monotonous local behavior



# A posteriori error estimation

For any quantity of interest $\phi(\mathfrak{s},\cdot)$, define the trajectory error norm over time steps $t_k : k = 1,\ldots,K$, with $\mathfrak{s}_k = \tilde{s}(t_k)$,

$$\mathcal{E}_\phi^{p,w} = \frac{\|\phi - \phi_d\|_{p,w}}{\|\phi_d\|_{p,w}} = \frac{\left(\sum_{k=1}^{K} w_k |\phi(\mathfrak{s}_k,\cdot) - \phi_d(\mathfrak{s}_k,\cdot)|^p\right)^{1/p}}{\left(\sum_{k=1}^{K} w_k |\phi_d(\mathfrak{s}_k,\cdot)|^p\right)^{1/p}}$$

where

$\phi_d$ refers to the detailed model

$w_k = w(\mathfrak{s}_k)$ is a weight function *e.g.* $= 1/\sigma_d(\mathfrak{s}_k)$

# A posteriori error estimation

Example quantities of interest for trajectory error estimation

- A per-trajectory error that is random

$$\phi(\mathfrak{s}, \lambda) := X_i(\mathfrak{s}, \lambda) \qquad \Rightarrow \qquad \mathcal{E}^{p,w}_{X_i}(\lambda; \tau, \theta)$$

–Trajectories need to be in the same probability space

- An error in the mean of uncertain trajectories

$$\phi(\mathfrak{s}) := \mu_i(\mathfrak{s}) = \mathsf{E}_\lambda[X_i(\mathfrak{s}, \lambda)] \qquad \Rightarrow \qquad \mathcal{E}^{p,w}_{\mu_i}(\tau, \theta)$$

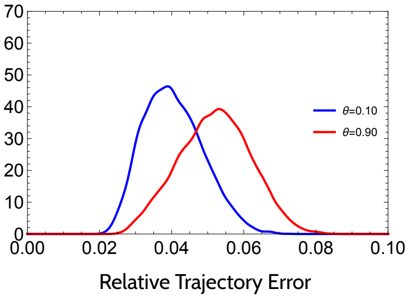
- An error in the standard deviation of uncertain trajectories

$$\phi(\mathfrak{s}) := \sigma_i(\mathfrak{s}) = (\mathsf{V}_\lambda[X_i(\mathfrak{s}, \lambda)])^{1/2} \qquad \Rightarrow \qquad \mathcal{E}^{p,w}_{\sigma_i}(\tau, \theta)$$

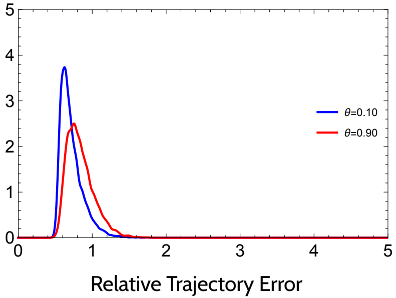
# Trajectory error PDF

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

Target species

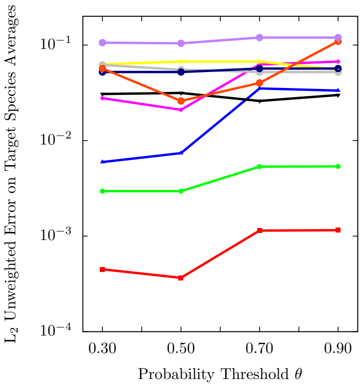



All species

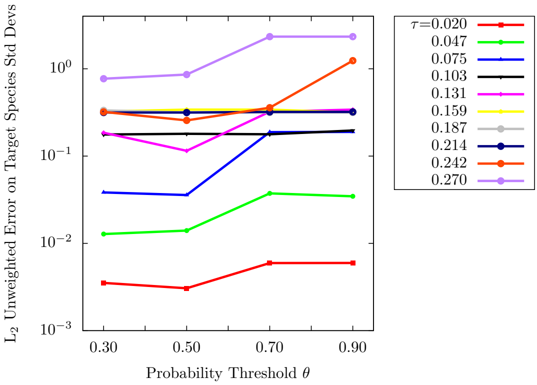



- Error increases with $\theta$
  - less reactions included with higher inclusion threshold

# Trajectory error statistics – target species




Mean




Standard Deviation

- General trend towards higher error with increasing $\tau$, $\theta$
- Trend is more evident at low $\tau$

# 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
- We employ a convenient indexing of models
- We use *a posteriori* error norms for entropic phase space trajectories
- We demonstrated the construction with an uncertain n-butane mechanism
  - Examined *a posteriori* errors
  - Convergence behavior given relevant tolerances