AIAA 2011–0513

49th AIAA Aerospace Sciences Meeting; Orlando, FL; January 2011.

Optimal Bayesian Experimental Design for Combustion Kinetics

Xun Huan* and Youssef Marzouk†

Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

Experimental diagnostics play an essential role in the development and refinement of chemical kinetic models, whether for the combustion of common complex hydrocarbons or of emerging alternative fuels. Questions of experimental design—e.g., which variables or species to interrogate, at what resolution and under what conditions—are extremely important in this context, particularly when experimental resources are limited. This paper attempts to answer such questions in a rigorous and systematic way. We propose a Bayesian framework for optimal experimental design with nonlinear simulation-based models. While the framework is broadly applicable, we use it to infer rate parameters in a combustion system with detailed kinetics. The framework introduces a utility function that reflects the expected information gain from a particular experiment. Straightforward evaluation (and maximization) of this utility function requires Monte Carlo sampling, which is infeasible with computationally intensive models. Instead, we construct a polynomial surrogate for the dependence of experimental observables on model parameters and design conditions, with the help of dimension-adaptive sparse quadrature. Results demonstrate the efficiency and accuracy of the surrogate, as well as the considerable effectiveness of the experimental design framework in choosing informative experimental conditions.

I. Introduction

An increasingly broad range of fuels, derived from diverse feedstocks, are shaping future transportation and power generation systems. In the realm of ground transportation, biofuels are entering the fuel supply in blends with conventional petroleum-derived fuels and as primary energy carriers. Alternative aviation fuels—synthesized through gasification and liquefaction processes from biomass, coal, natural gas, tar sands, and oil shale—are of great interest to engine and airframe manufacturers. Environmental requirements often dictate that these fuels be coupled with high-efficiency energy conversion technologies designed to minimize pollutant emissions. Emerging strategies include low-temperature combustion in homogeneous charge compression ignition (HCCI) engines, direct injection in spark-ignition (SI) engines, or staged ultra-lean combustion in gas turbines. Yet the consumption of alternative fuels in these systems involves chemical species and blends whose physical properties and combustion kinetics are poorly characterized. Additionally, new engines and propulsion technologies may operate in low temperature or extreme pressure regimes where current kinetic models—even for widely-used fuels, much less alternative fuels—are not validated.

Myriad sources of data contribute to reaction mechanisms and thermophysical properties. Ignition experiments, flow reactor experiments, and flame speed measurements are important, yet indirect, targets for developing and validating kinetic mechanisms. Optical diagnostics are essential to focused experimental studies of gas-phase combustion kinetics, but high pressure environments and direct interrogation of large molecules present significant challenges. *Ab initio* calculations, on the other hand, are limited in scale, and results often must be extrapolated to larger species or to faraway pressures and temperatures, with unknown levels of confidence. The rapid development of new fuels and new conversion technologies urges *more systematic* approaches to model development that synthesize all these sources of kinetic information, including indirect data such as ignition delays and flame speeds.

*Graduate Student, Department of Aeronautics and Astronautics, 77 Massachusetts Ave., 37-442, Cambridge, MA 02139, AIAA Student Member.

†Assistant Professor, Department of Aeronautics and Astronautics, 77 Massachusetts Ave., 33-305, Cambridge, MA 02139, AIAA Senior Member.

To facilitate such systematic approaches to model development, this paper focuses on **optimal experimental design.** We employ a Bayesian approach to constructing and calibrating kinetic models from experimental measurements. Bayesian statistics provides a rigorous foundation for inference from noisy, indirect, and incomplete data; a natural mechanism for incorporating physical constraints and heterogeneous sources of information; and a complete assessment of uncertainty at all stages of the model construction process. Given limited experimental resources, however, it is critical to choose the "best" set of observations or experimental conditions with which to probe a system. This choice encompasses not only questions of where and when to measure, but of which variables to interrogate. Optimal experimental design is particularly challenging in the context of combustion. Kinetic models contain very strong nonlinearities, rendering simple linear design theory inapplicable or ineffective. Moreover, these models may incorporate hundreds of uncertain rate parameters, with a complicated joint correlation structure. And repeatedly simulating canonical reacting flow experiments, in order to explore the variation of observables over candidate regions of parameter space, can be a computationally prohibitive undertaking.

We address these challenges as follows. *First,* we use the Bayesian setting to formulate an information-theoretic approach to optimal nonlinear experimental design, directly incorporating uncertainties in parameters and experimental measurements (§III). The optimal design is then the one yielding measurements that maximize information gain in particular kinetic parameters. This is quantitatively reflected through a utility function whose expectation must be maximized. *Second,* we introduce polynomial chaos surrogates to make the optimal design problem computationally tractable (§IV). Straightforward maximization of the expected utility involves evaluating high-dimensional integrals, where each single evaluation of the integrand may entail a time-dependent simulation. The polynomial chaos approach replaces these simulations entirely, in favor of an offline stochastic projection procedure. *Finally,* we combine these tools and demonstrate them in the context of shock tube ignition experiments—choosing optimal mixture conditions for the inference of kinetic parameters in a detailed hydrogen-oxygen mechanism (§V). An assessment of the current results and a summary of future directions is given in §VI.

II. Ignition Problem

Ignition delays, measured via shock tube experiments, are frequently used as validation and calibration targets for gas-phase combustion kinetics.[1,2] Indeed, shock tube studies have been performed for an enormous variety of fuels, including jet fuels and many alternative fuels of current interest.[3,4] We thus use ignition data to demonstrate our experimental design methodology. (Note, however, that the formulation described in subsequent sections is quite general and can be extended to numerous other combustion observables.)

II.A. Governing Equations

We adopt a simple zero-dimensional ignition model, assuming adiabatic conditions and constant pressure. In particular, we consider ignition of a hydrogen-oxygen mixture using the 9-species 19-reaction detailed mechanism of Yetter *et al.*[5] as a nominal kinetic model. (In subsequent experiments, we will treat reaction rate parameters in this mechanism as uncertain, but we will not alter the model structure, i.e., the choice of reactions themselves.)

The state of the chemical system can be completely described by the species mass fractions $Y_1, \ldots, Y_{n_s}$ (where n_s is the total number of species) and the temperature T. Dynamics of the system are then given by the following ordinary differential equations (ODEs):

$$\frac{dY_j}{dt} = \frac{\dot{\omega}_j W_j}{\rho}, \; j = 1 \ldots n_s \tag{1}$$

$$\frac{dT}{dt} = -\frac{1}{\rho c_p} \sum_{n=1}^{n_s} h_n \dot{\omega}_n W_n \tag{2}$$

$$\begin{aligned} Y_j(t=0) &= Y_{j,0} \\ T(t=0) &= T_0 \end{aligned} \quad , \tag{3}$$

where $\dot{\omega}_j \left[\mathrm{kmol} \cdot \mathrm{m}^{-3} \cdot \mathrm{s}^{-1}\right]$ is the molar production rate of the jth species, $W_j \left[\mathrm{kg} \cdot \mathrm{kmol}^{-1}\right]$ is the molecular weight of the jth species, $\rho \left[\mathrm{kg} \cdot \mathrm{m}^{-3}\right]$ is the mixture density, $c_p \left[\mathrm{J} \cdot \mathrm{K}^{-1} \cdot \mathrm{kg}^{-1}\right]$ is the mixture specific heat

capacity under constant pressure, and h_n $\left[\mathrm{J}\cdot\mathrm{kg}^{-1}\right]$ is the specific enthalpy of the nth species. The molar production rate is defined in terms of elementary reaction rates as

$$\dot{\omega}_j \equiv \frac{dC_j}{dt} = \sum_{m=1}^{n_r} (\nu_{mj}^{''} - \nu_{mj}^{'}) \left(k_{f,m} \prod_{n=1}^{n_s} C_n^{\nu_{mn}^{'}} - k_{r,m} \prod_{n=1}^{n_s} C_n^{\nu_{mn}^{''}} \right), \tag{4}$$

where C_j $\left[\mathrm{kmol}\cdot\mathrm{m}^{-3}\right]$ is the molar concentration of the jth species, n_r is the total number of reactions, and $\nu_{mn}^{'}$ and $\nu_{mn}^{''}$ are the (dimensionless) stoichiometric coefficients on the reactant and product sides of the equation, respectively, for the nth species in the mth reaction.

The forward and reverse rate constants of the mth reaction, denoted by $k_{f,m}$ and $k_{r,m}$ respectively, are assumed to have the modified Arrhenius form:

$$k_{f,m} = A_m T^{b_m} \exp\left(\frac{-E_{a,m}}{R_u T}\right) \tag{5}$$

$$k_{r,m} = \frac{k_{f,m}}{K_{c,m}} = \frac{k_{f,m}}{\exp\left(\frac{-\Delta G_{T,m}^{o}}{R_u T}\right)}, \tag{6}$$

where A_m $\left[\left(\mathrm{m}^3\cdot\mathrm{kmol}^{-1}\right)^{-1+\sum_{n=1}^{n_s}\nu_{mn}^{'}}\cdot\mathrm{s}^{-1}\cdot\mathrm{K}^{-b_m}\right]$ is the pre-exponential factor, b_m is the exponent of the temperature dependence, $E_{a,m}$ $\left[\mathrm{J}\cdot\mathrm{kmol}^{-1}\right]$ is the activation energy, $R_u = 8314.472$ $\left[\mathrm{J}\cdot\mathrm{kmol}^{-1}\cdot\mathrm{K}^{-1}\right]$ is the universal gas constant, $K_{c,m}$ $\left[\left(\mathrm{m}^3\cdot\mathrm{kmol}^{-1}\right)^{\left(\sum_{n=1}^{n_s}\nu_{mn}^{'}-\sum_{n=1}^{n_s}\nu_{mn}^{''}\right)}\right]$ is the equilibrium constant, and $\Delta G_{T,m}^{o}$ $\left[\mathrm{J}\cdot\mathrm{kmol}^{-1}\right]$ is the change in Gibbs free energy at standard pressure and temperature T. A_m, b_m, and $E_{a,m}$ are collectively called the kinetic parameters of reaction m.

