SAND2020-10821C

Uncertainty Propagation Using Conditional Random Fields in Large-Eddy Simulations of Scramjet Computations

Xun Huan* and Cosmin Safta†
Sandia National Laboratories, Livermore, California 94550

Zachary P. Vane‡
Northrop Grumman Corporation, Sunnyvale, California 94086

Guilhem Lacaze§
Space Exploration Technologies Corporation, Hawthorne, California 90250

Joseph C. Oefelein¶
Georgia Institute of Technology, Atlanta, Georgia 30332

Habib N. Najm‖
Sandia National Laboratories, Livermore, California 94550

The development of scramjet engines is crucial for attaining efficient and stable propulsion under hypersonic flight conditions. Design for well-performing scramjet engines requires accurate flow simulations in conjunction with uncertainty quantification (UQ). We advance the start-of-the-art in bringing together UQ and large-eddy simulations for scramjet computations, with a focus on the HIFiRE Direct Connect Rig combustor. In particular, we perform uncertainty propagation for spatially dependent "field" quantities of interest by treating them as random fields, and numerically compute low-dimensional Karhunen-Loève expansions (KLEs) using a finite number of simulations on non-uniform grids. We also describe a formulation and procedure to extract *conditional* KLEs that characterize the stochasticity induced by uncertain parameters at given designs. This is achieved by building a single KLE via samples drawn *jointly* from the parameter and design spaces, and leverage polynomial chaos expansions to insert input dependencies into the KLE. The ability to access conditional KLEs will be immensely useful for subsequent efforts in design optimization under uncertainty.

I. Nomenclature

$\sim \mathcal{U}(a,b)$	=	uniform distribution from a to b
a_1	=	primary injector angle, degrees
C_2H_4	=	ethylene
C_R	=	modified Smagorinsky constant
c_{β^n}	=	coefficient for the nth basis function
d	=	injector diameter, mm
f_{NO}	=	ratio between volume fractions of nitrogen and oxygen in the oxidizer streams
I_i, I_f	=	inlet and fuel turbulence intensity magnitudes
$\mathcal{J}$	=	index set
$K(\cdot,\cdot), K$	=	covariance kernel and matrix
k_t	=	truncation index

*Postdoctoral Appointee, Combustion Research Facility, MS9051. Member AIAA.
†Principal Member of Technical Staff, Quantitative Modeling and Analysis, MS9159. Senior Member AIAA.
‡Aerospace Engineer. Member AIAA.
§Principal Numerical Physicist. Member AIAA.
¶Professor, School of Aerospace Engineering, 270 Ferst Drive NW. Associate Fellow AIAA.
‖Distinguished Member of Technical Staff, Combustion Research Facility, MS9051. Member AIAA.

L_i, L_f	=	inlet and fuel turbulence length scales, mm
M	=	Mach number
M_0, M_f	=	inlet and fuel Mach numbers
N	=	number of samples
n_d	=	number of design variables
n_g	=	number of grid points
n_p	=	number of uncertain parameters
n_t	=	number of expansion terms
Pr_t	=	turbulent Prandtl number
p	=	polynomial degree
p_0	=	inlet stagnation pressure, MPa
Q	=	eigenvector matrix
$q_k(\cdot), q_k$	=	kth eigenfunction and eigenvector
R_i	=	ratio of turbulence intensity vertical to horizontal components
S	=	sample matrix
Sc_t	=	turbulent Schmidt number
T	=	temperature, K
T_0	=	inlet stagnation temperature, K
U	=	left-singular matrix
V	=	right-singular matrix
w_j, W	=	weight for the jth point, weight matrix
x, y, z	=	streamwise, wall-normal, and spanwise coordinates, mm
$x/d, y/d, z/d$	=	streamwise, wall-normal, and spanwise coordinate normalized by injector diameter
$\mathbf{x}$	=	vector of spatial coordinates, mm
x_1	=	primary injector x-location, m
x_2	=	secondary injector x-location, m
Y_{CO}	=	carbon monoxide mass fraction
β	=	multi-index
δ_{kl}	=	Kronecker delta
ζ_k	=	random variables in Karhunen-Loève Expansion
Λ	=	eigenvalue matrix
λ_k	=	kth eigenvalue
μ	=	regularization parameter
ξ_j	=	polynomial chaos expansion random variables
Σ	=	singular value matrix
ϕ_G	=	global equivalence ratio (total of primary and secondary)
ϕ_R	=	ratio of primary to secondary injector equivalence ratios
$\Psi_\beta, \psi_{\beta_j}$	=	multivariate and univariate orthonormal polynomial basis functions
ω	=	sample space random event

