# *Analysis of Uncertain Dynamical Network Models*

R.D. Berry[1] H.N. Najm[1] B. Sonday[1]
B.J. Debusschere[1] H. Adalsteinsson[1] Y.M. Marzouk[2]

[1]Sandia National Laboratories, Livermore, CA

[2]Mass. Inst. of Tech., Cambridge, MA

Stochastic Multiscale Methods: Bridging the Gap Between Mathematical Analysis and Scientific and Engineering Applications

BANF International Research Station, BANF, Canada
March 27 – April 1, 2011

# Acknowledgement

R.G. Ghanem — Univ. Southern California, Los Angeles, CA
O.M. Knio — Johns Hopkins Univ., Baltimore, MD

This work was supported by:

- The US Department of Energy (DOE), Office of Advanced Scientific Computing Research (ASCR), Applied Mathematics program, 2009 American Recovery and Reinvesment Act.
- The DOE Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences.
- BS acknowledges DOE Computational Science Graduate Fellowhip support, which is provided under grant number DE-FG02-97ER25308.

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

# Motivation

- Many physical systems are governed by network models
  - Electric grids
  - Biochemical/chemical models
  - Internet, communication networks, ...
- Resulting models are complex
  - Large number of governing equations (dimension $n$)
  - Large number of connections/reactions
  - Strong non-linearity – ODEs/DAEs
  - Large range of time scales – stiffness
- Need for analysis and model reduction methods
  - Krylov projection methods
  - Methods based on dynamical analysis
    - Automated identification of slow manifolds

# Uncertainty in Network Models

- Network ODE models typically rely on empirically-based parameters/inputs
  - Uncertain parameters/inputs
  - Uncertain network structure
- Need for dynamical analysis methods that
  - Can handle uncertainty
  - Provide model reduction with quantified fidelity
    - *accounting* for uncertainty
- Uncertain ODE systems, $x(t) \in \mathbb{R}^n$

$$
\begin{aligned}
\frac{dx}{dt} &= f(x;\lambda) \\
x(0) &= x_0
\end{aligned}
$$

# UQ Challenges in complex Network models

- Bifurcations
  - Transitions between operating regimes, switching
  - Instability; Ignition

  $\Rightarrow$ MC; Smooth observables; Multi-element local PC methods
- Phase error growth and oscillatory dynamics
  - Uncertain dynamics over long time horizons

  $\Rightarrow$ MC; Smooth observables; Time-shifting
- High Dimensionality
  - Large number of uncertain parameters or degrees of freedom

  $\Rightarrow$ MC; Non-intrusive Sparse-Quadrature; Adaptive bases

# Deterministic Nonlinear 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 right/left eigenvectors provide first-order estimates of the CSP vectors/covectors that define decoupled fast/slow subspaces
- 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 *time-scale-aware* Importance indices provide means for elimination of “unimportant" network nodes and connections for a selected observable

# 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
- One can analyze/reduce each system realization
  - Statistics of $x(t;\lambda)$ trajectories
- This can be expensive!
- Explore alternate means

# Spectral Stochastic Representations

Let $(\Omega, \sigma, \rho)$ be a probability space.
Let $\xi : \Omega \to \mathbb{R}^m$ be an $L^2$ RV.
Let $(\Xi, s, \mu) = \xi_\sharp(\Omega, \sigma, \rho)$.
Let $\{\varphi_\alpha(\xi) : \alpha = 0, 1, 2, \dots\}$ be an orthonormal basis of $L^2(\Xi)$.
Let $X : A \times \Omega \to \mathbb{R}$ be an $L^2(\Omega)$ $A$-process. Its closest representative in $L^2(\Xi)$ is

$$X(a, \omega) \simeq \sum_\alpha X_\alpha(a)\, \varphi_\alpha(\xi(\omega))$$

where

$$X_\alpha(a) = \int_\Omega X(a, \omega)\, \varphi_\alpha(\xi(\omega))\, d\rho(\omega) = \langle \varphi_\alpha, X \rangle.$$

Take $m = 1$ for simplicity. $m > 1$ holds by tensor product arguments.

