Bayesian optimal experimental design for the Shock-tube experiment

G Terejanu, C M Bryant and K Miki
Institute for Computational Engineering and Sciences, The University of Texas at Austin, 1 University Station C0200, Austin, TX, 78712, USA
E-mail: cbryant@ices.utexas.edu

Abstract. The sequential optimal experimental design formulated as an information-theoretic sensitivity analysis is applied to the ignition delay problem using real experimental. The optimal design is obtained by maximizing the statistical dependence between the model parameters and observables, which is quantified in this study using mutual information. This is naturally posed in the Bayesian framework. The study shows that by monitoring the information gain after each measurement update, one can design a stopping criteria for the experimental process which gives a minimal set of experiments to efficiently learn the Arrhenius parameters.

1. Introduction
A mathematical description of the gas-phase combustion kinetic problem is very challenging in context of Bayesian optimal experimental design. The system is composed of very stiff nonlinear ordinary differential equations (ODEs). Relevant studies on Bayesian optimal experimental design applied to such a problem are still sparse in literature, see Ref’s [1, 14, 7], and often test their models using “manufactured” data, which is generated using the model with artificial noise added.

Terejanu et al.[4] have shown that Bayesian experimental design can be formulated as an information-theoretic sensitivity analysis, where the optimal design maximizes the mutual information between the parameters and the future observations.

In this study, we present the application of the sequential BOED as presented in [4] to the ignition delay problem in the hydrogen-oxygen combustion. We explore the behavior of BOED when used to calibrate the model in the presence of model inadequacy and real data. It is shown that by closely monitoring the sequential experimental process using information theoretic measures such as information gain, one can perform a more efficient calibration using fewer experiments.

2. The Proposed Experimental Design Methodology
Given a set of observations \( D_{n-1} = \{ \tilde{d}_1, \tilde{d}_2, \ldots, \tilde{d}_{n-1} \} \) we are concerned with finding the next experimental design \( \xi_n \in \Xi \) such that the model parametric uncertainty is reduced after the experiment is performed and the associated measurement data \( \tilde{d}_n \) is collected. The mathematical
models used in this paper are generally represented by the following abstract model:

\[
\begin{align*}
\dot{\mathbf{r}}(\mathbf{u}, \Theta, \xi_n) &= 0 \\
\mathbf{d}_n &= \mathbf{d}(\mathbf{u}, \Theta, \xi_n) + \epsilon_{\text{model}}(\xi_n) + \epsilon_{\text{meas}}
\end{align*}
\]

Here \( \mathbf{u} \) is the state of the system which obeys the governing equations defined by \( \mathbf{r}(\cdot) \), \( \xi_n \in \Xi \) is the control variables associated with the experimental scenario, \( \Xi \) is the design space and \( \Theta \in \Theta \) are the model parameters, where \( \Theta \) is the parameter space of the model. Here, the model prediction \( \mathbf{d}_n \) calculated using the measurement model \( \mathbf{d}(\cdot) \) is comparable with the experimental data \( \hat{\mathbf{d}}_n \) for a particular scenario input \( \xi_n \). The random vector or random field \( \epsilon_{\text{meas}} \) is the observation noise due to sensor imprecision or the discretization of computer simulations. The random vector or random field \( \epsilon_{\text{model}}(\xi_n) \) captures discrepancy between model predictions and experimental data, and is due to missing physics in the governing equations.

When the objective of the experiment is to infer the parameters \( \Theta \), then the optimal experiment is given by maximizing the mutual information between the parameters and the model predictions, \( \mathbf{d}_n \), see Ref. [4]. In other words we would like to sample where the expected observations have the highest impact on model parameters.

\[
\xi^*_n = \arg \max_{\xi_n \in \Xi} \int_D \int_{\Theta} p(\Theta, \mathbf{d}_n|D_{n-1}, \xi_n) \log \frac{p(\Theta, \mathbf{d}_n|D_{n-1}, \xi_n)}{p(\Theta|D_{n-1})p(\mathbf{d}_n|D_{n-1}, \xi_n)} d\Theta d\mathbf{d}_n
\]

Compared with maximum entropy sampling for parameter estimation proposed in Ref.[16], the sampling based on maximizing the mutual information is more general as no assumption is made about the independence between model discrepancy and design variables. For this study, the estimation of the mutual information is done using the \( kNN \) (Nearest Neighbor) method proposed in Ref. [9].

