SAND2017-3413C

Quantifying Uncertainty from Model Error in Turbulent Combustion Applications

Xun Huan, Khachik Sargsyan, Zachary Vane,
Guilhem Lacaze, Joe Oefelein, Habib N. Najm

Sandia National Laboratories

April 3, 2017

Acknowledgement

- Defense Advanced Research Projects Agency (DARPA)
  Enabling Quantification of Uncertainty in Physical Systems (EQUiPS)

- Sandia National Laboratories[1]

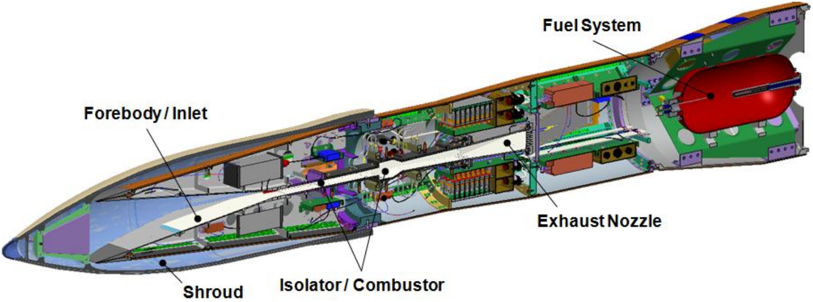
[1] Sandia National Laboratories is a multi-program laboratory managed and operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

HIFiRE-II Scramjet

Development of scramjet[2] engine involves

- flow simulations
- uncertainty quantification (UQ)
- design optimization

We focus on the HIFiRE-II[3] configuration:



[2] supersonic combustion ramjet

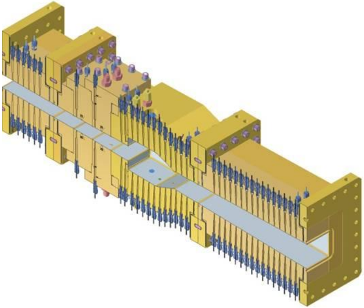
[3] Hypersonic International Flight Research and Experimentation-II

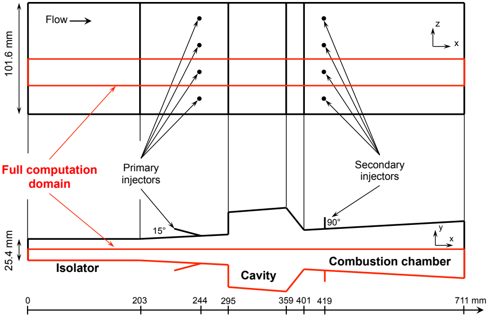
Computational domain of initial unit problem

Final: simulation of full combustor domain, match experimental setup —HIFiRE Direct Connect Rig (HDCR) [bottom-left figure]

Initial unit problem: primary injection section

- omit cavity
- no combustion
- focus on interaction of fuel jet and supersonic air crossflow



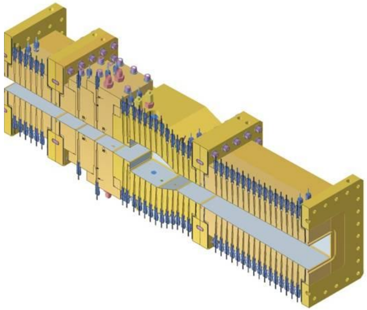


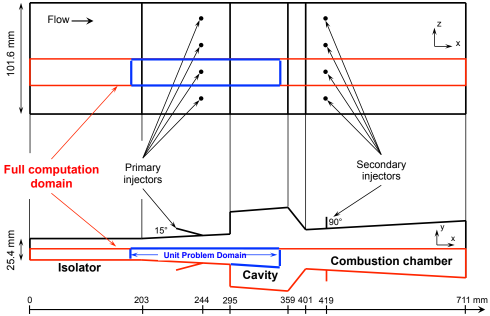
Computational domain of initial unit problem

Final: simulation of full combustor domain, match experimental setup —HIFiRE Direct Connect Rig (HDCR) [bottom-left figure]