# Galerkin Reformulation

Consider an ODE

$$\dot{x} = f(\xi, x) \qquad x(\xi, 0) = x_0(\xi)$$

with $x(t, \omega) \in \mathbb{R}^n$. Represent $x$ as

$$x(\xi, t) = \sum_{\alpha} x_{\alpha}(t)\, \varphi_{\alpha}(\xi)$$

where

$$x_{\alpha}(t) = \langle \varphi_{\alpha}(\xi),\, x(\xi, t) \rangle$$

and so these coefficients have dynamics

$$\begin{aligned} \dot{x}_{\alpha} &= \left\langle \varphi_{\alpha}(\xi),\, \frac{d}{dt} x(\xi, t) \right\rangle \\ &= \langle \varphi_{\alpha}(\xi),\, f(\xi, x) \rangle \end{aligned}$$

# Jacobian of Sampled System

The dynamical system can be locally characterized by the eigenstructrure of the Jacobian matrix. The entries of the Jacobian matrix $J$ of the sampled system are given by

$$J_{ij}(\xi, t) = \frac{\partial f^i}{\partial x^j}(\xi, x(\xi, t))$$

At each value of time, $J(\xi, t)$ is a random matrix.

# Jacobian Matrix of Reformulated System

The Jacobian matrix of the coefficient system can be thought of as a block matrix with blocks

$$\begin{aligned}\mathcal{J}_{\alpha\beta}(t) &= D_{x_\beta} \int_\Xi f(\xi, x(\xi,t))\, \varphi_\alpha(\xi)\, d\mu(\xi) \\ &= \int_\Xi \varphi_\alpha(\xi)\, J(\xi,t)\, \varphi_\beta(\xi)\, d\mu(\xi) \\ &= \langle \varphi_\alpha, J\varphi_\beta \rangle\end{aligned}$$

Truncate the representation so that $\alpha, \beta = 0, \ldots, P$.
$\mathcal{J}$ is then a $n(P+1) \times n(P+1)$ matrix.

# Dynamical Analysis of the Galerkin PC System

Key questions:

- How do the eigenvalues and eigenvectors of the Galerkin system relate to those of the sampled original system
- What can we learn about the sampled dynamics of the original system from analysis of the Galerkin system
  - fast/slow subspaces
  - slow manifolds
- Can CSP analysis of the Galerkin system be used for analysis and reduction of the original uncertain system

# Relevant Prior Work

- Ghosh and Ghanem (2002-2005)
- Homescu, Petzold, and Serban (Siam Review, 2007)
- Tryoen and Le Maître (JCP, JCAM, 2010)
- Fisher and Bhattacharya, PC Galerkin system eigenvalues (2008)
  - First numerical illustrations that the Galerkin system eigenvalues seem to exist in the support of the eigenvalues of the sampled system
  - System with continguous locus of each stochastic eigenvalue
- Nevai (1980)

# Infinite Jacobian $\mathcal{J}$ ($P = \infty$)

For $P = \infty$, we can prove that the eigenvalues $\lambda_i$ of $J(\omega, t)$ are also eigenvalues of $\mathcal{J}(t)$ in the $L^2$-sense.

We can construct vectors $w_i = \{w_{i,k\gamma}\}$, $k = 1, \dots, n$; $\gamma = 0, \dots, \infty$, where

$$Jw_i = \lambda_i w_i, \qquad i = 1, \dots, n$$

with equality in $L^2$, *i.e.* for each $j\alpha$,

$$\lim_{P\to\infty} \left\| \sum_{k=1}^{n} \sum_{\gamma=0}^{P} \mathcal{J}_{jk\alpha\gamma}(t) w_{i,k\gamma}(\omega, t) - \lambda_i(\omega, t) w_{i,j\alpha}(\omega, t) \right\|_{L^2(\Omega)} = 0.$$

# Essential Numerical Range of $J$

The numerical range of a matrix $M$ is

$$W(M) = \{v^* M v : v \in \mathcal{C}^m, v^* v = \|v\|^2 = 1\}.$$

Note that

$$\text{spect}(M) \subset W(M).$$

The essential numerical range of $J(\xi)$ is

$$\tilde{W}(J) = \bigcup_{\text{a.e. } \xi} W(J(\xi)).$$

# Eigenpolynomials

Let $\lambda_{i\alpha}$, $v_{i\alpha}$ be an eigenvalue/vector pair of $\mathcal{J}^P$:

$$\mathcal{J}\, v_{i\alpha} = \lambda_{i\alpha} v_{i\alpha}.$$

Alternatively,

$$\langle \varphi_\beta(\xi),\, (J(\xi) - \lambda_{i\alpha})\, \mathfrak{v}_{i\alpha}(\xi) \rangle = 0 \qquad \text{for } \beta = 0 \ldots P$$

where $\mathfrak{v}_{i\alpha}(\xi)$ in an $n$-vector with components

$$\mathfrak{v}^k_{i\alpha}(\xi) = \sum_{\gamma=0}^{P} v^{k\gamma}_{i\alpha}\, \varphi_\gamma(\xi).$$