3. Physical Model and Experimental data
Assuming that the temperature and pressure behind the reflected shock waves are constant during the reaction, the chemical reactions can be expressed in ordinary differential equations for each species (i.e. a well-stirred reactor model), see Ref.[6]. The forward rate coefficient of the \( r \)-th reaction, \( k_{f,r} \), is calculated using the modified Arrhenius equation: \( k_{f,r} = A_r T_r^{m_r} \exp(-\frac{\Theta_r}{T_r}) \), \( r = 1, \ldots, N_r = 35 \), where \( A_r \) and \( m_r \) are the pre-exponential Arrhenius parameters and \( \Theta_r \) is the characteristic temperature of the activation energy. We use the detailed chemical kinetic mechanism previously verified in [6] to model the ignition delay of \( \text{O}_2/\text{H}_2/\text{Ar} \) mixture behind the reflected shock wave. This mechanism considers 35 elementary reactions and 11 species.

Experimental data were recently collected at the High Temperature Gasdynamics Laboratory at Stanford University [6]. The data is composed of six profiles of the \( \text{H}_2\text{O} \) concentration history behind reflected shock waves over the temperature range 1100 - 1472 K. All six experimental scenarios have an initial \( \text{O}_2 \) concentration of 0.001%. Three have an initial \( \text{H}_2 \) concentration of 0.029% with \( T_5 = \{1100, 1197, 1256\} \)K and corresponding \( P_5 = \{1.95, 1.84, 2.01\} \) atm. The other three have initial \( \text{H}_2 \) concentration of 0.009% and \( T_5 = \{1317, 1448, 1472\} \)K and \( P_5 = \{1.91, 1.85, 1.83\} \) atm. Given our assumption that the model discrepancy is not time dependent we choose five data points where the level of the error is assumed to be constant. The detailed description of the shock tube experiment is provided in [5].

4. Description of Stochastic Models
In the current study we treat the pre-exponential Arrhenius parameters, \( A_r \) and \( m_r \), and the characteristic temperature, \( \Theta_r \), as uncertain and thus opened for calibration. The prior
distribution for these parameters is given by $A_r \sim \mathcal{U}(0.64289, 1.21429)$, $m_r \sim \mathcal{U}(0, 1)$, $\Theta_r \sim \mathcal{U}(0, 2)$. The measurement model for the reaction is assumed to have the following probabilistic formulation,

$$
\tilde{d}(t) = d(A_r, m_r, \Theta_r, t) + \epsilon_{\text{model}} \tag{3}
$$

where $d(A_r, m_r, \Theta_r, t)$ is the concentration of $H_2O$ given by the kinetic mechanism and $\tilde{d}(t)$ is the experimental observation given in Ref.[6]. The discrepancy between model predictions and observations, $\epsilon_{\text{model}}$, accounts for the small experimental uncertainty and any missing physics in the governing equation. Kennedy & O’Hagan [8] have emphasized the importance of model discrepancy, $\epsilon_{\text{model}}$, in Bayesian calibration when dealing with approximate models of complex mathematical models. In this work, we model this model discrepancy using an additive Gaussian model, $\epsilon_{\text{model}} \sim \mathcal{N}(0, \sigma^2)$. The hyper-parameter $\sigma$ of the Gaussian model will also be inferred from the data with prior distribution for $\sigma$ given by $\sigma \sim \mathcal{U}(0, 1)$.

5. Results and Discussions

The estimated mutual information between model parameters and the model predictions are shown in Table 1. At some stages, stage 1 in particular, the mutual information is nearly identical for multiple experiments. This may indicate that any one of these experiments could be added to inform the parameters.

Table 1. Mutual information at each design stage and the optimal sequence of designs. In bold are the maximums found at each stage of the design process.

| Index | stage 1 | stage 2 | stage 3 | stage 4 | stage 5 | stage 6 |
|-------|---------|---------|---------|---------|---------|---------|
| 1     | 0.454   | 0.429   | 1.500   | 1.375   | 1.212   |         |
| 2     | 0.441   | 0.306   | 1.517   |         |         |         |
| 3     | 0.434   | 0.466   |         |         |         |         |
| 4     | 0.422   | 0.411   | 1.060   | 1.071   | 0.897   | 0.882   |
| 5     | 0.451   | 0.458   | 1.074   |         |         |         |
| 6     |         |         |         |         |         |         |

Additionally Table 1 also lists the information gain after each stage of the process. A large number signifies that the posterior has dramatically changed with respect to the prior given the experimental data. We can see that there is a convergence in the information gain which is reached after the fourth stage. This convergence signals that the inference process is slowing down and that additional data provides less information compared with the first data points.