The initial mass fractions $Y_{j,0}$ are expressed compactly using the dimensionless equivalence ratio ϕ:

$$\phi = \frac{(Y_{O_2}/Y_{H_2})_{\mathrm{stoic}}}{(Y_{O_2}/Y_{H_2})} = \frac{(X_{O_2}/X_{H_2})_{\mathrm{stoic}}}{(X_{O_2}/X_{H_2})}, \tag{7}$$

where the subscript "stoic" refers to the stoichiometric ratios, and X_j is the molar fraction of the jth species, related to the mass fraction through

$$X_j = \frac{Y_j}{W_j \sum_{n=1}^{n_s} Y_n/W_n}. \tag{8}$$

For the rest of this paper, we use X_j in place of the Y_j as the species state variables. We assume a perfect gas mixture, thus closing the system with the following equation of state:

$$\rho = \frac{p}{R_u T \sum_{n=1}^{n_s} Y_n/W_n}, \tag{9}$$

where p [Pa] is the (assumed constant) pressure.

These equations are solved with the help of Cantera version 1.7.0,[6,7] an open source chemical kinetics software package. Since detailed kinetics render the ODE system stiff, time integration proceeds via CVODE,[8] which uses (implicit) backward differentiation formulas.

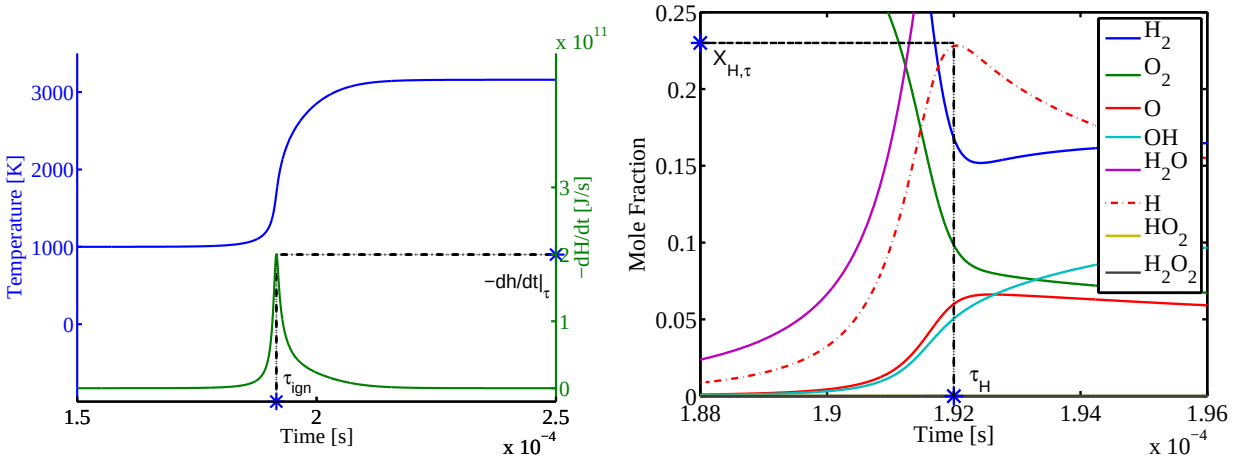
II.B. Experimental Goal & Selection of Observables

In this study, the experimental goal is to infer the values of (some or all of) the kinetic parameters (A_m, b_m, $E_{a,m}$) of a subset of the elementary reactions governing the system.

The most complete and detailed set of observables comprises simply the state variables as a function of time. However, discretization of the time domain may lead to resolution and storage problems; moreover, species concentrations and temperature far from the ignition event may not be particularly informative. Therefore, one should transform the state variables to some new observables that in some sense compress information, while retaining the information relevant to the experimental goal. Given the experimental goal of inferring the kinetic parameter values, the observables must reflect the variations in the kinetic

parameters, and thus equilibrium (steady-state) quantities are poor candidates. Furthermore, it is desirable for the observables to be common and easy to measure.

Taking the above factors into consideration, we select 10 distinct observables for this study. The first five are characteristic times: an igntion delay, defined as the time of peak heat release rate; and the times at which four radical species reach their peak concentrations. The next five observables are peak values themselves: peak heat release rate and peak radical concentrations. Examples of these observables are shown in Figure 1. As to be discussed later, log delay-times, $\ln \tau$, shall be used instead of delay times τ.



(a) τ_{ign} (b) τ_H and $X_{H,\tau}$

Figure 1. Selected experimental observables: ignition delay time τ_{ign}, H radical delay τ_H, and H radical peak $X_{H,\tau}$.

III. Bayesian Inference & Optimal Experimental Design

The experimental goal of inferring kinetic parameters has been defined in §II.B. This goal can be used to quantify the quality of an experimental design. First, however, we must introduce a method of solving the inference problem, as the measure of quality is typically related to the solution of the inference problem itself. Furthermore, solving the inference problem is necessary to validate the final results of optimal experimental design. Once the metric for the quality of an experiment is formulated, numerical optimization can be invoked to find the optimal experimental design. This experiment can then be carried out and its data analyzed as intended by the original experimental goal.

III.A. Bayesian Inference

III.A.1. Framework & numerical methods

In the Bayesian setting, a probability density is used to represent a state of knowledge about the values of parameters. We thus treat the model parameters and observables as random variables. Let the vector of random variables $\boldsymbol{\theta} \in \mathbb{R}^{n_{\boldsymbol{\theta}}}$ represent the uncertain parameters whose values are to be inferred; $\mathbf{y} = \{\mathbf{y}_l\}_{l=1}^{n_{\text{meas}}}$ be the set of n_{meas} data, where $\mathbf{y}_l \in \mathbb{R}^{n_{\mathbf{y}}}$ is one particular datum; and $\mathbf{d} \in \mathbb{R}^{n_{\mathbf{d}}}$ represent the experimental conditions (i.e., the design). Here, $n_{\boldsymbol{\theta}}$ is the number of uncertain parameters, $n_{\mathbf{y}}$ is the number of distinct observables (e.g., the 10 observables discussed in the previous section), and $n_{\mathbf{d}}$ is the number of design variables. At the heart of the Bayesian inference framework is, of course, Bayes' theorem. It can be written in this context as

$$p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d}) = \frac{p(\mathbf{y}|\,\boldsymbol{\theta},\mathbf{d})\,p(\boldsymbol{\theta}|\,\mathbf{d})}{p(\mathbf{y}|\,\mathbf{d})}, \tag{10}$$

where $p(\boldsymbol{\theta}|\,\mathbf{d})$ is the prior probability density function (PDF), $p(\mathbf{y}|\,\boldsymbol{\theta},\mathbf{d})$ is the likelihood function, $p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d})$ is the posterior PDF, and $p(\mathbf{y}|\,\mathbf{d})$ is the evidence. We can easily assume that the prior knowledge of $\boldsymbol{\theta}$ is independent of the experimental design, simplifying $p(\boldsymbol{\theta}|\,\mathbf{d}) = p(\boldsymbol{\theta})$.

The goal of inference is to characterize the posterior PDF, or to evaluate point or interval estimates based on the posterior. Samples from posterior can be obtained using Markov chain Monte Carlo (MCMC)

methods,[9–11] which require only pointwise evaluations of the *unnormalized* posterior. The resulting samples can then be used to either visually present the posterior or its marginals, or to construct sample estimates of posterior expectations. One simple and flexible type of of MCMC is the Metropolis-Hastings (MH) algorithm.[12,13] Two useful improvements to MH are the concepts of delayed rejection (DR)[14,15] and adaptive Metropolis (AM),[16] which are combined together in the DRAM algorithm.[17] These tools are used to solve inference problems in this study.

III.A.2. Prior and likelihood

For demonstration purposes, the unknown kinetic parameters are chosen to be A_1 and $E_{a,3}$ (i.e., the pre-exponential factor of reaction R1, $H + O_2 \rightleftharpoons O + OH$, and the activation energy of reaction R3, $H_2 + OH \rightleftharpoons H_2O + H$). The other kinetic parameters are set to their recommended values.[5] Instead of controlling A_1 directly, a transformation to $\tilde{A}_1 \equiv \ln(A_1/A_1^*)$ (where A_1^* is a nominal value of A_1) is used instead, which is useful in constraining A_1 to positive values only. Extension to additional parameters is straightforward.

The support of the parameter prior usually reflects constraints of physical admissibility, experimental limits, or regions of interest. Some preliminary tests have been performed to determine "interesting regions" of these parameters using the error models constructed in the next section. A uniform prior is assigned across the support adhering to the principles of indifference and maximum entropy.[18] The prior support is summarized in Table 1.