# $n$-dimensional system – Key Results

1. The spectrum of $\mathcal{J}^P$ is contained in the convex hull of the essential range of the random matrix $J$.

$$\text{spect}(\mathcal{J}^P) \subset \text{conv}(\widetilde{W}(J))$$

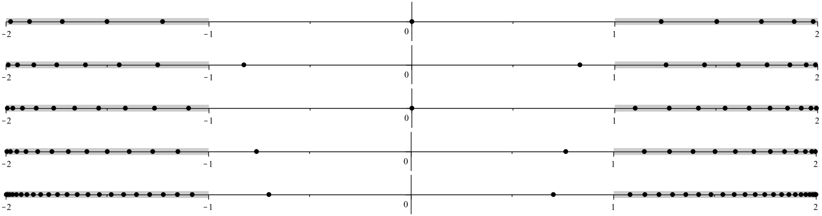
2. For any orthonormal basis $\{\varphi_\alpha\}_{\alpha=0}^{\infty}$:

   As $P \to \infty$, the eigenvalues of $\mathcal{J}^P(t)$ converge weakly, *i.e.* in the sense of measures, toward $\bigcup_{\omega\in\Omega} \text{spect}(J(\omega))$.

3. The $\mathcal{J}^P$ eigenvalues and eigenpolynomials can be used to construct a polynomial approximation of the PCE for the random eigenvalues.
   - for continuous and separated $\lambda_i(\xi)$ in $\mathbb{C}$.

Sonday *et al.*, *SISC*, in press; Berry *et al.*, in review

# 1D Example



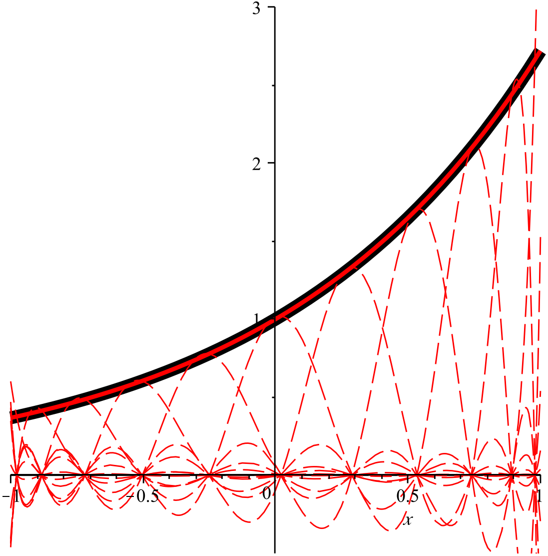

$$\dot{x}(\xi,t) = a(\xi)x(\xi,t); \quad \xi(\omega) \sim U[-1,1];$$

$$J = a(\xi) \equiv \begin{cases} \xi + 1 & \text{for} \quad \xi \geq 0, \\ \xi - 1 & \text{for} \quad \xi < 0. \end{cases}$$

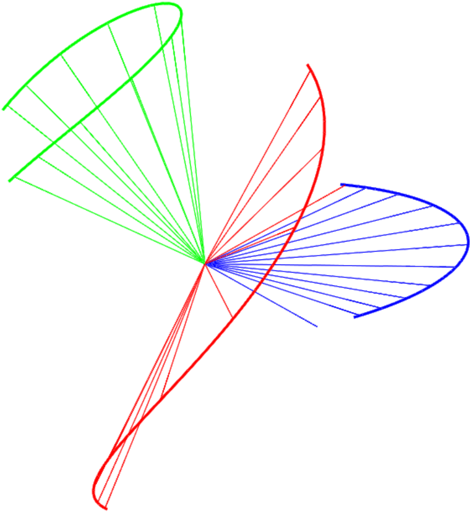
$$\widetilde{W}(J) = [-2,-1] \cup [1,2]; \qquad \text{conv}(\widetilde{W}(J)) = [-2,2].$$