Posterior means and marginal 95% confidence intervals for the four parameters are shown in Figure 1. Note that for $\Theta_r$ and $\sigma$ the mean value remains relatively unchanged after stage four, see Figs 1(a),1(b). In addition the confidence intervals for both parameters narrow initially and then remain relatively stable as additional experiments are added. This indicates a convergence of the distributions to a stable representation of the uncertainty for these parameters.

On the other hand, the results for $A_r$ and $m_r$ are not as stable, see Figs1(c),1(d). While the mean values seem to fluctuate only slightly after two stages, the confidence intervals only narrow significantly after the sixth stage of the design process. Even though in the 5th stage of optimal experimental deign, there is an increase in the marginal uncertainty for $A_r$ and $m_r$, see Figs.1(c),1(d), there is also an increase in the correlation between these two parameters, see
Fig. 1(e), which results in an overall decrease in the uncertainty given by the joint distribution of $A_r$ and $m_r$.

6. Conclusions
In this paper the sequential Bayesian optimal experimental design as presented by Terejanu et al. 2012 is applied to an ignition delay problem in hydrogen-oxygen combustion. Monitoring the sequential experimental process using information theoretic measures we are able to determine an efficient calibration using fewer experiments. Model inadequacy is also taken into account in the methodology. More work needs to be done to propagate this uncertainty to the reaction rate of interest.

Acknowledgments
This work has been supported by the Department of Energy [National Nuclear Security Administration] under Award Number [DE-FC52-08NA28615].

References
[1] Blau G, Lasinski M, Orcun S, Hsu S H, Caruthers J, Delgass N and Venkatasubramanian V. \textit{Comput. Chem. Eng.} 32 (2008) 971–989.
[2] Cheung S H and Beck J L. \textit{Computer-Aided Civil and Infrastructure Engineering} 25 (2010) 304–321.
[3] Franceschini G and Macchietto S. \textit{Chem. Eng. Sci.} 63 (2008) 4846–4872.
[4] Terejanu G, Upadhyay R R and Miki K. \textit{Exp. Therm. Fluid Sci} 36 (2012) 178–193.
[5] Hong Z, Barbour E A, Davidson A and Hanson R K. \textit{J. Chem. Phys.} 113 (2009) 12919–12925.
[6] Hong Z, Davidson D E, Barbour E A and Hanson R K. \textit{Proc. Combust. Inst.} 33 (2011) 309–316.
[7] Huan X and Marzouk Y M. \textit{J. Comput. Phys.} (2011) under review.
[8] Kennedy M C and O’Hagan A. \textit{J. Royal Statisti. Soc.: Series B (Statistical Methodology)} 63(3) (2001) 1467-9868.
[9] Kraskov A, Stogbauer H and Grassberger P. \textit{Phys. Rev. E} 69(6) (2004) 066138.
[10] Lindley D V. \textit{Ann. Math. Statist.} 27(4) (1956) 986-1005.
[11] Manion J A, et al. NIST chemical kinetics database, NIST standard reference database 17, version 7.0 (web version), release 1.4.3, data version 2008.12, National Institute of Standards and Technology, Gaithersburg, Maryland, 20899-8320. Web address: http://kinetics.nist.gov/.
[12] Masten D A, Hanson R K, Bowman C T. \textit{The Journal of Physical Chemistry} 91 (1990) 7119–7128.
[13] Miki K, Cheung S H, Prudencio E E and Varghese P L. \textit{International Journal of Chemical Kinetics} (2012), to appear.
[14] Mosbach S, Braumann A, Man P L W, Kastner C A, Brownbridge G P E and Kraft M. \textit{Combust. Flame} doi:10.1016/j.combustflame.2011.10.019 (2011).
[15] Pirraglia A N, Michael J V, Sutherland R B and Klemm J. \textit{The Journal of Physical Chemistry} 93 (1989) 282–291.
[16] Sebastini P and Wynn H P. \textit{J. R. Statist. Soc. B} 62(1) (2000) 145-157.
[17] Sheen D and Wang H. \textit{Combust. Flame} 158 (2011) 2358–2374.
[18] Smith G P, et al. GRI-Mech 3.0. available at http://www.me.berkeley.edu/grimech.