Optimal Design and Model Validation for Combustion Experiments in a Shock Tube

Quan Long Daesang Kim Raúl Tempone
Fabrizio Bisetti Aamir Farooq Omar Knio Serge Prudhomme

Computer, Electrical and Mathematical Sciences & Engineering, King Abdullah University of Science and Technology, KSA
Clean Combustion Research Center, King Abdullah University of Science and Technology, KSA

quan.long@kaust.edu.sa, daesang.kim@kaust.edu.sa, raul.tempone@kaust.edu.sa,
fabrizio.bisetti@kaust.edu.sa, aamir.farooq@kaust.edu.sa, omar.knio@kaust.edu.sa,
serge.prudhomme@kaust.edu.sa

Introduction

We develop a Bayesian framework for the optimal experimental design of the shock tube experiments which are being carried out at the KAUST Clean Combustion Center. The unknown parameters are the pre-exponential parameters and the activation energies in the reaction rate functions. The control parameters are the initial hydrogen concentration and the temperature. First, we build a polynomial based surrogate model for the observable related to the reactions in the shock tube. Second, we use a novel MAP based approach to estimate the expected information gain in the proposed experiments and select the best experimental set-ups corresponding to the optimal expected information gains. Third, we use the synthetic data to carry out virtual validation of our methodology.

H_2 oxidation in shock tubes

- The goal of the H_2 oxidation experiment is to determine the rate constants k of H+O_2 $\leftrightarrow$ OH+O. The rate constant k has two parameters, pre-exponential factor A and activation energy E_a and is evaluated by the Arrehnius equation $k(T) = A\exp(-E_a/RT)$.
- The maximum slope from the graph of the measured H_2O determines a value of the reaction rate constant k through a relation between the maximum slope and k provided by a simulation model.
- Fuel mixtures consist of H_2 (0.5% - 5%) and O_2 (0.1%) with a diluent of Ar. The temperature changes over a range from 1100 K to 1500 K and pressure is fixed at 2 atm.
- SENKIN is a Fortran program for chemical kinetics simulations of a homogenous gas mixtures in a closed system and is applied to H_2 oxidation in shock tubes with the reaction mechanism from Hong et al (2011) containing 10 chemical species and 20 reactions.

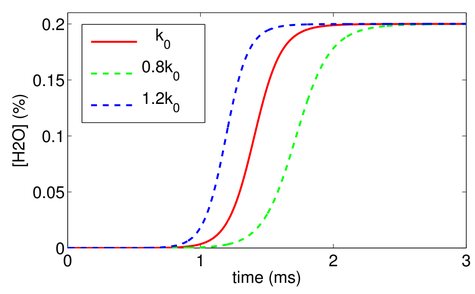


Figure 1: *Time histories of H_2O concentrations from three SENKIN simulations at 1300 K and 2 atm with an intial fuel mixture of 2.75% H_2, 0.1% O_2, and 97.15% Ar. Three simulations have different reaction rate constants, the nomial value k_0 from the reaction mechanim file, Hong et al (2011) and $\pm$20% deviations ($0.8k_0$ and $1.2k_0$) from the nominal value. The maximum slopes from the simulations are respectively, 3.90, 4.91, and 5.91 at $0.8k_0$, $1.0k_0$, and $1.2k_0$.*

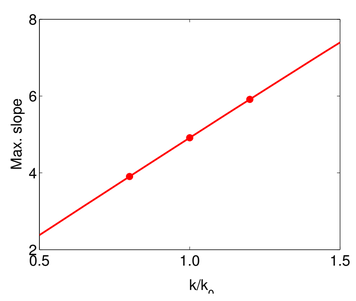


Figure 2: *The solid curve shows the behavior of the maximum slope of [H2O] over a range of k from $0.5k_0$ to $1.5k_0$ at the same condition for the above Figure 1. The three points represent the values from the three simulations of Figure 1. The graph shows an almost linear relation between k and maximum slope of [H_2O].*

Adaptive pseudo spectral analysis

- The simulaion model can be approximated by a truncated polynomial chaos expansion

$$g(\boldsymbol{\xi}) \approx \sum_{k=0}^{P} c_k \Psi_k(\boldsymbol{\xi}) \tag{1}$$

where g is the simulation model, $\boldsymbol{\xi}$ is a vector in $\Omega = [-1, +1]^d$, P is the number of terms, and Ψ_k is multidimensional Legendre polynomials orthogonal with respect to an inner product

$$\langle f, g\rangle = \int_\Omega f(\boldsymbol{\xi}) g(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{2}$$

Numerical integration of $c_k = \langle g, \Psi_k\rangle / \langle \Psi_k, \Psi_k\rangle$ for the PC expansion requires quadrature grids and model realizations at the grid points.