Parameter	Lower Bound	Upper Bound
$\tilde{A}_1$	-0.05	0.05
$E_{a,3}$	0	2.7196×10^7

Table 1. Prior support of the uncertain kinetic parameters $\tilde{A}_1$ and $E_{a,3}$. Uniform prior is assigned.

In constructing the likelihood, an additive error model is assumed for the observables

$$\mathbf{y}_l(\boldsymbol{\theta}, \mathbf{d}) = \mathbf{g}(\boldsymbol{\theta}, \mathbf{d}) + \boldsymbol{\epsilon}_l(\mathbf{d}),\ l = 1, \cdots, n_{\text{meas}}, \tag{11}$$

where $\mathbf{g}(\boldsymbol{\theta}, \mathbf{d})$ is the output from the computational model (defined in §II), and $\boldsymbol{\epsilon}_l(\mathbf{d}) = \left(\epsilon_{l,1}(\mathbf{d}), \ldots, \epsilon_{l,n_{\mathbf{y}}}(\mathbf{d})\right)$ is the additive error. Furthermore, the error shall be assumed to be i.i.d. zero-mean Gaussians $\epsilon_{l,i} \sim \mathcal{N}\left(0, \sigma_i^2(\mathbf{d})\right)$. This independence property conveniently makes the entries of a datum vector conditionally independent given the parameters $\boldsymbol{\theta}$. Additionally, different realizations of the data (i.e., realizations of $\mathbf{y}_l$) can be reasonably assumed to be independent conditioned on $\boldsymbol{\theta}$, and hence the likelihood factors as follows:

$$p(\mathbf{y}|\boldsymbol{\theta}, \mathbf{d}) = \prod_{l=1}^{n_{\text{meas}}} p(\mathbf{y}_l|\boldsymbol{\theta}, \mathbf{d}) = \prod_{l=1}^{n_{\text{meas}}} \prod_{i=1}^{n_{\mathbf{y}}} p(y_{l,i}|\boldsymbol{\theta}, \mathbf{d}) = \prod_{l=1}^{n_{\text{meas}}} \prod_{i=1}^{n_{\mathbf{y}}} \frac{1}{\sigma_i\sqrt{2\pi}} \exp\left\{-\frac{(y_{l,i} - g_i(\boldsymbol{\theta}, \mathbf{d}))^2}{2\sigma_i^2}\right\}. \tag{12}$$

The noise magnitudes σ_i are allowed to be a function of the design variables (but not of the kinetic parameters), because the observables in the combustion problem can vary over orders of magnitude as the design variables are adjusted within reasonable ranges. For instance, ignition delays can change by several orders of magnitude as the mixture's initial temperature is adjusted over a 100 K range. Because of the large variation of the characteristic time observables, we use their natural-log values, $\ln\tau$, in subsequent numerical implementations. The observational errors $\epsilon_{l,i}$, however, are still additive Gaussian perturbations to the non-log values. The key impact of using the log-delays is that the polynomial chaos expansions (described in §IV) directly approximate $\ln\tau$ values instead of τ.

Values of σ_i are determined as follows. At some desired design conditions $\mathbf{d}$, a simulation at the recommended parameter values is performed to obtain the nominal observable values. For the peak-value observables, σ_i's are simply chosen to be 10% of the nominal observable values. For the characteristic-time observables, a common value of σ shall be used (because all five of these characteristic times typically have very similar magnitudes), which is represented by an affine model:

$$\sigma = a + b\tau_{0.75}. \tag{13}$$

$\tau_{0.75}$ is the characteristic interval length of the 75% peak values in the profile of the heat release rate, illustrated in Figure 2. The purpose of a is to establish some minimum level of error, reflecting the resolution

limit of timing technology. The purpose of b is to represent an assumed linear dependence on some width measure of a peak. The rationale is that pinpointing the maximum of a flatter peak would be experimentally more challenging, since the measurement error on the magnitude of the quantity would be more significant compared to the variation of the quantity across the peak. For this study, $a = 10^{-5}$ seconds is a realistic choice, while $b = 10$ appears to work well after some experimenting.

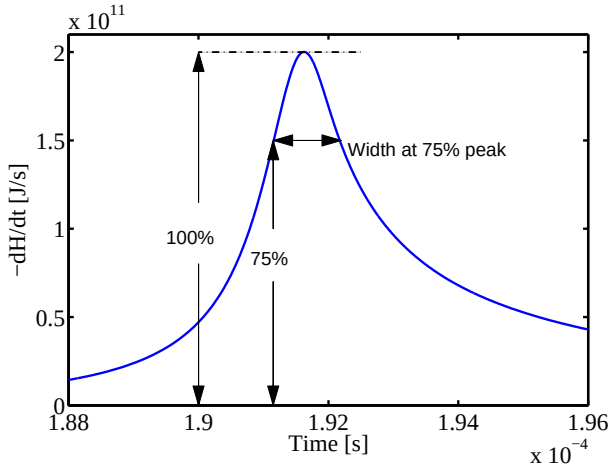


Figure 2. Illustration of $\tau_{0.75}$, the characteristic interval length of the 75% peak values in the profile of the heat release rate.

III.B. Optimal Bayesian Experimental Design

III.B.1. Background

Optimal design theory for linear models is well established,[19,20] since the analysis can be carried out in closed form. For nonlinear models, a closed-form analysis is typically not available. Various approximation methods have been proposed, such as locally linearizing the model, or approximating the posterior with different Gaussian distributions in which closed-form analysis may still be derived. However, a more general framework[21] that is largely independent of the linear design theory is preferred. This involves developing a utility function to be optimized, whose value can often only be approximated through numerical techniques.

Applications of optimal nonlinear experimental design are becoming more common—for example, in the fields of astrophysics,[22] material science,[23] and geophysics.[24] However, the models used in these applications are all relatively simple compared to the stiff ODE system described in §II. To date, no known literature has applied optimal nonlinear experimental design in such computationally intensive models; as will be shown below, additional developments are needed to make this framework practical.

III.B.2. Expected utility

Our approach to quantifying the value of an experiment employs concepts from information theory. In particular, the value indicator will be based on the Kullback-Leibler (KL) divergence (or relative entropy), whose usage in experimental design was first demonstrated by Lindley.[25] The KL divergence from PDF $p_B(\theta)$ to $p_A(\theta)$ is defined by

$$D\left(p_A(\theta)\,||p_B(\theta)\right) = \int_\Theta p_A(\theta) \ln\left[\frac{p_A(\theta)}{p_B(\theta)}\right] d\theta. \tag{14}$$

This quantity reflects the difference in information carried by the two PDFs, and is thus also called the information gain about θ if $p_A(\theta)$ is used instead of $p_B(\theta)$.

The KL divergence from the posterior to the prior is adopted as the utility u to reflect the value of an experiment:

$$u(\mathbf{d},\mathbf{y}) \equiv D\left(p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d})\,||p(\boldsymbol{\theta})\right) = \int_{\boldsymbol{\Theta}} p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d}) \ln\left[\frac{p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d})}{p(\boldsymbol{\theta})}\right] d\boldsymbol{\theta}. \tag{15}$$

The intuition is that a large KL divergence implies that the data $\mathbf{y}$ have decreased entropy (or increased information) in our state of knowledge of $\boldsymbol{\theta}$ by a large amount; and hence these data are more valuable with respect to the experimental goal of inference. When applied to a linear design problem, the KL divergence reduces to a popular optimality condition (linear d-optimality).

Equation 15 cannot be used directly, because it is a function of the data $\mathbf{y}$, which of course are unknown when designing the experiment. Thus, an expectation is taken over $\mathbf{y}|\mathbf{d}$ to obtain the expected utility function (i.e., a weighted average over the prior predictive distribution of the data):

$$U(\mathbf{d}) = \mathbb{E}_{\mathbf{y}|\mathbf{d}}\left[u(\mathbf{d},\mathbf{y})\right] = \int_{\mathcal{Y}} \int_{\boldsymbol{\Theta}} p(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}) \ln\left[\frac{p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d})}{p(\boldsymbol{\theta})}\right] d\boldsymbol{\theta}\, p(\,\mathbf{y}|\,\mathbf{d})\; d\mathbf{y}, \tag{16}$$

where $\mathcal{Y}$ is the support of $p(\,\mathbf{y}|\,\mathbf{d})$. Finally, the expected utility needs to be maximized to find the optimal experimental design

$$\mathbf{d}^* = \arg\, \max_{\mathbf{d}\in\mathcal{D}} U(\mathbf{d}), \tag{17}$$

where $\mathcal{D}$ is the design space. Simple modifications can extend this formulation to the simultaneous design of multiple experiments, or to sequential experimental design.

Note that the formulation above focuses on information gain in the model parameters θ. Alternatively, an optimal design may be defined in terms of some model output(s) of interest; experiments can thus be targeted to maximize information about the results of uncertainty propagation—i.e., a posterior predictive distribution—rather than the inputs. This definition is not employed in the present work, but will be explored in the future.

