SAND2018-7829C

Sandia National Laboratories

Chemical Model Comparison via Bayesian Model Evidence

Cem Gormezano, Department of Mechanical Engineering, McGill University
Tiernan Casey, Habib Najm, Khachik Sargsyan, Org. 08351, Reacting Flow Research, Cosmin Safta, Org. 08754 Quantitative Modeling & Analysis

Objectives

The goal of this work is to explore the use of Bayesian model evidence for comparing chemical models of different complexity/size, with given parametric uncertainty. The comparison will be done employing ignition delay time for Hydrogen – Oxygen reaction as a reference. The objective is to explore the utility of this approach for providing judgement on the inclusion of additional reactions in a given chemical model.

Hydrogen- Oxygen Reaction

- A chemical model is a thermodynamically feasible step by step sequence of elementary reactions by which the overall change occurs.

Elementary Reactions

$H + O_2 \rightleftharpoons OH + O$
$O + H_2 \rightleftharpoons OH + H$
$OH + H_2 \rightleftharpoons H + H_2O$
$H_2 + M \rightleftharpoons H + H + M$
$O + H_2O \rightleftharpoons OH + OH$
…

$$K(T) = AT^n \exp\left(-\frac{E_a}{RT}\right)$$

$$\vec{\lambda} = [\ln(A), n, E_a]$$

$$\frac{d\vec{X}}{dt} = \vec{f}(\vec{X}, T, \vec{\lambda}) \qquad \vec{X}(0) = \vec{X}_0 \qquad T(0) = T_0$$

$$\frac{dT}{dt} = g(\vec{X}, T, \vec{\lambda})$$

- In Hydrogen combustion, there are eight major species and at least 16 elementary reactions.
- The reaction rate parameters are subject to uncertainty. However, given a reaction rate mechanism one can estimate the the reaction rate parameters by using experimental data on the observables of a system.
- A function "F" is constructed by solving the above dynamical system with the initial values specified by initial temperature and equivalence ratio and computing the ignition time.

$$t_{ignition} = F(T_0, \phi; \vec{\lambda}, M)$$

- Given an experimental data of $(T_o, \phi, t_{ignition})$, which chemical model is most useful for prediction ?

Parameter Estimation & Model Comparison

Bayesian Inference

Bayesian view defines probability of a proposition as the degree of belief or plausibility that the proposition is true based on the evidence at hand.

$$P(\lambda|D,M) = \frac{P(D|\lambda,M)}{P(D|M)} P(\lambda|M)$$

$$P(\lambda|D,M) \propto P(D|\lambda,M) P(\lambda|M)$$

M: Model, D: Data, λ: Parameter

$P(\lambda|D,M)$: Posterior Distribution
$P(\lambda|M)$: Prior Distribution
$P(D|\lambda,M)$: Likelihood
$P(D|M)$: Evidence

Prior probability represents our state of knowledge before the current data is analyzed. This is refined by the experimental data through the likelihood , and yields the posterior probability, representing our state of knowledge in light of the data. In other words, Bayes theorem encapsulates the process of learning.

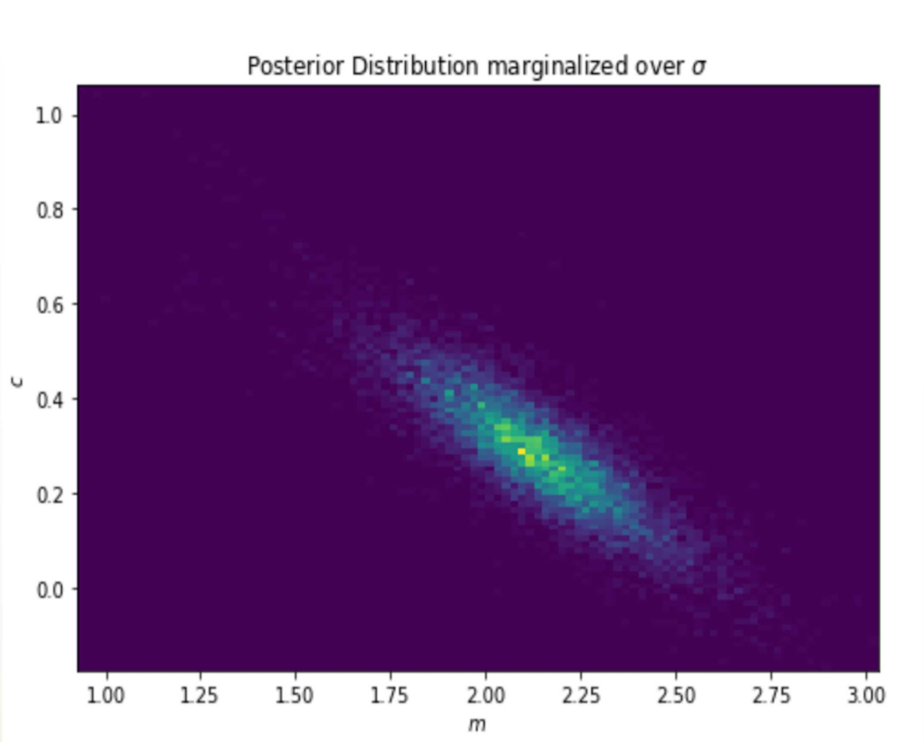
Parameter Estimation Example (Line Fitting)