Initial unit problem: primary injection section

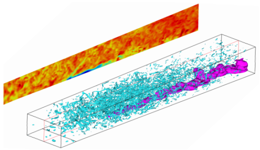
- omit cavity
- no combustion
- focus on interaction of fuel jet and supersonic air crossflow

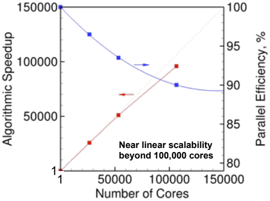


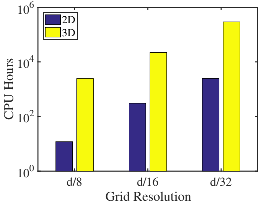


RAPTOR LES solver: model variants

RAPTOR: LES solver by Oefelein *et al.* at Sandia [Oefelein 06]







Highly-scalable but still **very expensive**

Motivation: different "model variants" are available under RAPTOR, trading off between solution accuracy and computational costs:

- Different grid resolutions
- Emulation using 2D geometry
- Static approximation of Smagorinsky model for SGS eddy viscosity
- Varying degree of detail in chemical mechanisms
- ...

⇒ to use/combine them, need to quantify the error due to model structure

Overview

Challenge:
LES is expensive, many flow solves needed to perform UQ and design

Would like to:
incorporate results from less expensive simulations of lower-fidelity models, while capturing model errors

Objective: characterize uncertainty due to model error resulting from using lower-fidelity models

Plan: represent the model error stochastically, by embedding a discrepancy term in the model parameters in a non-intrusive manner, where predictions automatically satisfy the governing equations

Embedded representation of model error

Traditional additive form: [Kennedy & O'Hagan 01]

$$q_k = \rho_k f_k(\lambda) + \delta_k + \epsilon_{d_k}$$

- Flexible for fitting model discrepancy
- Only applies corrections on model output, not on input parameters
- δ_k not transferable for prediction of QoIs outside calibration set
- Predictions generally do not obey governing equations
- Difficult to distinguish uncertainty contributions between model error and measurement noise

Embedding approach: [Sargsyan *et al.* 15]

$$q_k = f_k(\lambda + \delta_k) + \epsilon_{d_k}$$

$\Rightarrow$ physically-meaningful predictions that auto satisfy governing equations

Model error representation via polynomial chaos

Represent the discrepancy term δ in a stochastic manner using a polynomial chaos expansion (PCE):

$$\lambda + \delta(\alpha, \xi) = \lambda + \sum_{\beta \neq 0} \alpha_\beta \Psi_\beta(\xi)$$

$\tilde{\alpha} \equiv (\lambda, \alpha)$–calibration parameters ξ–aleatoric source (rep. model error)

Safer extrapolation of δ over k since is always a λ correction

---

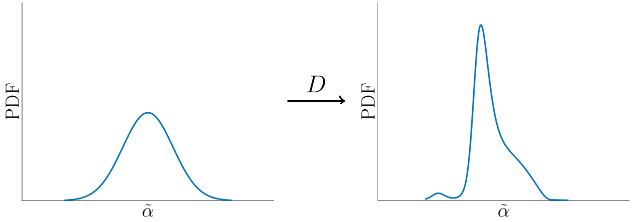
PCE in a nutshell: an expansion for a random variable:

$$\theta(\xi) = \sum_{\beta \in \mathcal{J}} c_\beta \Psi_\beta(\xi)$$

- c_β: PCE coefficients
- ξ: reference random vector (e.g., uniform, Gaussian)
- Ψ_β: multivariate orthonormal polynomial (e.g., Legendre, Hermite)
- β: multi-index, reflects order of polynomial basis

Calibrate by inferring all parameters $\tilde{\alpha} \equiv (\lambda, \alpha)$ via

Bayesian inference: $$\underbrace{p(\tilde{\alpha}|D)}_{posterior} \propto \underbrace{p(D|\tilde{\alpha})}_{likelihood} \underbrace{p(\tilde{\alpha})}_{prior}$$



Calibration data D is often higher-fidelity model simulations

$\Rightarrow$ capturing model error of **low-fidelity model w.r.t. high-fidelity model**