III.B.3. Numerical methods

The expected utility almost always needs to be approximated numerically, due to the nonlinearity of the simulation model. To do that, first, Equation 16 is rewritten as

$$\begin{aligned} U(\mathbf{d}) &= \int_{\mathcal{Y}} \int_{\boldsymbol{\Theta}} p(\boldsymbol{\theta}|\mathbf{y},\mathbf{d}) \ln\left[\frac{p(\boldsymbol{\theta}|\,\mathbf{y},\mathbf{d})}{p(\boldsymbol{\theta})}\right] d\boldsymbol{\theta}\, p(\,\mathbf{y}|\,\mathbf{d})\; d\mathbf{y}, \\ &= \int_{\mathcal{Y}} \int_{\boldsymbol{\Theta}} \ln\left[\frac{p(\,\mathbf{y}|\,\boldsymbol{\theta},\mathbf{d})}{p(\,\mathbf{y}|\,\mathbf{d})}\right] p(\,\mathbf{y}|\,\boldsymbol{\theta},\mathbf{d})\, p(\boldsymbol{\theta})\; d\boldsymbol{\theta}\, d\mathbf{y} \\ &= \int_{\mathcal{Y}} \int_{\boldsymbol{\Theta}} \left\{\ln\left[p(\,\mathbf{y}|\,\boldsymbol{\theta},\mathbf{d})\right] - \ln\left[p(\,\mathbf{y}|\,\mathbf{d})\right]\right\} p(\,\mathbf{y}|\,\boldsymbol{\theta},\mathbf{d})\, p(\boldsymbol{\theta})\; d\boldsymbol{\theta}\, d\mathbf{y}, \end{aligned} \tag{18}$$

where the second equality is due to the application of Bayes' theorem to the quantities both inside and outside the ln. Monte Carlo (MC) integration is then used to approximate the integral

$$U(\mathbf{d}) \approx \frac{1}{n_{\text{out}}} \sum_{i=1}^{n_{\text{out}}} \left\{\ln\left[\mathbf{p}\left(\mathbf{y}^{(i)}\middle|\,\boldsymbol{\theta}^{(i)},\mathbf{d}\right)\right] - \ln\left[p\left(\mathbf{y}^{(i)}\middle|\,\mathbf{d}\right)\right]\right\}, \tag{19}$$

where $\boldsymbol{\theta}^{(i)}$ are drawn from the prior $p(\boldsymbol{\theta})$; $\mathbf{y}^{(i)}$ are drawn from $p\left(\mathbf{y}|\,\boldsymbol{\theta}=\boldsymbol{\theta}^{(i)}\right)$ given the $\boldsymbol{\theta}^{(i)}$ just sampled; and n_{out} is the number of samples in this "outer" MC approximation. The evidence $p\left(\mathbf{y}^{(i)}\middle|\,\mathbf{d}\right)$ is typically unknown, and an additional "inner" MC approximation needs to be made to approximate it:

$$p\left(\mathbf{y}^{(i)}\middle|\,\mathbf{d}\right) = \int_{\boldsymbol{\Theta}} p\left(\mathbf{y}^{(i)}\middle|\,\boldsymbol{\theta},\mathbf{d}\right) p(\boldsymbol{\theta})\; d\boldsymbol{\theta} \approx \frac{1}{n_{\text{in}}} \sum_{j=1}^{n_{\text{in}}} p\left(\mathbf{y}^{(i)}\middle|\,\boldsymbol{\theta}^{(i,j)},\mathbf{d}\right), \tag{20}$$

where $\boldsymbol{\theta}^{(\cdot,j)}$ are drawn from the prior $p(\boldsymbol{\theta})$, and n_{in} is the number of samples in this "inner" MC approximation. As shown by Ryan,[23] this MC estimator is biased. The selection of n_{out} and n_{in} provides a tradeoff between estimator bias and variance.

This doubly-nested MC estimation can quickly become expensive. To alleviate the burden, we let $n_{\text{out}} = n_{\text{in}}$ and use the same batch of $\boldsymbol{\theta}^{(i)}$, $i = 1 \ldots n_{\text{out}}$ samples for every outer MC approximation, as well as for inner MC approximations at every $\mathbf{y}^{(i)}$. The resulting computational cost is then $\mathcal{O}(n_{\text{out}})$. Reusing samples adds bias, but tests have shown that this effect is very small.

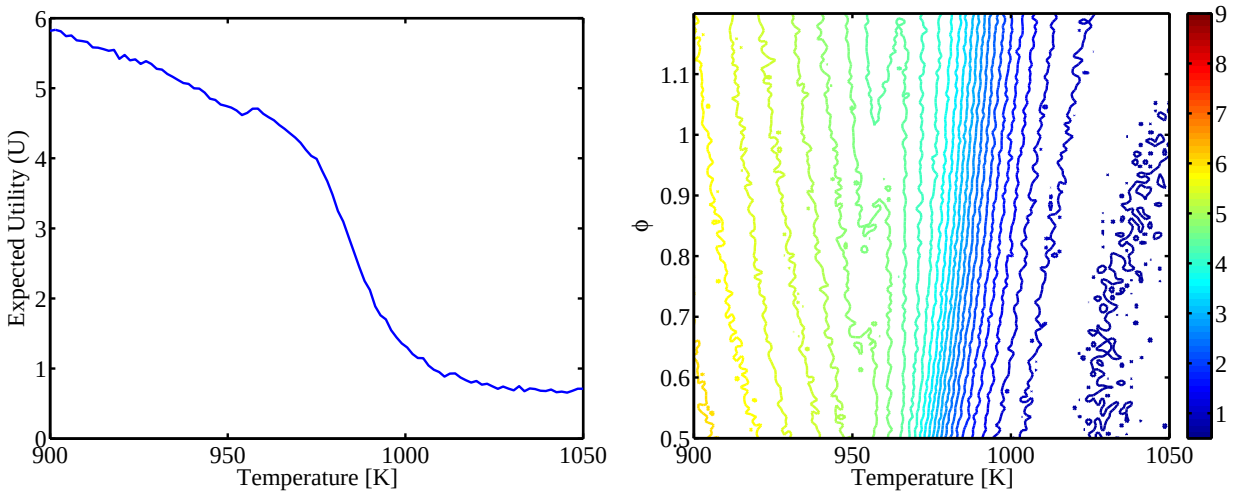
III.B.4. Stochastic optimization

One last numerical tool needed for the experimental design process is an optimization algorithm for Equation 17. More specifically, since only the MC approximation to the expected utility (which is the objective function) is available, the optimization method needs to be able to handle noisy objective functions. This topic is known as stochastic optimization (not to be confused with optimization methods that use random search techniques). Stochastic optimization is not yet implemented in the current code, pending further research and testing. Some potential algorithms include:

- using a large number of MC samples at each design point $\mathbf{d}$ to obtain a good estimate of the objective function, and then optimizing according to any deterministic optimization algorithm;
- classical stochastic approximation (SA) methods such as the Robbins-Monro[26] and Kiefer-Wolfowitz[27] algorithms; and
- simultaneous perturbation stochastic approximation (SPSA), developed by Spall and co-workers.[28,29]

III.C. Results and Issues

The nonlinear optimal experimental design framework is applied to the combustion problem with its underlying inference goal. First, only the initial temperature is allowed to vary in the design space $T_0 \in [900, 1050]$. The MC sampling size is chosen to be $n_{\text{out}} = n_{\text{in}} = 1001$. Second, the design space is extended to two dimensions, now including the equivalence ratio $\phi \in [0.5, 1.2]$ as well. The same number of MC samples is used. The expected utilities are plotted on a 1001-node discretization in each dimension, shown in Figure 3 for both cases.



(a) 1D design (T_0)

(b) 2D design (T_0 and ϕ)

Figure 3. Expected utilities of the combustion problem with 1 design variable T_0, and 2 design variables T_0 and ϕ.

The optimal design appears to be at a low initial temperature and low equivalence ratio. One possible explanation for choosing a low temperature design is that the ignition time scales are much longer at low temperatures (because reactions occur more slowly), but the assumed error model increases at a slow rate (recall the affine error model in Equation 13) such that the error variance becomes very small compared to the characteristic times. However, this may be only one of many factors that dictate the highly nonlinear relationship between the design condition and the expected utility, which is difficult to assess using only simplified arguments.

The computational resources required for these two tasks are enormous. The 1D design case requires hours to run when parallelized on approximately 50 3-GHz CPUs; the 2D design case requires days to complete under the same parallelization. Given that these examples use only a 19-reaction mechanism, substantial speedups are needed in order for the framework to be practical.

IV. Polynomial Chaos

As demonstrated in the previous section, for the proposed experimental design framework to have applicability to more complex real-world experiments, it must be accelerated substantially. One way to achieve this speedup is to capture the input-output relationships using polynomial chaos (PC), a particular spectral expansion for random variables.