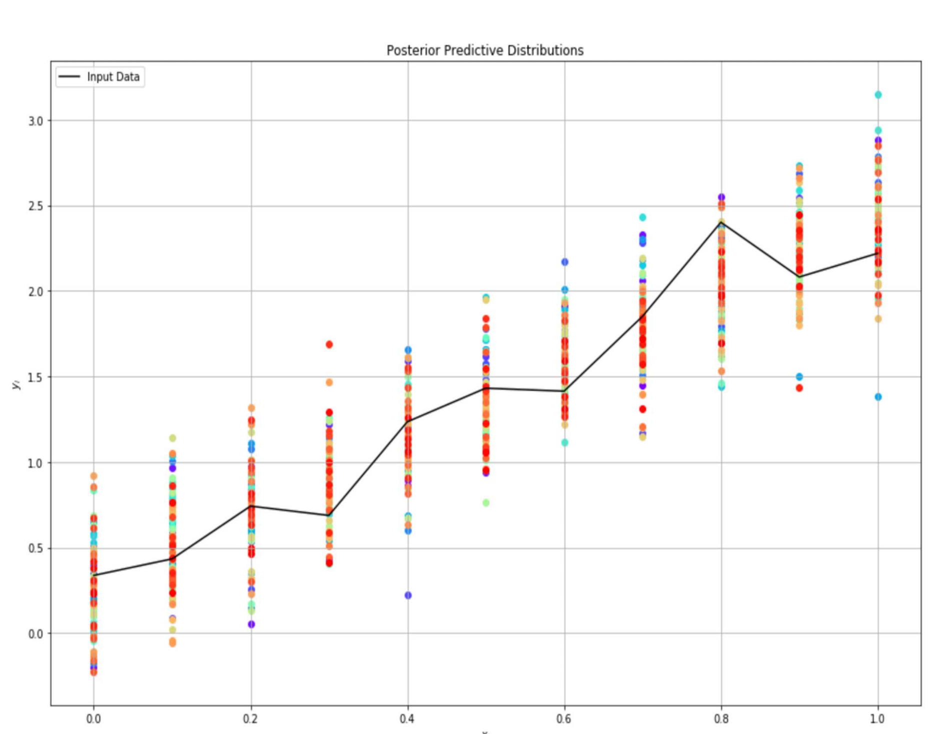
Data: $(x_i, y_i)\ i = 1, 2 .. n$ Model: $y = mx + c + \sigma\epsilon, \quad \epsilon \sim N(0,1)$ Parameters: (m, c, σ)

$$\epsilon_i = \frac{y_i - (mx_i + c)}{\sigma} \qquad P(D|m,c,\sigma,M) = \prod_{i=1}^{n} P(X = \epsilon_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-\epsilon_i^2)$$

If there is no information on the parameters before the data is observed, one can use an uninformed prior distribution . Through normalization one can obtain the posterior such that,

$$P(\lambda|D,M) = \frac{P(D|\lambda,M)P(\lambda|M)}{\int_{-\infty}^{\infty} P(D|\lambda,M)P(\lambda|M)\, d\lambda}$$





Model Comparison

$$\text{Bayes factor} = \frac{P(M_1|D)}{P(M_2|D)} = \frac{P(M_1)}{P(M_2)} \frac{P(D|M_1)}{P(D|M_2)} \qquad P(D|M) = \int_{-\infty}^{\infty} P(D|\lambda,M)P(\lambda|M)\, d\lambda$$

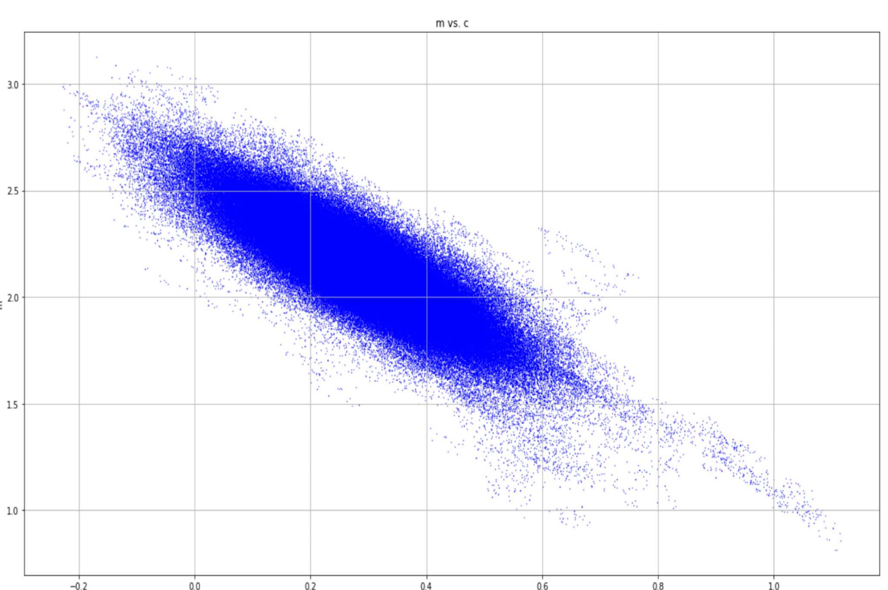
Bayes factor offers a mathematical ground to compare two chemical models in terms of how well they are supported by the data.