- Simple and easy full tensor products of 1D quadrature rules can provide mulitidimensional quadrature grids for the numerical integration, but the cost of the computation increases exponetally with the dimension of the parameter ξ. The curse of dimension can be avoided by using an adaptive sparse quadrature with nested 1D quadrature rule of Gauss-Kronrod-Pattersen's.
- The pseudo spectral projection provides not only the sensitivities and uncertainty propagation between parameters and observables, but also a polynomial surrogate model of high efficiency.

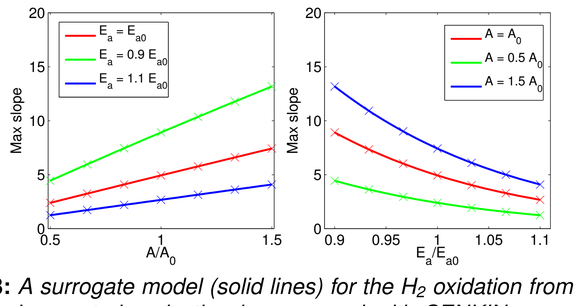


Figure 3: *A surrogate model (solid lines) for the H_2 oxidation from the adaptive pseudo spectral projection is compared with SENKIN computation results(crosses) for different A and E_a values at T = 1300 K and initial H_2 concentration of 2.75%.*

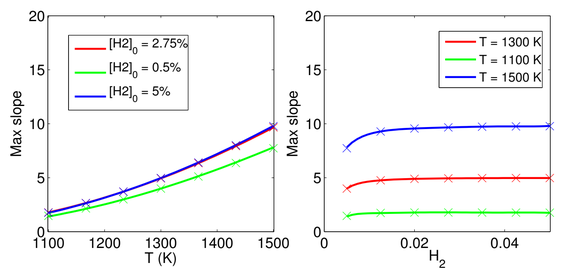


Figure 4: *For experimental design studies, possible design variables, temperature and initial H_2 concentration were added to the surrogate model. The surrogate model and SENKIN results are compared for different temperature and initial H_2 concentration at the nomial values, A_0 and E_{a0}.*

Experimental design

case 1: the design of a single experiment

Assume that the measured maximum slope data can be treated as the combination of the true model (the ODE model) and a Gaussian random noise:

$$y_i = g(\boldsymbol{\theta}, \boldsymbol{\xi}) + \epsilon_i \quad i = 1, ..., M\,. \tag{3}$$

where $\boldsymbol{\theta} = [A,\ E_a]^T$, $\boldsymbol{\xi} = [T,\ [H_2]]^T$, M is the number of repetitive experiments. We use our strategy proposed in [1, 2] to compute the expected information gain, i.e.,

$$I(T, [H_2]) = \int_y \int_\Theta \log\left(\frac{p(\boldsymbol{\theta}|\boldsymbol{y})}{p(\boldsymbol{\theta})}\right) p(\boldsymbol{\theta}|\boldsymbol{y}) d\boldsymbol{\theta} p(\boldsymbol{y}) d\boldsymbol{y} \tag{4}$$

where $\boldsymbol{y} = \{y_i\}$ is the set of all M measurements.

Now, we design a single experiment, i.e., we are interested in optimizing one point in the control parameter space such that the I reaches its maximum. Since there is only one observable, it is essentially an under determined problem. We assume a uniform prior of $\boldsymbol{\theta} \sim U(\boldsymbol{\theta}_l, \boldsymbol{\theta}_u)$ with $\boldsymbol{\theta}_l = [0.5A_0,\ 0.9E_{a0}]^T$ and $\boldsymbol{\theta}_u = [1.5A_0,\ 1.1E_{a0}]^T$.

We use the following approximation [2] to compute I:

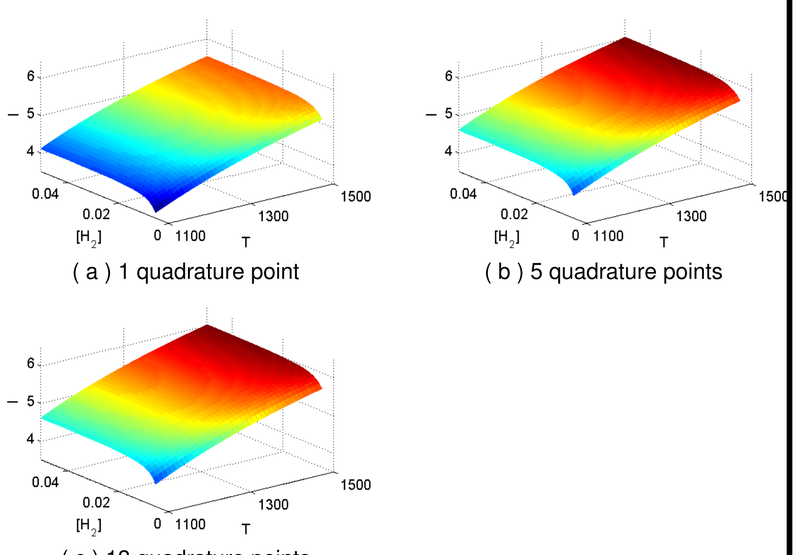
$$I = \int_\Theta \left[-\log\left[\tilde{p}_{\boldsymbol{s}}(\mathbf{0})\right] - \frac{1}{2}\log|\tilde{\boldsymbol{\Sigma}}_{\boldsymbol{s}|\boldsymbol{t}}| - \frac{r}{2}\log(2\pi) - \frac{r}{2}\right] p(\boldsymbol{\theta}_0) d\boldsymbol{\theta}_0 + O(1)\,. \tag{5}$$

where $\tilde{\boldsymbol{\Sigma}}_{s|t}$ is the post–experimental conditional covariance matrix, r is the dimension of the parameter vector $\boldsymbol{\theta}$ subtracted by the dimension of the indistinguishable manifold, here it equals 1. $\tilde{p}_s(\mathbf{0})$ is the approximated marginal prior. Additionally, the posterior conditional covariance matrix is computed by projecting the posterior covariance matrix of $\boldsymbol{\theta}$, i.e., $\boldsymbol{\Sigma}$, in the direction orthogonal to the indistinguishable manifold, that is

$$\tilde{\boldsymbol{\Sigma}}_{s|t} = \boldsymbol{U}^T \boldsymbol{\Sigma} \boldsymbol{U}\,,$$

where $\boldsymbol{U}$ is the matrix of eigenvectors corresponding to the nonzero eigenvalues of $\boldsymbol{\Sigma}$.

We use 1, 5 and 13 quadrature points to carry out the numerical integration in (5) and visualize the results in Figures 5(a), 5(b) and 5(c) respectively. Note that 5 quadrature points have already lead to converged results. We observe that the expected information gain increases as the temperature increases. In a small region near $[H_2] = 0.005$, the expected information gain decreases sharply together with the value of $[H_2]$. However, despite of this small region, the value of expected information gain is not very sensitive to a change in $[H_2]$, given a fixed value of T.



(a) 1 quadrature point (b) 5 quadrature points

(c) 13 quadrature points

Figure 5: *The expected information gain computed on a 40×40 discrete mesh in the control parameter domain. Increasing numbers of Gauss-Legendre quadrature points of Smolyak type are used in (a), (b) and (c).*

case 2: the design of multiple experiments

Next, we design two different experiments simultaneously, in which we use a single mixture and two different temperatures. Now our measurements can be expressed by

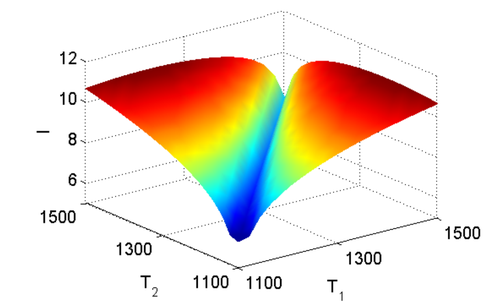
$$\boldsymbol{y}_i = \begin{bmatrix} y_{1_i} \\ y_{2_i} \end{bmatrix} = \begin{bmatrix} g(\boldsymbol{\theta}, \boldsymbol{\xi}_1) \\ g(\boldsymbol{\theta}, \boldsymbol{\xi}_2) \end{bmatrix} + \boldsymbol{\epsilon}_i\,, \tag{6}$$

where $\boldsymbol{\xi}_1 = [T_1, [H_2]]^T$ and $\boldsymbol{\xi}_2 = [T_2, [H_2]]^T$. Since the parameters now can be completely determined by the experiments (when $T_1 \neq T_2$). We use the following fast estimator developed in [1] for the expected information gain:

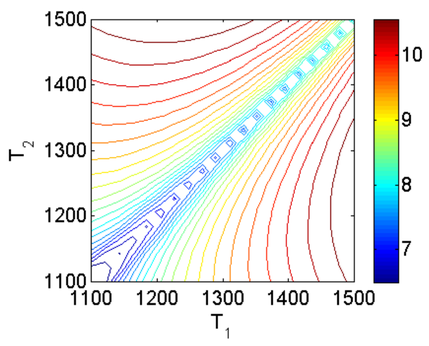
$$I = \int_\Theta \left[-\frac{1}{2}\log((2\pi)^d |\boldsymbol{\Sigma}|) - \frac{d}{2} - h(\boldsymbol{\theta}_0)\right] p(\boldsymbol{\theta}_0) d\boldsymbol{\theta}_0 + O\left(\frac{1}{M}\right)\,, \tag{7}$$

where $h(\boldsymbol{\theta}_0) = \log(p(\boldsymbol{\theta}_0))$.

We visualize, in Figure 6(a), the surface of expected information gain parameterized by T_1 and T_2, while $[H_2]$ is chosen as 0.05. We achieved the maximum information gain at $T_1 \approx 1226$ and $T_2 = 1500$ or its symmetrical location.



(a) The expected information gain



(b) Iso-contour of the expected information gain

Figure 6: *The expected information gain surface w.r.t. T_1 and T_2 and $[H_2]$ is a fixed value of 0.05.*

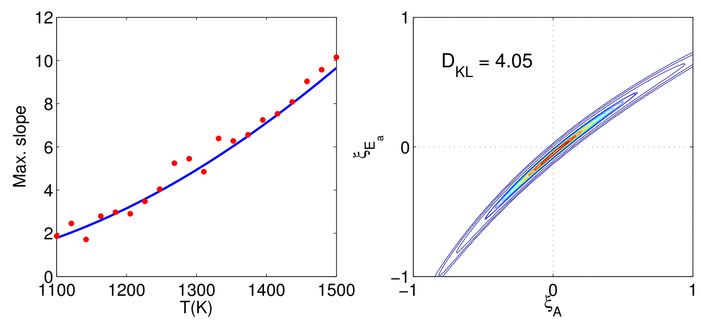


Figure 7: *Synthetic data of shock tube experiment are generated from (3) with $\sigma_\epsilon = 0.25$, using a naive design: 20 runs at uniform temperatures over the range from 1100 K to 1500 K (red points in the left). A posterior distribution of A and E_a is calculated from a uniform prior and Gaussian likelihood, and a contour plot of the posterior is shown (right). The information gain, D_{KL} is computed to be 4.05 for the generated synthetic data. The average of D_{KL} for 10 random synthetic data generations is 4.01.*

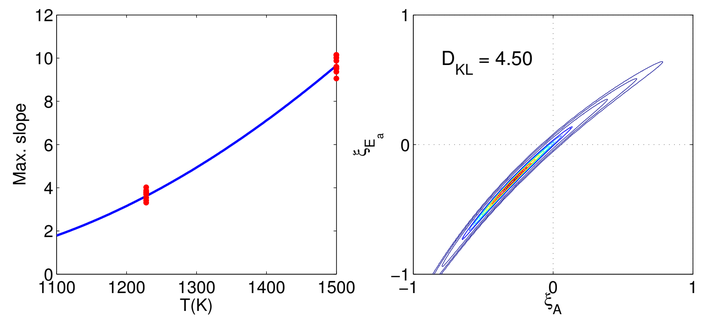


Figure 8: *Synthetic data are generated from the same model, (3) with $\sigma_\epsilon = 0.25$, for the previous plot, using the optimized design: 10 runs at 1226 K and 10 runs at 1500K (red points in the left). A contour plot of the computed posterior (right) from the identical uniform prior and Gaussian likelihood shows a narrower distribution of the posterior. The information gain, D_{KL} from the plot is 4.50 and the average of 10 random synthetic data generations is 4.44, which is approximately 10% greater than that from the naive design.*

Acknowledgement

Quan Long, Daesang Kim, Raúl Tempone, Fabrizio Bisetti, Omar Knio, and Serge Prudhomme are members of the KAUST SRI Center for Uncertainty Quantification in Computational Science and Engineering.

Future Work

- Validation using experimental data collected at the KAUST Clean Combustion Research Center and extending experimental design methodology to the cases where the control parameters are noisy.

References

[1] Q. Long, M. Scavino, R. Tempone, and S. Wang, *Fast estimation of expected information gains for bayesian experimental designs based on laplace approximations*, Computer Methods in Applied Mechanics and Engineering, 259 (2013), pp. 24–39.

[2] —, *A laplace method for under determined bayesian experimental design*, (2013). Working paper.