IV.A. Background

Polynomial chaos expansions were first introduced by Wiener,[30] motivated by physical problems in statistical mechanics. In subsequent years, PC expansions and associated spectral methods,[31–33] have been used extensively for uncertainty propagation—characterizing the probability distribution of the output of a model given a known distribution on the input. These methods constitute attractive alternatives to Monte Carlo simulation in numerous applications, including computational fluid dynamics[34,35] and reacting flow.[36] PC expansions have also been used to accelerate the Bayesian solution of inverse problems.[37–39] Several comprehensive reviews are available.[40–42]

IV.B. Formulation

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, where Ω is the sample space, $\mathcal{F}$ is the σ-field, and $\mathbb{P}$ is the probability measure. Then any real-valued random variable $\theta : \Omega \to \mathbb{R}$, $\theta \in L_2(\Omega, \mathbb{P})$ (i.e., with finite variance) can be expanded as follows:

$$\theta(\omega) = \sum_{|\mathbf{i}|=0}^{\infty} \theta_{\mathbf{i}} \Psi_{\mathbf{i}}(\xi_1, \xi_2, \ldots), \tag{21}$$

where $\omega \in \Omega$ is an element of the sample space, $\xi_i(\omega)$ are i.i.d. random variables defined on the same probability space as θ, $\theta_{\mathbf{i}}$ are the expansion coefficients, $\mathbf{i} = (i_1, i_2, \ldots)$, $\forall i_j \in \mathbb{N}$, is an infinite-dimensional multi-index, $|\mathbf{i}| = i_1 + i_2 + \ldots$ is the ℓ_1 norm, and $\Psi_{\mathbf{i}}$ are multivariate orthogonal polynomials, written as products of univariate orthogonal polynomials

$$\Psi_{\mathbf{i}}(\xi_1, \xi_2, \ldots) = \prod_{j=1}^{\infty} \psi_{i_j}(\xi_j). \tag{22}$$

This expansion is known as the polynomial chaos expansion, and is convergent in the mean-square sense.[43] The basis functions ψ_{i_j} are polynomials of order i_j in the independent variable ξ_j, orthogonal with respect to the density of ξ_j (i.e., $p(\xi_j)$). More specifically,

$$\mathbb{E}_\xi [\psi_m \psi_n] = \int_\Xi \psi_m(\xi) \psi_n(\xi) p(\xi) \, d\xi = \delta_{m,n} \mathbb{E}_\xi \left[\psi_m^2\right], \tag{23}$$

where Ξ is the support of $p(\xi)$. For computational purposes, the infinite sum and infinite dimension must be truncated to some finite order p_θ and stochastic dimension n_s, leading to an expansion with n_{PC} terms:

$$n_{PC} = \begin{pmatrix} n_s + p_\theta \\ p_\theta \end{pmatrix} = \frac{(n_s + p_\theta)!}{n_s! p_\theta!}. \tag{24}$$

The choice of p_θ is often influenced by the expected smoothness of the random variable, and the choice of n_s reflects the number of independent directions needed to capture the stochasticity of the system.

Methods to compute the PC expansion coefficients are broadly divided into two groups—intrusive and non-intrusive. The former involves substituting the expansions into the governing equations, resulting in a larger, modified system which can be difficult to solve; but, it only needs to be solved once. The latter involves directly projecting the unknown random variable onto the basis functions, which is done by reusing the solver of the deterministic system; but, it needs to be solved many times. For the combustion problem, the non-intrusive spectral projection (NISP), which is a type of non-intrusive method, is chosen, mainly for the flexibility it offers in choosing the observables in the combustion problem context.

IV.C. Non-Intrusive Spectral Projection (NISP) for the Combustion Problem

The random variables with known prior PDFs are the input parameters, namely, $\boldsymbol{\theta} = \left(\tilde{A}_1, E_{a,3}\right)$. As discussed in §III.A.2, a uniform prior has been assigned, with the support described in Table 1. Therefore, the Legendre chaos is a convenient choice, with $\xi_1, \xi_2 \sim \mathcal{U}(-1, 1)$. The PC expansions for these parameters are simply

$$\theta_1 = 0.05\xi_1 \tag{25}$$
$$\theta_2 = 1.3598 \times 10^7 + 1.3598 \times 10^7 \xi_2. \tag{26}$$

The goal now is to construct PC expansions for the outputs, which are the model-predicted delay times and peak values $\mathbf{g}(\boldsymbol{\theta}, \mathbf{d})$. As discussed at the end of §II.B, the log of the characteristic time observables will be used in the PC expansions. Before this is done, we must include $\mathbf{d}$ among the input variables. To do this, we write a PC expansion for $\mathbf{d}$, effectively increasing the stochastic dimension and treating the design parameters as random variables. Denoting $\xi_3 = T_0$ and $\xi_4 = \phi$, the new stochastic dimension is $n_s = 4$. For the sake of approximation, uniform "priors" are assigned to $\mathbf{d}$ over the regions where the experiments are physically realizable. In particular, the support for the design space is shown in Table 2, where the bounds are arbitrarily but reasonably chosen. Using the Legendre chaos, expansions for the design variables are simply

$$d_1 = 975 + 75\xi_3 \tag{27}$$
$$d_2 = 0.85 + 0.35\xi_4. \tag{28}$$

Note that one would never sample ξ_3 and ξ_4, because there is no actual randomness in $\mathbf{d}$. These expansions are for the purposes of approximating the observables' dependence on $\mathbf{d}$. When a particular $\mathbf{d}$ is desired, the corresponding ξ_3 and ξ_4 are simply computed by inverting Equations 27 and 28.

Parameter	Lower Bound	Upper Bound
T_0	900	1050
ϕ	0.5	1.2

Table 2. "Prior" support of the design variables T_0 and ϕ. Uniform prior is assigned.

Now the PC expansions of the unknown variables can finally be constructed. Performing a standard Galerkin projection of the random variables onto the desired basis functions and taking advantage of the orthogonality property, the coefficients are simply

$$g_{m,\mathbf{i}} = \frac{\mathbb{E}\left[g_m \Psi_{\mathbf{i}}\right]}{\mathbb{E}\left[\Psi_{\mathbf{i}}^2\right]} = \frac{\int_{\boldsymbol{\Xi}} g_m\left(\boldsymbol{\theta}(\boldsymbol{\xi}), \mathbf{d}(\boldsymbol{\xi})\right) \Psi_{\mathbf{i}}(\boldsymbol{\xi}) p(\boldsymbol{\xi})\, d\boldsymbol{\xi}}{\int_{\boldsymbol{\Xi}} \Psi_{\mathbf{i}}^2(\boldsymbol{\xi})\, p(\boldsymbol{\xi})\, d\boldsymbol{\xi}}, \tag{29}$$

where $g_{m,\mathbf{i}}$ is the PC expansion with multi-index $\mathbf{i}$ for the mth observable. The denominators are easily evaluated with analytical formulas that depend on the particular polynomial basis. The numerator, on the other hand, almost always has no analytic form and must be computed via numerical integration techniques—this is the heart of the NISP computation. Due to the potentially high stochastic dimension as well as the expensive evaluations of the forward model $\mathbf{g}$, an efficient method of high-dimensional numerical integration is essential.

IV.D. Sparse Quadrature & Dimension-Adaptive Sparse Quadrature

Sparse quadrature (SQ) is an efficient numerical integration method based on the Smolyak rule.[44,45] Compared to Monte Carlo and even quasi Monte Carlo (QMC), SQ takes great advantage of integrand smoothness, and thus has faster convergence properties; however, it still has (a weak) dependence on dimensionality, and will be ultimately surpassed by MC as dimension is increased.

Dimension-adaptive sparse quadrature (DASQ)[46] is a natural extension to the isotropic Smolyak rule used to construct sparse grids. It constructs the quadrature rule intelligently by selecting "index sets" based on a heuristic error indicator; the selection exploits anisotropy, in that certain dimensions can be refined more than others. One drawback of DASQ is that parallelization can only be implemented within each index set, which is not as efficient as the parallelization of MC and SQ. The original DASQ algorithm[46] does

not address how nested quadrature points can be reused as adaptation proceeds. We have proposed a new algorithm[47] to solve this problem.

Figure 4 shows results from a 20-dimensional anisotropic integration problem;[48] the problem has an analytical solution, which allows comparison between the various integration methods. Details are given elsewhere.[47] Here, we simply note that the relative L_1 error for DASQ is better than all other methods, indicating that adaptation is indeed helpful. The error indicator for DASQ appears to approximate the true error, but it is not guaranteed to be always a lower or higher estimate. The non-monotonic convergence of the dimension-adaptive case implies that in some cases, refinement can actually lead to a worse estimate. This suggests that the error indicator can be further improved.

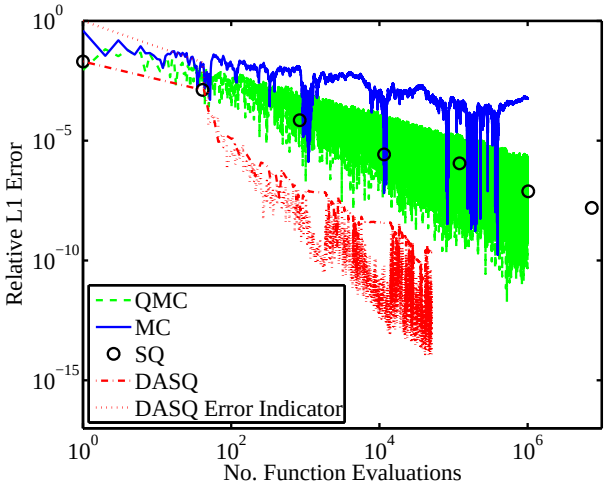


Figure 4. Relative L_1 error of a 20D integral using various numerical integration schemes.

Our explorations have shown that ignition models give rise to extremely anisotropic integrands, and thus the computational advantages of DASQ over SQ in this context can be enormous.[47]

V. Results

This section combines the optimal design and polynomial chaos tools introduced above and applies them to the combustion kinetics problem, thus providing the main results of this paper. First, we construct PC expansions for the observables of the combustion system, using DASQ to compute the integrals in Equation 29. Second, the PC expansions are used in the optimal nonlinear experimental design framework, yielding optimal experimental conditions as in Equation 17. Third, the inference problem is solved at a selected number of experimental conditions from the design space in order to validate the experimental design results. Finally, we include a short discussion on computational efficiency.