True likelihood is intractable

Posterior explored via **Markov chain Monte Carlo (MCMC)**

- adaptive Metropolis [Haario 01]
- efficient Gaussian proposal constructed from chain samples

MCMC requires likelihood evaluation $p(D|\tilde{\alpha})$, but have no analytical form

---

Enable tractable likelihood evaluation via two approximations:

1. Deterministic surrogate for low-fidelity model, built using regression

$$f_k(\cdot) \approx \hat{f}_k(\cdot) + \epsilon_{k,\mathrm{LOO}}$$

$\epsilon_{k,\mathrm{LOO}} \sim \mathcal{N}(0, \sigma^2_{k,\mathrm{LOO}})$ models the discrepancy between $\hat{f}_k$ and f_k, $\sigma^2_{k,\mathrm{LOO}}$ approximated from average leave-one-out cross validation error

Purpose: to replace the many subsequent low-fidelity model evaluations needed in MCMC; can skip if low-fidelity model inexpensive

Approximate likelihood using Gauss-marginal form

2 Gauss-marginal approximation to likelihood

$$p(D|\tilde{\alpha}) \approx \frac{1}{(2\pi)^{\frac{N}{2}}} \prod_{k=1}^{N} \frac{1}{\sigma_k(\tilde{\alpha})} \exp\left[-\frac{(\mu_k(\tilde{\alpha}) - D_k)^2}{2\sigma_k^2(\tilde{\alpha})}\right]$$

where $\mu_k(\tilde{\alpha})$ and $\sigma_k^2(\tilde{\alpha})$ are obtained by constructing a PCE for $\hat{f}_k$ at given $\tilde{\alpha}$ and in the argument of ξ:

$$\hat{f}_k(\lambda + \delta(\alpha, \xi)) = \hat{f}_k\left(\lambda + \sum_{\beta \neq 0} \alpha_\beta \Psi_\beta(\xi)\right) \approx \sum_{\beta} \hat{f}_{k,\beta}(\tilde{\alpha}) \Psi_\beta(\xi)$$

and so $\mu_k(\tilde{\alpha}) \approx \hat{f}_{k,0}(\tilde{\alpha})$ and $\sigma_k^2(\tilde{\alpha}) \approx \sum_{\beta \neq 0} \hat{f}_{k,\beta}^2(\tilde{\alpha})$

This PCE needs to be constructed at every $\tilde{\alpha}$ encountered in the MCMC, but not too expensive if we are working with $\hat{f}_k$ instead of f_k (use NISP)

Attribution of total predictive variance

A nice result: **attribute total predictive variance** to different sources

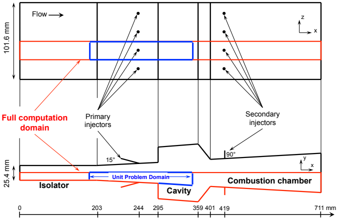
Recall
$$\begin{aligned} q_k &= f_k(\lambda + \delta(\alpha, \xi)) + \epsilon_{d_k} \\ &\approx \hat{f}_k(\lambda + \delta(\alpha, \xi)) + \epsilon_{k,\mathrm{LOO}} + \epsilon_{d_k} \\ &\approx \sum_{\beta} \hat{f}_{k,\beta}(\tilde{\alpha})\Psi_{\beta}(\xi) + \epsilon_{k,\mathrm{LOO}} + \epsilon_{d_k} \end{aligned}$$

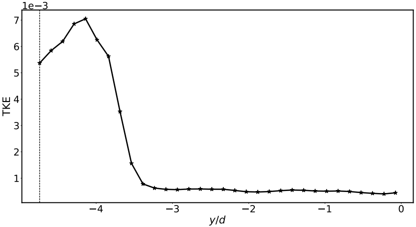
$$\mathrm{Var}[q_k] = \underbrace{\mathbb{E}_{\tilde{\alpha}}\left[\sigma_k^2(\tilde{\alpha})\right]}_{\text{model error}} + \underbrace{\mathrm{Var}_{\tilde{\alpha}}\left[\mu_k(\tilde{\alpha})\right]}_{\text{posterior uncertainty}} + \underbrace{\sigma_{k,\mathrm{LOO}}^2}_{\text{surrogate error}} + \underbrace{\sigma_{d_k}^2}_{\text{data noise}}$$