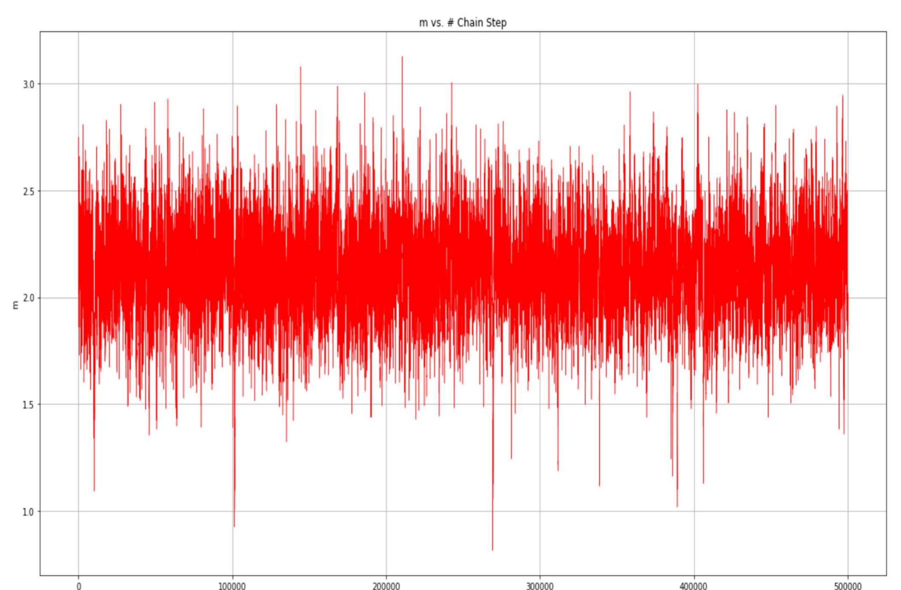
MCMC (Markov Chain Monte Carlo) Sampling

- Evaluating the posterior on a grid becomes increasingly expensive as the dimensionality of the parameter space increases. In many applications, rather than the posterior itself, characteristics such as the mean, the standard deviation and the moments are of interest. Monte Carlo sampling methods make use of the law of large numbers in order to calculate these quantities. The properties of interest (μ, σ …) are calculated from the samples of the posterior and not by integrating the posterior.
- It is counterintuitive to draw samples from the posterior when all one knows is the likelihood. However, this can be achieved through constructing a Markov chain whose equilibrium distribution is the posterior distribution.
- Adaptive Metropolis- Hastings Algorithm:
  - The information of posterior is incorporated into the Markov chain through the acceptance probability:

$$p = \min\left(\frac{P(D|\lambda_{n+1})Q(\lambda_n|\lambda_{n+1})}{P(D|\lambda_n)Q(\lambda_{n+1}|\lambda_n)}, 1\right)$$

  - The proposal distribution is adaptive and must be tuned carefully.
  - The initial steps of the chain are "burnt in" since they are not good representations of the posterior. The adaptive scheme uses the entire chain history to adapt the proposal distribution.





Polynomial Surrogate Construction

- The MCMC procedure requires many evaluations of the likelihood. When computing the likelihood involves execution of a complicated computational model, successive evaluations of the likelihood become expensive.
- Polynomial Chaos (PC) expansion can be employed to approximate the computational model in a specific range.

$$F(T_o, \phi, \vec{\lambda}) = \sum_{k=1}^{K-1} c_k \prod_{i=1}^{d} \psi_{k,i}(\xi_i)$$

- The coefficients c_k can be calculated by either quadrature or least squares approximation. Generally, quadrature would require less samples depending on the dimensionality. Moreover, an independent Bayesian inference problem can be formulated for the coefficients of the expansion.
- The accuracy of the PC expansion depends on the dimensionality and structure of the computational model, ranges of variables and the size of the training set (# of samples).
- Higher order PC expansion yields a better approximation as long as the PC expansion does not over fit the data.
- In order to avoid running TChem many times during the MCMC process, UQTk and TChem will be coupled to build a PC surrogate of "F". At each step of the MCMC, the surrogate will be evaluated. If MCMC chain drifts away from the range of the surrogate, TChem will be called.

Future Work

The reaction rate parameters of the $H + O_2 \rightleftharpoons OH + O$ elementary reaction will be learned from $(T_o, \phi, t_{ignition})$ data via Bayesian Inference. This data will be generated artificially using a different chemical model than the one inference assumes. Posterior of rate parameters will be constructed via MCMC procedure. Within MCMC, a PC surrogate will replace the TChem solver for evaluating likelihood. Finally, the inference problem will be performed in other chemical models and their Bayes factor will be calculated.