V.A. PC Expansion Construction

PC expansions for the log-delay and peak-value observables are constructed using DASQ, as detailed in §IV.D. A global order of $p_\theta = 12$ is chosen for all of the expansions, and quadrature in the 4-dimensional space of $\mathbf{y}, \boldsymbol{\theta}$ is stopped once a total of $n_{\text{quad}} = 25,000$ function evaluations has been exceeded. These choices are not intuitive to reach, and are instead a result of some trial and error.

A more rigorous analysis is performed by computing the relative L_2 error for different combinations of p_θ and n_{quad}:

$$e_m = \frac{\int_{\boldsymbol{\Xi}} \left| g_m\left(\boldsymbol{\theta}\left(\boldsymbol{\xi}\right), \mathbf{d}\left(\boldsymbol{\xi}\right)\right) - g_m^p\left(\boldsymbol{\xi}\right)\right|^2 p\left(\boldsymbol{\xi}\right)\, d\boldsymbol{\xi}}{\int_{\boldsymbol{\Xi}} \left| g_m\left(\boldsymbol{\theta}\left(\boldsymbol{\xi}\right), \mathbf{d}\left(\boldsymbol{\xi}\right)\right)\right|^2 p\left(\boldsymbol{\xi}\right)\, d\boldsymbol{\xi}}, \; m = 1, \ldots, n_{\mathbf{y}}, \tag{30}$$

where g_m and g_m^p are the output from the original ODE model and the PC expansion, respectively, for the mth observable; and $\boldsymbol{\theta}$ and $\mathbf{d}$ as functions of $\boldsymbol{\xi}$ are from Equations 25, 26 and 27, 28, respectively. The integrals in the L_2 error expression are computed using a level 15 SQ rule (3,502,081 abscissas); such a

high accuracy integration rule is used in an attempt to prevent additional integration errors from polluting the L_2 errors. Figure 5 shows the $\log_{10}$ of the L_2 error contours for τ_{ign} with respect to p_θ and n_{quad}. As p_θ is increased, n_θ also needs to be increased for further improvement in accuracy; a delicate balance between truncation and aliasing (integration) error needs to be reached to achieve the highest efficiency of the quadrature rule. Note that n_{quad} values are approximate, as DASQ is immediately terminated at the iteration after passing n_{quad}, which might not be exactly equal to n_{quad}.

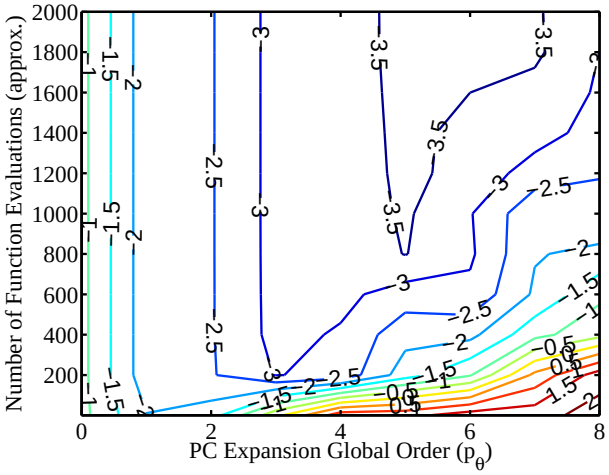
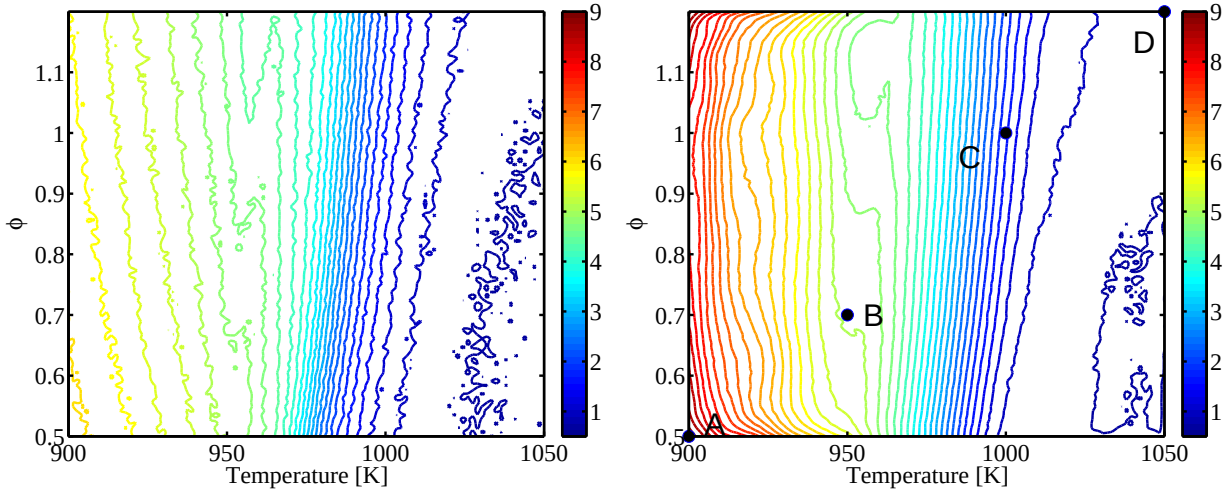


Figure 5. $\log_{10}$ **of the** L_2 **errors of** $\ln \tau_{\text{ign}}$**.**

V.B. Optimal Experimental Design

The optimal experimental design problem stated in Equation 17 is solved with the PC expansions replacing the original ODE model, but now with a higher $n_{\text{in}} = n_{\text{out}} = 10,001$. The expected utility contours for both cases are presented in Figure 6 using a 1001×1001 uniform grid, with the first subfigure reproduced from Figure 3(b) for comparison. The contours from the PC expansions are overall quite similar to those from the original model, and most importantly, they yield the same optimal experimental design at $(T_0^*, \phi^*) = (900, 0.5)$. They are even able to capture the ridge near $T_0 = 950$, although not perfectly.



(a) Using original ODE model

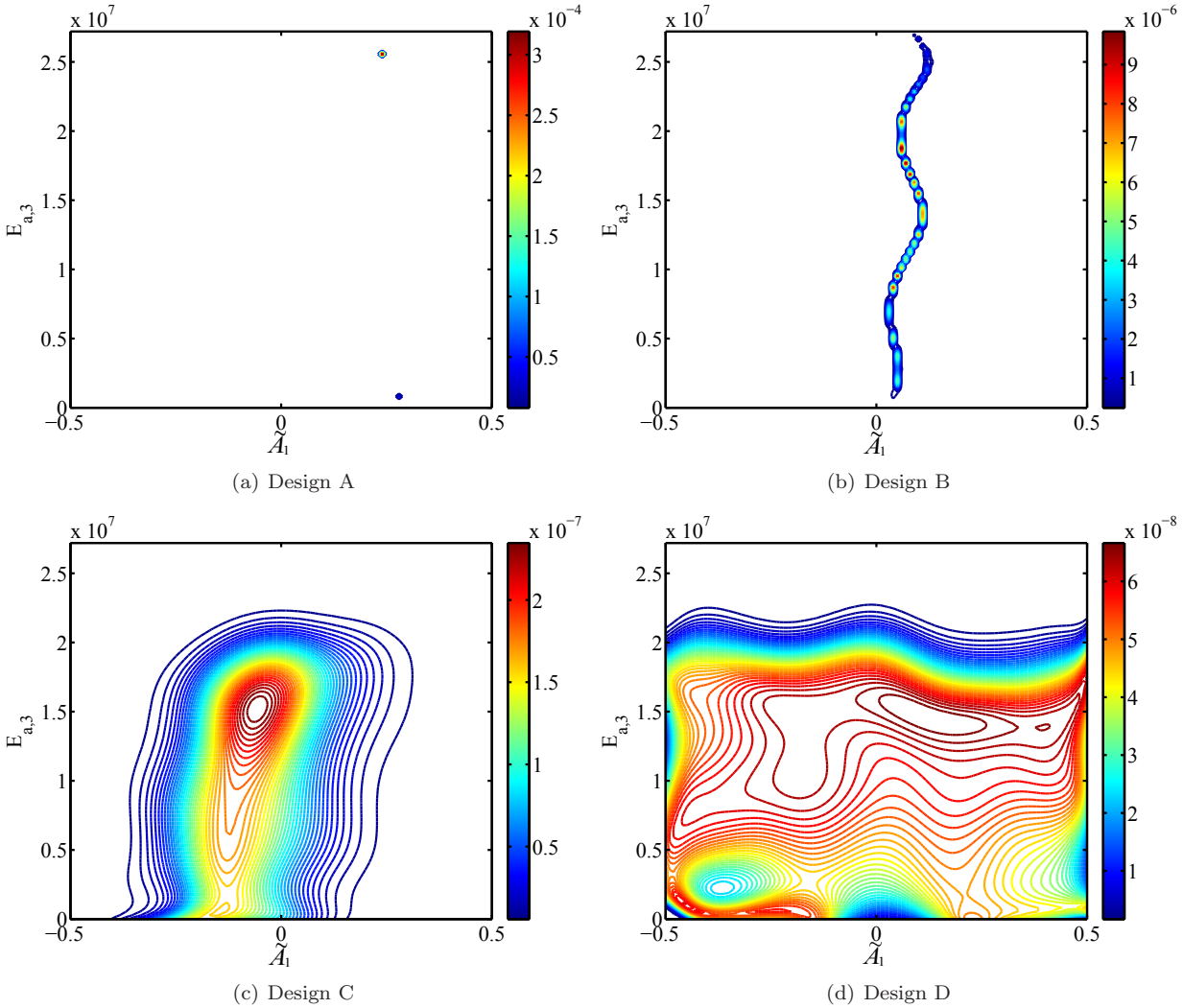
(b) Using PC expansions

Figure 6. Expected utility contours of the combustion problem with 2 design variables T_0 **and** ϕ**, using the original ODE model (reproduced from Figure 3(b)), and using the PC expansions with** $p_\theta = 12$ **and** $n_{\text{quad}} = 25,000$**.**

V.C. Validation of the Experimental Design Results

The validity of the expected utility calculations can be assessed by solving the inference problem at a number of design points. In particular, designs at points A, B, C, and D are selected; these experimental conditions are illustrated on the design space in Figure 6(b). Since the expected utility is decreasing from designs A through D, then intuitively, the posterior from the inference problem should reflect the most information at A, less at B and C, and least at D.