Dynamic-vs-Static Smagorinsky turbulence model

Calibrate static Smagorinsky model with dynamic treatment simulations

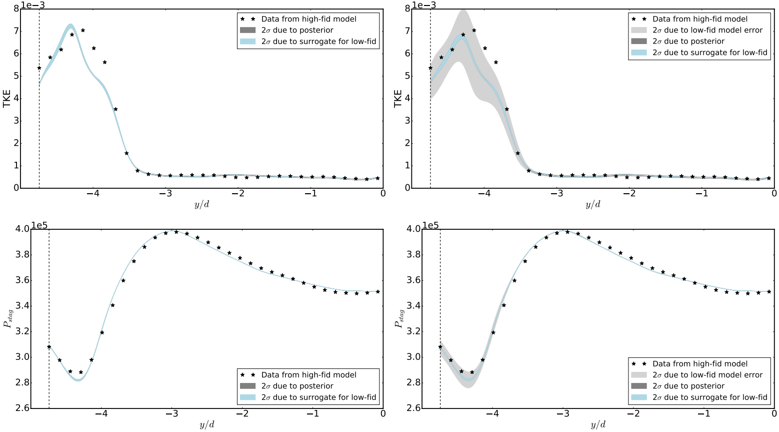
- 3D geometry
- Calibrate using TKE y-profile (t-averaged, at fixed x, centerline z)





- Embed in parameter $\lambda = C_R$
  - 1st-order expansion for $\delta = \alpha\xi$ (i.e., Gaussian)
- Surrogates: 500 regression points, 3rd-order PCEs

Dynamic-vs-Static Smagorinsky turbulence model



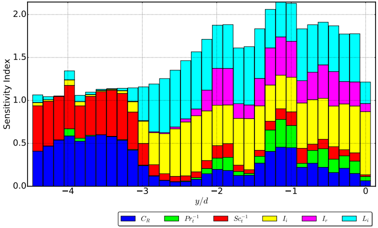
No model error treatment

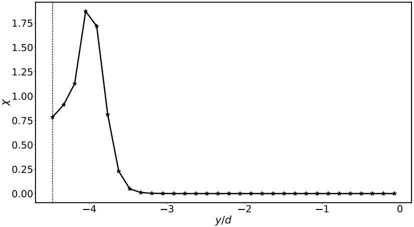
Embedded model error treatment

2D-vs-3D: choice of embedding parameters

Calibrate 2D model using 3D model simulations

- Calibrate using χ profile
- $\lambda = (C_R, Pr_t^{-1}, Sc_t^{-1}, I_i, I_r, L_i)$
- We do not want to embed δ for all λ, too many terms
    - Embed δ in select parameters
    - Employ triangular form of PCE for multivariate embedding
    - Target parameters where embedding is most “effective”
        - Global sensitivity analysis on calibration QoIs
        - Bayesian model selection (evidence computation)





2D-vs-3D: choice of embedding parameters

Calibrate 2D model using 3D model simulations

- Calibrate using χ profile
- $\lambda = (C_R, Pr_t^{-1}, Sc_t^{-1}, I_i, I_r, L_i)$
- We do not want to embed δ for all λ, too many terms
  - Embed δ in select parameters
  - Employ triangular form of PCE for multivariate embedding
  - Target parameters where embedding is most "effective"
    - Global sensitivity analysis on calibration QoIs
    - Bayesian model selection (evidence computation)

Embed Param	GSA $\bar{S}_{T_i}$	Log-evidence
$\mathbf{C_R}$	$\mathbf{5.24 \times 10^{-1}}$	$\mathbf{2.82 \times 10^{2}}$
Pr_t^{-1}	1.58×10^{-2}	-2.55×10^{3}
$\mathbf{Sc_t^{-1}}$	$\mathbf{4.90 \times 10^{-1}}$	$\mathbf{2.30 \times 10^{2}}$
I_i	3.63×10^{-2}	-9.68×10^{2}
I_r	2.24×10^{-3}	-3.74×10^{3}
L_i	5.32×10^{-2}	-4.15×10^{2}
C_R, Sc_t^{-1}		2.79×10^{2}