LU PC: eigenvalues of $\mathcal{J}^P$ shown for $P = 10, 15, 20, 25, 45$

# Eigenpolynomials approximate the PCE of $\lambda(\xi)$



# Stochastic Vectors composed of Galerkin eigenvectors approximate the stochastic eigenvectors well



# CO Oxidation Example

The oxidation of CO on a surface can be modeled as (Makeev et al., JCP, 2002)

$$\dot{u} = az - cu - 4duv \qquad \dot{v} = 2bz^2 - 4duv$$

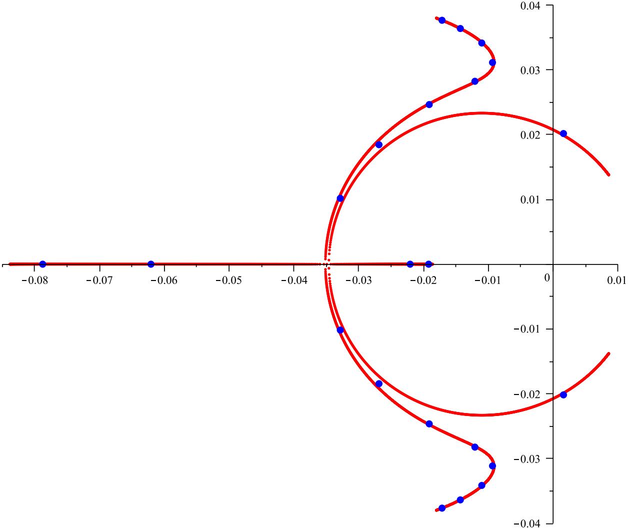
$$\dot{w} = ez - fw \qquad z = 1 - u - v - w$$

$$a = 1.6, b = 20.75 + .45\xi, c = 0.04, \quad d = 1.0, e = 0.36, f = 0.016$$

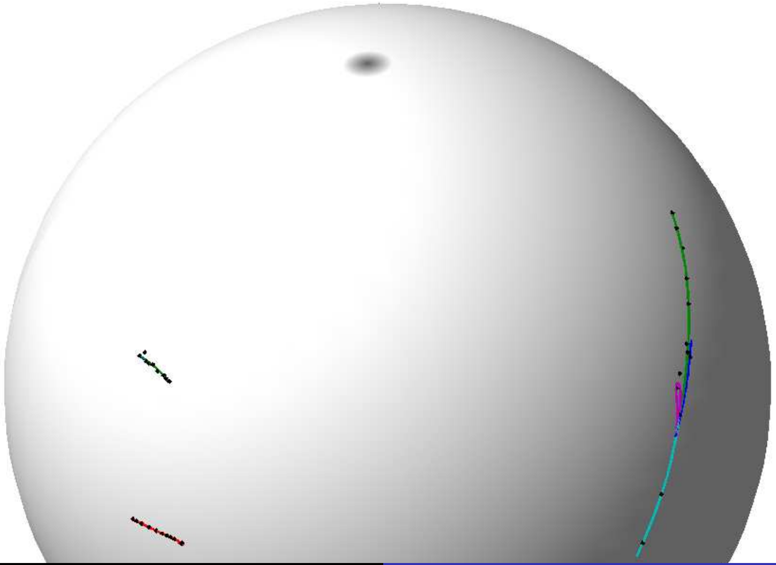
$$u(0) = 0.1, v(0) = 0.2, w(0) = 0.7$$

exhibits Hopf bifurcations for $b \in [20.3, 21.2]$

# CO Oxidation: PC order 10. Slow eigenvalues.



# CO Oxidation: PC order 10. Eigenvectors.



# Data Free Inference (DFI)

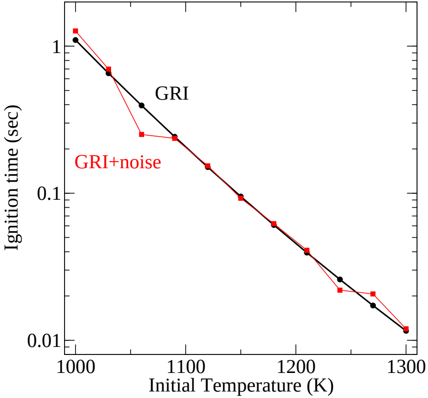
- Input uncertainties are not well characterized in many practical network models
- May have nominal parameter values and bounds
  - No information on correlations
  - No joint PDF on parameters
- Joint PDF structure can have a drastic effect on resulting uncertainties in predictions
- When original raw data is available, Bayesian inference provides the requisite posterior
- When original data is **not** available, what can be done?
  - DFI: discover a consensus joint PDF on the parameters consistent with given information