At each design, one single datum is generated from $g_m(\boldsymbol{\theta}, \mathbf{d})$ (not $g_m^p(\boldsymbol{\theta}, \mathbf{d})$) using the recommended values of the kinetic parameters, then perturbed with synthetic measurement noise according to the likelihood described in §III.A.2. The posteriors computed using the PC expansions are shown in Figure 7, and for comparisons, the posteriors computed using the original ODE model (and with the same data) are shown in Figure 8. With the exception of design A, these two sets of posteriors agree reasonably well, implying that the PC expansions are not only suitable for experimental design, they can also be used for solving the inference problems as well. The posterior for design A is especially sensitive to the forward model approximation error; at this point, the design conditions are so informative that the assumed measurement error is small compared to the PC approximation error.



(a) Design A

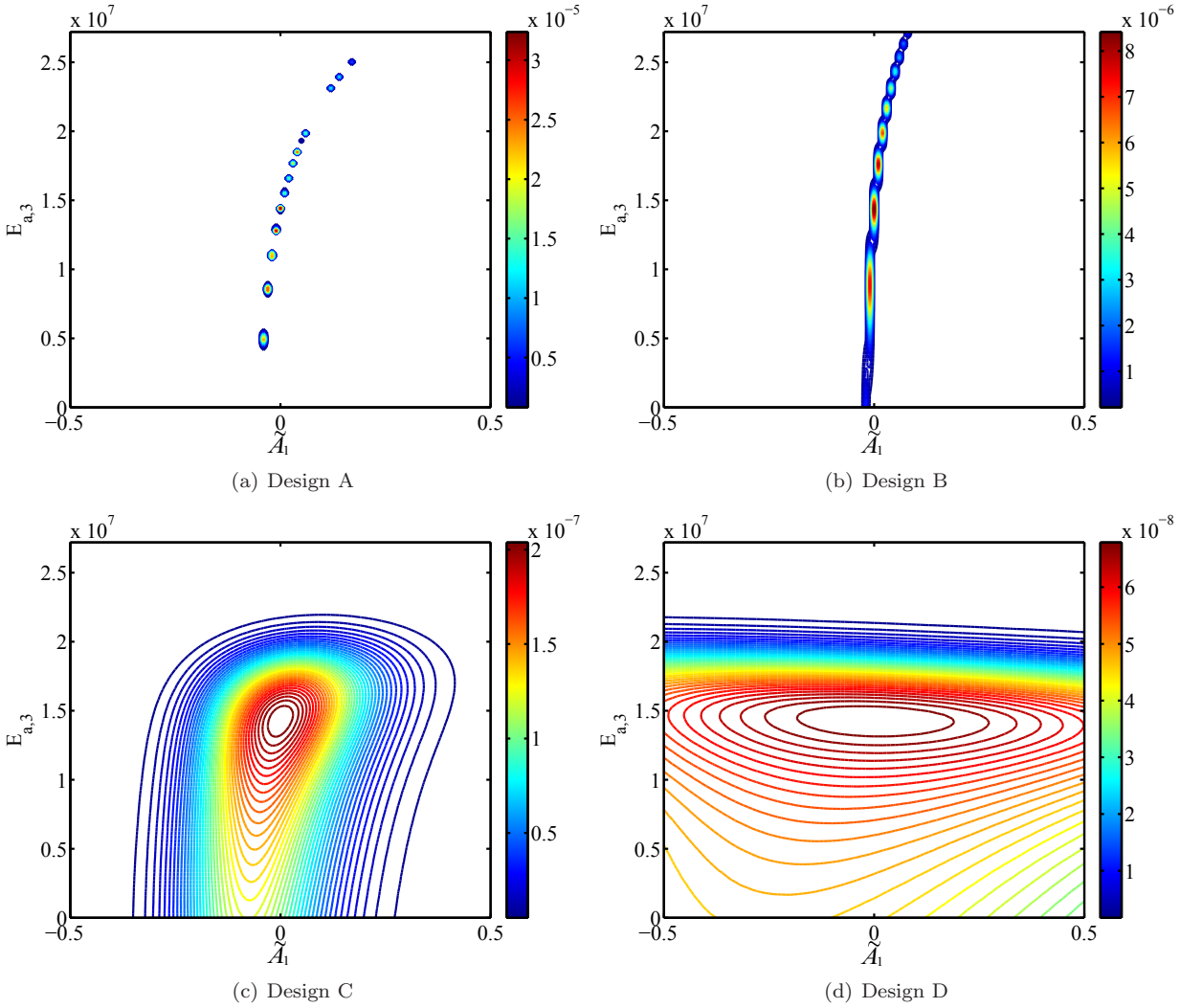
(b) Design B

(c) Design C

(d) Design D

Figure 7. Posterior of the inference problem at the four chosen designs to validate the experiment design methodology, constructed using PC expansions with $p_\theta = 12$, $n_{\text{quad}} = 25,000$.

As expected, the kinetic parameters can be most confidently inferred from design A, less confidently in B and C, and least confidently in D. This implies that the expected utility is indeed a good indicator of the quality or usefulness of an experiment, and that the experimental design methodology is functioning.



(a) Design A

(b) Design B

(c) Design C

(d) Design D

Figure 8. Posterior of the inference problem at the four chosen designs to validate the experiment design methodology, constructed with direct evaluations of the original ODE model.

One reason why the posterior at design A is so close to being "deterministic" is that the error width from the likelihood model is too small compared to the error due to the approximation from the PC expansions. In the future, error models should not only try to capture measurement noise, but also inadequacy in the model itself. Design A's posterior is also clearly not centered around the recommended values, but this can be explained by the fact that the datum is perturbed by random noise, and that only a finite number (one in this case) is available.

V.D. Computational Savings

The original Cantera-based ODE model has a cost of about 0.2 seconds per evaluation on a 3-GHz CPU, whereas evaluations of the $p_\theta = 12$, n_{quad} =25,000 PC expansions are about 600 times faster. However, the PC expansions require an up-front investment for construction. This cost can be reduced by using a lower order polynomial and fewer quadrature points, but the tradeoff is accuracy. PC expansions are thus worthwhile to invest in given a sufficiently large number of required samples. For example, the break-even point for the $p_\theta = 12$, n_{quad} =25,000 PC expansions is 25,038 forward evaluations. In this study, about 100 million samples are used to perform the final experimental design computations, which translates to a saving of almost 8 months if a single CPU were used.

VI. Conclusions

The development and calibration of chemical kinetic models often relies on parameter inference, using noisy observations of a limited number of indirect diagnostics. The relationship between these observables and the kinetic parameters of interest is highly nonlinear and multivariate; each observable depends on multiple kinetic parameters and experimental conditions, and the nature of these dependencies is difficult to elucidate *a priori*. Optimal experimental design implicitly evaluates these dependencies to determine the most informative experimental conditions for inferring a set of target parameters. In the Bayesian framework, this determination also incorporates prior uncertainty in the kinetic parameters and a detailed statistical model for observational errors.

This paper introduces a general Bayesian formulation for optimal experimental design. The formulation is made computationally tractable through the use of polynomial chaos surrogates, which are constructed non-intrusively via dimension-adaptive sparse quadrature. We demonstrate the methodology on hydrogen-oxygen ignition experiments, using ignition delays to infer selected rate parameters. The methodology developed here is quite general, however, and can be extended to numerous other combustion observables.

In the future, we will implement stochastic search methods to efficiently identify the optimal design conditions given noisy (small-sample) estimates of the objective function. These methods are imperative for high-dimensional design spaces. Furthermore, we will extend our formulation to capture uncertainties in the design conditions themselves, as real-world experiments rarely achieve the desired conditions exactly. Finally, we will explore extensions of this methodology to the relatively new area of sequential experimental design.

Acknowledgments

The authors would like to acknowledge support from the KAUST Global Research Partnership and from the US Department of Energy, Office of Science, Office of Advanced Scientific Computing Research (ASCR).

References

[1]Davidson, D. F. and Hanson, R. K., "Interpreting Shock Tube Ignition Data," *International Journal of Chemical Kinetics*, Vol. 36, No. 9, 2004, pp. 510–523.