2D-vs-3D: choice of embedding parameters

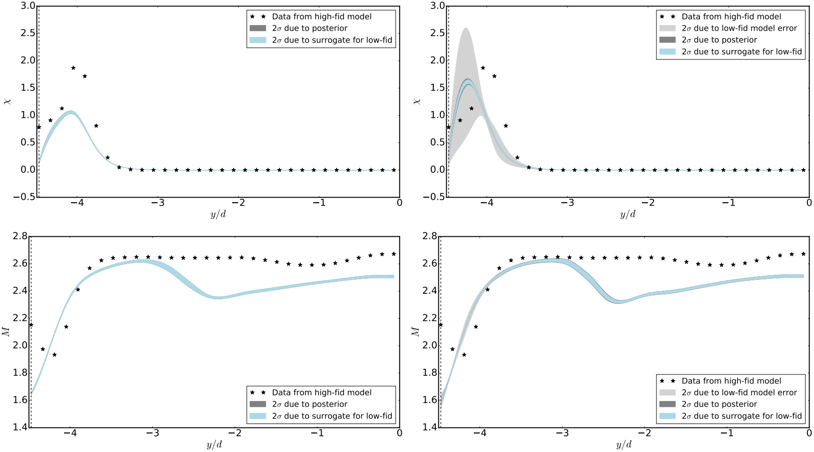
Calibrate 2D model using 3D model simulations

- Calibrate using χ profile
- $\lambda = (C_R, Pr_t^{-1}, Sc_t^{-1}, I_i, I_r, L_i)$
- We do not want to embed δ for all λ, too many terms
  - Embed δ in select parameters
  - Employ triangular form of PCE for multivariate embedding
  - Target parameters where embedding is most "effective"
    - Global sensitivity analysis on calibration QoIs
    - Bayesian model selection (evidence computation)

$\Rightarrow$ embed in C_R and Sc_t^{-1}

$$(\lambda + \delta(\alpha, \xi)) = \begin{cases} C_R + \alpha_{(1)}\xi_1 \\ Pr_t^{-1} \\ Sc_t^{-1} + \alpha_{(1,0)}\xi_1 + \alpha_{(0,1)}\xi_2 \\ I_i \\ I_r \\ L_i \end{cases}$$

2D-vs-3D: predictive quantities



No model error treatment

Embedded model error treatment

Conclusions

Conclusions:

- Introduced a framework for characterizing uncertainty from model error
  - embed discrepancy in model parameters; non-intrusive
  - predictions automatically satisfy governing equations
- Attributed total predictive variance to different contributing sources
- Illustrated good capture of model-to-model discrepancy, and also limitations when models are too different
- Demonstrated method in a non-reactive unit problem in scramjet design involving expensive LES:
  - Dynamic vs. static Smagorinsky model treatment
  - 2D vs. 3D geometry

Future work:

- Bayesian model selection for optimal model error embedding
- More sophisticated forms of embedding
- Full HDCR geometry with combustion

References I

Heikki Haario, Eero Saksman & Johanna Tamminen.
An adaptive Metropolis algorithm.
Bernoulli, vol. 7, no. 2, pages 223–242, 2001.

Marc. C. Kennedy & Anthony O'Hagan.
Bayesian calibration of computer models.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), vol. 63, no. 3, pages 425–464, 2001.

Joseph C. Oefelein.
Large eddy simulation of turbulent combustion processes in propulsion and power systems.
Progress in Aerospace Sciences, vol. 42, no. 1, pages 2–37, 2006.

Khachik Sargsyan, Habib N. Najm & Roger G. Ghanem.
On the Statistical Calibration of Physical Models.
International Journal of Chemical Kinetics, vol. 47, no. 4, pages 246–276, 2015.