    (Berry *et al.*, JCP, in review)
  - Demonstrate on a chemical ignition problem (ODE)

# Generate ignition "data" using a detailed model+noise

- Ignition using a detailed chemical model for methane-air chemistry
- Ignition time versus Initial Temperature
- Multiplicative noise error model
- 11 data points:

$$d_i = t^{\mathsf{GRI}}_{\mathrm{ig},i}(1 + \sigma\epsilon_i)$$
$$\epsilon \sim N(0, 1)$$



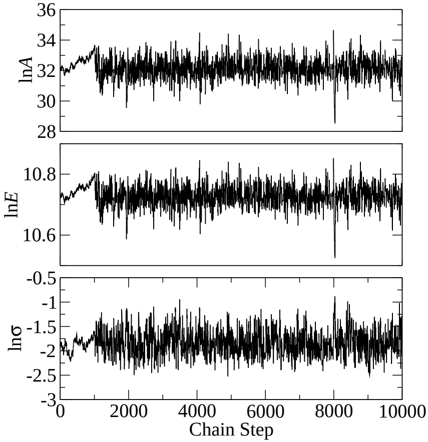
# Fitting with a simple chemical model

- Fit a global single-step irreversible chemical model

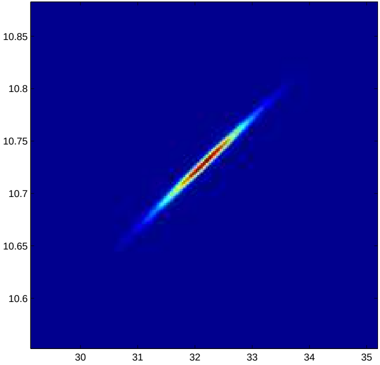
$$CH_4 + 2O_2 \rightarrow CO_2 + 2H_2O$$

$$\begin{aligned} \mathfrak{R} &= [CH_4][O_2]k_f \\ k_f &= A\exp(-E/R^oT) \end{aligned}$$

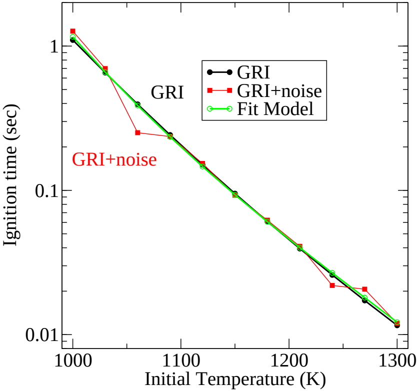
- Infer 3-D parameter vector $(\ln A, \ln E, \ln \sigma)$
- Good mixing with adaptive MCMC when start at MLE



# Bayesian Inference Posterior and Nominal Prediction

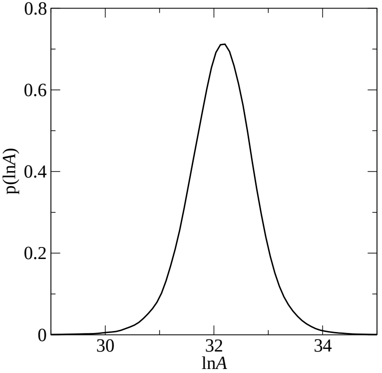



Marginal joint posterior on $(\ln A, \ln E)$ exhibits strong correlation

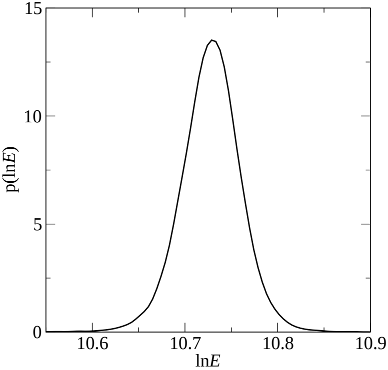



Nominal fit model is consistent with the true model

# Marginal Posteriors on $\ln A$ and $\ln E$




$$\ln A = 32.15 \pm 3 \times 0.61$$




$$\ln E = 10.73 \pm 3 \times 0.032$$

# Data Free Inference Challenge

Discarding initial data, reconstruct marginal $(\ln A, \ln E)$ posterior using the following information