[2]Davidson, D. F. and Hanson, R. K., "Recent Advances in Shock Tube/Laser Diagnostic Methods for Improved Chemical Kinetics Measurements," *Shock Waves*, Vol. 19, No. 4, 2009, pp. 271–283.

[3]Vasu, S., Davidson, D., and Hanson, R., "Jet Fuel Ignition Delay Times: Shock Tube Experiments over Wide Conditions and Surrogate Model Predictions," *Combustion and Flame*, Vol. 152, No. 1–2, 2008, pp. 125–143.

[4]Kahandawala, M., Corera, S., Williams, S., Carter, C., and Sidhu, S., "Investigation of Kinetics of Iso-Octane Ignition under Scramjet Conditions," *International Journal of Chemical Kinetics*, Vol. 38, 2006, pp. 194–201.

[5]Yetter, R. A., Dryer, F. L., and Rabitz, H. A., "A Comprehensive Reaction Mechanism for Carbon Monoxide/Hydrogen/Oxygen Kinetics," *Combustion Science and Technology*, Vol. 79, 1991, pp. 97–128.

[6]Goodwin, D. G., *Cantera C++ User's Guide*, California Institute of Technology, October 2002.

[7]"Cantera 1.7.0 Website," http://sourceforge.net/projects/cantera/.

[8]Cohen, S. D. and Hindmarsh, A. C., "CVODE, a Stiff/Nonstiff ODE Solver in C," *Computers in Physics*, Vol. 10, No. 2, 1996, pp. 138–143.

[9]Andrieu, C., de Freitas, N., Doucet, A., and Jordan, M. I., "An Introduction to MCMC for Machine Learning," *Machine Learning*, Vol. 50, No. 1, January 2003, pp. 5–43.

[10]Robert, C. P. and Casella, G., *Monte Carlo Statistical Methods*, Springer Verlag, 2004.

[11]Gilks, W. R., Richardson, S., and Spiegelhalter, D. J., *Markov Chain Monte Carlo in Practice*, Interdisciplinary Statistics, Chapman & Hall/CRC, 1996.

[12]Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E., "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, Vol. 21, No. 6, 1953, pp. 1087–1092.

[13]Hastings, W. K., "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, Vol. 57, No. 1, 1970, pp. 97–109.

[14]Green, P. J. and Mira, A., "Delayed Rejection in Reversible Jump Metropolis-Hastings," *Biometrika*, Vol. 88, No. 4, 2001, pp. 1035–1053.

[15]Mira, A., "On Metropolis-Hastings Algorithms with Delayed Rejection," *Metron – International Journal of Statistics*, Vol. 59, No. 3–4, 2001, pp. 231–241.

[16]Haario, H., Saksman, E., and Tamminen, J., "An Adaptive Metropolis Algorithm," *Bernoulli*, Vol. 7, No. 2, 2001, pp. 223–242.

[17]Haario, H., Laine, M., Mira, A., and Saksman, E., "DRAM: Efficient Adaptive MCMC," *Statistics and Computing*, Vol. 16, No. 4, December 2006, pp. 339–354.

[18]Jaynes, E. T. and Bretthorst, G. L., *Probability Theory: the Logic of Science*, Cambridge University Press, 2003.

[19]Atkinson, A. C. and Donev, A. N., *Optimum Experimental Designs*, Oxford Statistical Science Series, Oxford University Press, 1992.

[20]Chaloner, K. and Verdinelli, I., "Bayesian Experimental Design: a Review," *Statistical Science*, Vol. 10, No. 3, 1995, pp. 273–304.

[21]Müller, P., "Simulation Based Optimal Design," *Bayesian Statistics 6: Proceedings of the Sixth Valencia International Meeting*, Oxford University Press, 1998, pp. 459–474.

[22]Loredo, T. J. and Chernoff, D. F., "Bayesian Adaptive Exploration," *Statistical Challenges of Astronomy*, Springer, 2003, pp. 57–69.

[23]Ryan, K. J., "Estimating Expected Information Gains for Experimental Designs With Application to the Random Fatigue-Limit Model," *Journal of Computational and Graphical Statistics*, Vol. 12, No. 3, September 2003, pp. 585–603.

[24]Van den Berg, J., Curtis, A., and Trampert, J., "Optimal Nonlinear Bayesian Experimental Design: an Application to Amplitude Versus Offset Experiments," *Geophysical Journal International*, Vol. 155, No. 2, November 2003, pp. 411–421.

[25]Lindley, D. V., "On a Measure of the Information Provided by an Experiment," *The Annals of Mathematical Statistics*, Vol. 27, No. 4, 1956, pp. 986–1005.

[26]Robbins, H. and Monro, S., "A Stochastic Approximation Method," *The Annals of Mathematical Statistics*, Vol. 22, No. 3, 1951, pp. 400–407.

[27]Kiefer, J. and Wolfowitz, J., "Stochastic Estimation of the Maximum of a Regression Function," *The Annals of Mathematical Statistics*, Vol. 23, No. 3, 1952, pp. 462–466.

[28]Spall, J. C., "An Overview of the Simultaneous Perturbation Method for Efficient Optimization," *Johns Hopkins APL Technical Digest*, Vol. 19, No. 4, 1998, pp. 482–492.

[29]Spall, J. C., "Implementation of the Simultaneous Perturbation Algorithm for Stochastic Optimization," *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 34, No. 3, 1998, pp. 817–823.

[30]Wiener, N., "The Homogeneous Chaos," *American Journal of Mathematics*, Vol. 60, No. 4, 1938, pp. 897–936.

[31]Xiu, D. and Karniadakis, G. E., "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM Journal of Scientific Computing*, Vol. 24, No. 2, 2002, pp. 619–644.

[32]Debusschere, B. J., Najm, H. N., Pébay, P. P., Knio, O. M., Ghanem, R. G., and Le Maître, O. P., "Numerical Challenges in the Use of Polynomial Chaos Representations for Stochastic Processes," *SIAM Journal on Scientific Computing*, Vol. 26, No. 2, 2004, pp. 698–719.

[33]Le Maître, O. P., Knio, O. M., Najm, H. N., and Ghanem, R. G., "Uncertainty Propagation Using Wiener-Haar Expansions," *Journal of Computational Physics*, Vol. 197, No. 1, 2004, pp. 28–57.

[34]Walters, R. W., "Towards Stochastic Fluid Mechanics via Polynomial Chaos," *41st Aerospace Sciences Meeting and Exhibit*, 2003, AIAA paper 2003-413.

[35]Hosder, S., Walters, R., and Perez, R., "A Non-Intrusive Polynomial Chaos Method For Uncertainty Propagation in CFD Simulations," *44th AIAA Aerospace Sciences Meeting and Exhibit*, 2006, AIAA paper 2006-891.

[36]Reagan, M. T., Najm, H. N., Ghanem, R. G., and Knio, O. M., "Uncertainty Quantification in Reacting-Flow Simulations Through Non-Intrusive Spectral Projection," *Combustion and Flame*, Vol. 132, No. 3, 2003, pp. 545–555.

[37]Marzouk, Y. M., Najm, H. N., and Rahn, L. A., "Stochastic Spectral Methods for Efficient Bayesian Solution of Inverse Problems," *Journal of Computational Physics*, Vol. 224, No. 2, June 2007, pp. 560–586.

[38]Marzouk, Y. M. and Xiu, D., "A Stochastic Collocation Approach to Bayesian Inference in Inverse Problems," *Communications in Computational Physics*, Vol. 6, No. 4, October 2009, pp. 826–847.

[39]Marzouk, Y. M. and Najm, H. N., "Dimensionality Reduction and Polynomial Chaos Acceleration of Bayesian Inference in Inverse Problems," *Journal of Computational Physics*, Vol. 228, No. 6, April 2009, pp. 1862–1902.

[40]Najm, H. N., "Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics," *Annual Review of Fluid Mechanics*, Vol. 41, No. 1, 2009, pp. 35–52.

[41]Xiu, D., "Fast Numerical Methods for Stochastic Computations: a Review," *Communications in Computational Physics*, Vol. 5, No. 2–4, 2009, pp. 242–272.

[42]Le Maître, O. P. and Knio, O. M., *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics*, Springer, 2010.

[43]Cameron, R. H. and Martin, W. T., "The Orthogonal Development of Non-Linear Functionals in Series of Fourier-Hermite Functionals," *The Annals of Mathematics*, Vol. 48, No. 2, 1947, pp. 385–392.

[44]Barthelmann, V., Novak, E., and Ritter, K., "High Dimensional Polynomial Interpolation on Sparse Grids," *Advances in Computational Mathematics*, Vol. 12, No. 4, March 2000, pp. 273–288.

[45]Gerstner, T. and Griebel, M., "Numerical Integration Using Sparse Grids," *Numerical Algorithms*, Vol. 18, 1998, pp. 209–232.

[46]Gerstner, T. and Griebel, M., "Dimension-Adaptive Tensor-Product Quadrature," *Computing*, Vol. 71, No. 1, 2003, pp. 65–87.

[47]Huan, X., *Accelerated Bayesian Experimental Design for Chemical Kinetic Models*, Master's thesis, Massachusetts Institute of Technology, 2010.

[48]Morokoff, W. J. and Caflisch, R. E., "Quasi-Monte Carlo Integration," *Journal of Computational Physics*, Vol. 122, 1995, pp. 218–230.