- Form of fit model
- Range of initial temperature
- Nominal fit parameter values of $\ln A$ and $\ln E$
- Marginal 5% and 95% quantiles on $\ln A$ and $\ln E$

Further, for now, presume

- Multiplicative Gaussian errors
- $N = 8$ data points

# DFI Algorithm Structure

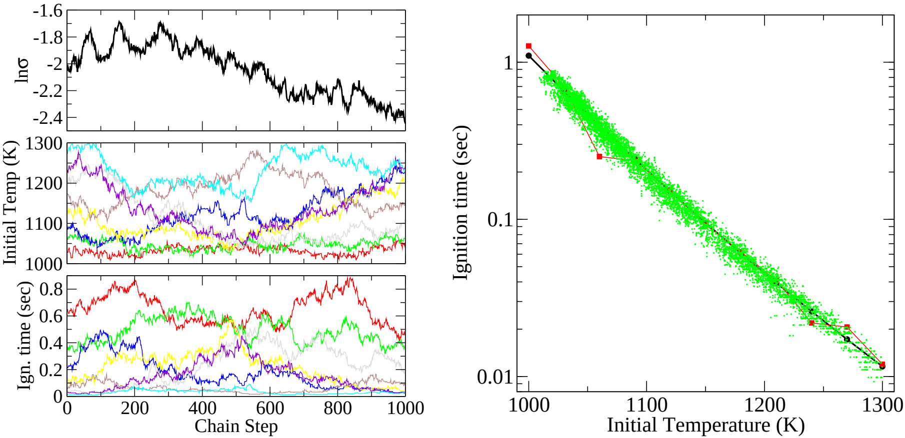
Basic idea:

- Explore the space of hypothetical data sets
- Accept data sets that lead to posteriors that are consistent with the given information
- Evaluate pooled posterior from all acceptable posteriors

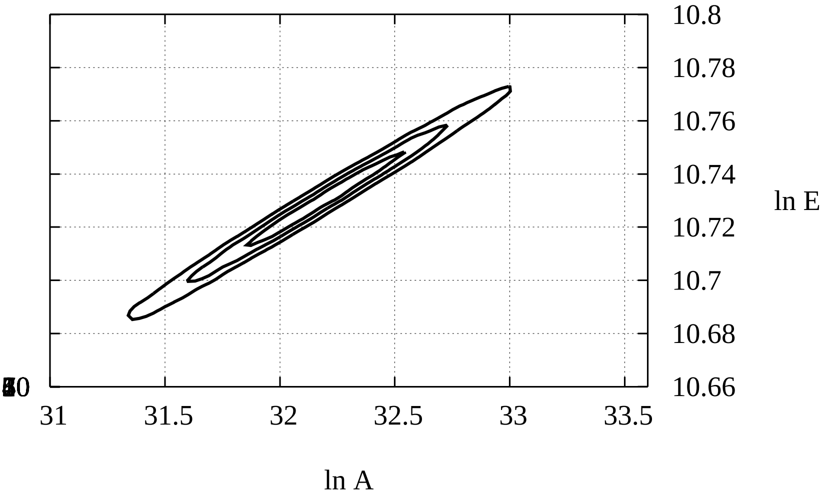
Algorithm uses two nested MCMC chains

- An outer chain on the data, $(2N+1)$–dimensional
  - $N$ data points $(x_i, y_i)$ + $\sigma$
  - Likelihood function captures constraints on parameter nominals+bounds
- An inner chain on the model parameters
  - Likelihood based on fit-model
  - parameter vector $(\ln A, \ln E, \ln \sigma)$

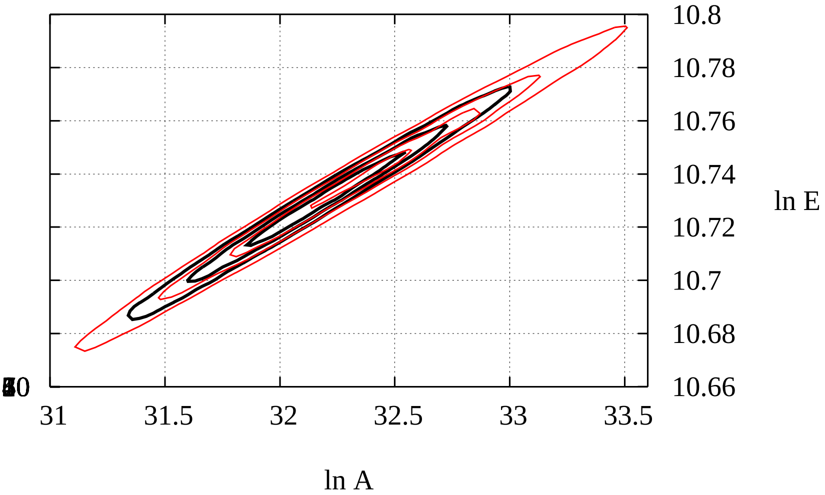
# Short sample from outer/data chain



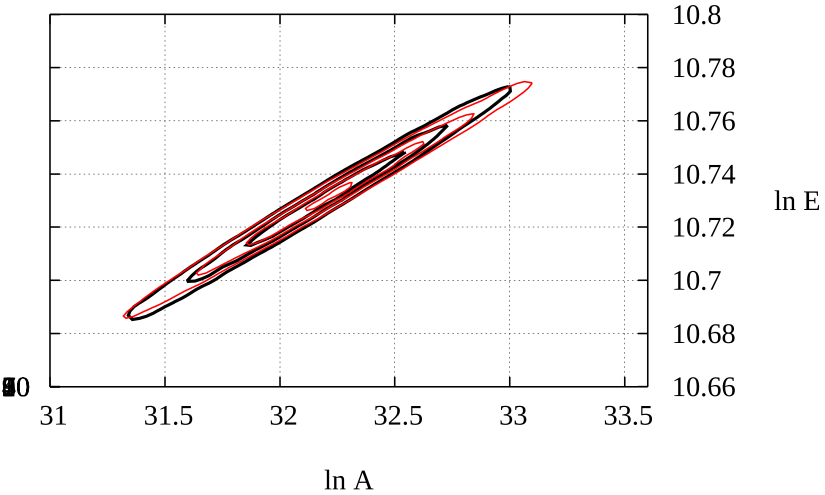
# Reference Posterior – based on actual data



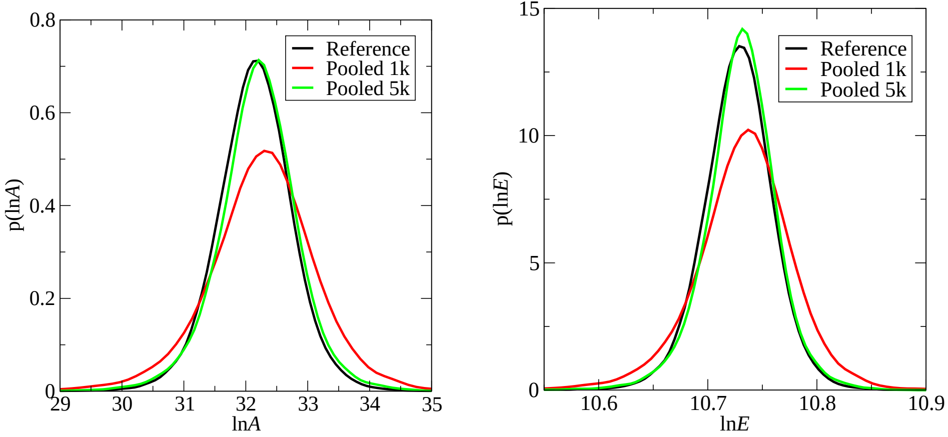
# Ref + DFI posterior based on a 1000-long data chain



# Ref + DFI posterior based on a 5000-long data chain



# Marginal Pooled DFI Posteriors on $\ln A$ and $\ln E$



# Closure

- Analysis of uncertain network model dynamics:
  - Outlined relationship between eigen-analysis of a sampled stochastic ODE system and the Galerkin PC system.
  - Galerkin system eigenvalues/eigenvectors can be used to analyze the dynamics of the stochastic system
  - Work in progress on
    - associated stochastic model reduction strategies
    - structural uncertainty in network models
- Data Free Inference:
  - Developed a DFI procedure for estimation of self-consistent parametric posteriors in the absence of data
  - Demonstrated effective and convergent estimation of missing posterior in a chemical ignition problem
  - In progress: algorithm optimization and generalization to handle a range of different constraints