II. Introduction

Research in powered hypersonic flight has thrived in the past decades with strong interests from both military and civilian aerospace applications [1, 2]. Among others, one significant technical challenge involves the need for propulsion systems that can sustain operations under hypersonic flight conditions (Mach > 5 and typically below low Earth orbit) in an efficient and stable manner [2]. In these environments, one type of air-breathing propulsion systems known as *supersonic combustion ramjet (scramjet)* engines seeks to burn fuel using atmospheric air at supersonic speeds, and so sidesteps the need for carrying on-board oxidizer or decelerating airflow to subsonic speeds. Scramjets thus can potentially achieve much higher efficiencies compared to traditional technologies such as rockets or turbojets.

The design for well-performing scramjet engines is still in its early stages. It faces major scientific difficulties in characterizing and predicting combustion properties for multiscale and multiphysics turbulent flows under extreme conditions, and whereby mixing and combustion must occur on time scales on the order of milliseconds while flow in the combustion chamber is supersonic. Designing an optimal engine typically involves maximizing combustion efficiency while minimizing pressure losses, thermal loading, and the risk of unstart and flame blow-out. Achieving this

while also producing designs that are robust and reliable in the presence of uncertainty and noise presents an extremely challenging undertaking. An important step for advancing scramjet design thus involves conducting accurate flow simulations together with uncertainty quantification (UQ).

While UQ in general has received substantial attention in the past decades, UQ for scramjet applications is much less developed but gaining traction in recent years [3–10]. A comprehensive assessment of uncertainty in such systems has been prohibitive due to the high cost of simulating turbulent reacting flows compounded with the multi-query nature of UQ investigations that generally requires some form of exploration in the stochastic space. These challenges are exacerbated when one wishes to carry out large-eddy simulations (LES), which, although more computationally intensive, can describe detailed turbulent flows and features often not accessible through the more widely used Reynolds-averaged Navier-Stokes (RANS) models. Indeed, one of several recognized computational grand challenges of LES for scramjets is "the quantification of associated uncertainties in the computational results with regard to both aleatoric and epistemic errors in input simulation parameters and physical models" [2].

We advance the start-of-the-art in bringing together UQ and LES for scramjet computations. While our previous work [6–10] explored different aspects of this endeavor, most of them studied an initial unit test problem focusing on the jet-in-crossflow physics with reaction switched off. This paper thus presents new results with combustion enabled. We also set the scope of our UQ investigation to performing uncertainty propagation for spatially dependent "field" quantities of interest (QoIs). We treat these QoIs as random fields (RFs), and numerically compute Karhunen-Loève expansions (KLEs) (see e.g., [11]) using a finite number of simulations on non-uniform grids. A truncated KLE treatment is advantageous compared to a direct grid-discretized random variable approach, since a RF can effectively capture correlations across spatial dimensions while KLE further reduces the stochastic degrees of freedom (DOFs) needed for an accurate representation of the RF toward numbers much smaller than the number of grid points. We also describe a formulation and procedure to extract *conditional* KLEs that characterize the stochasticity induced by uncertain parameters at given designs. A naive approach may entail repeating the KLE construction multiple times at select design points and interpolate, which would be computationally burdensome. Instead, we build a single KLE via samples drawn *jointly* from the parameter and design spaces, and leverage polynomial chaos expansions (PCEs) to insert input dependencies into the KLE [11, 12]. Once the joint KLE is constructed, we can then assess statistical properties and generate samples of the RF *at different designs*. This will be immensely useful as a statistically consistent surrogate model for subsequent efforts in optimal experimental design and design optimization under uncertainty. In summary, the key contributions of this paper are:

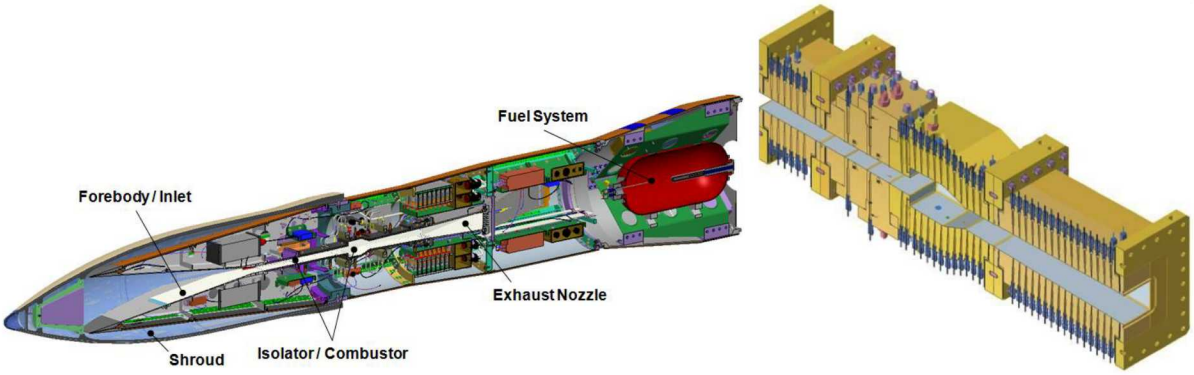
1) to illustrate advances in bringing together UQ and LES for scramjet computations;
2) to perform uncertainty propagation for spatially dependent "field" QoIs using RF treatment and low-dimensional KLE representation;
3) to demonstrate the formulation and procedure for extracting conditional KLEs that characterize the stochasticity induced by uncertain parameters at given design variables.

This paper is organized as follows. Section III describes the physical setup and LES solver. We then introduce KLEs in Sec. IV for representing RFs of spatially dependent QoIs, including their numerical construction from samples and the procedure for extracting conditional KLEs. Descriptions of simulation results along with their KLEs are presented in Sec. V. Lastly, the paper concludes with summary and discussions in Sec. VI.

III. Large-Eddy Simulations for the HIFiRE Direct Connect Rig

We concentrate on a scramjet configuration studied under the HIFiRE (Hypersonic International Flight Research and Experimentation) program [13, 14], which has been the target of a mature experimental campaign through its HIFiRE Flight 2 (HF2) project [15–17]. The HF2 payload, depicted in Fig. 1(a), consists of a cavity-based hydrocarbon-fueled dual-mode scramjet, and was tested under flight conditions of Mach 6–8+. A ground test rig, designated the HIFiRE Direct Connect Rig (HDCR) (Fig. 1(b)), was developed to duplicate the isolator/combustor layout of the flight test hardware, and to provide ground-based measurements for comparisons with flight test data, verifying engine performance and operability, and designing fuel delivery schedule [18, 19]. While data from flight tests are not publicly released, HDCR ground test data are available [18, 20]. Therefore, we aim to simulate and assess reactive flows inside the HDCR, and intend to leverage existing experimental datasets to drive modeling developments in the future.

The HDCR consists of a constant-area isolator (planar duct) attached to a combustion chamber. It includes four primary injectors that are mounted upstream of flame stabilization cavities on both the top and bottom walls. Four secondary injectors along both walls are positioned downstream of the cavities. Flow travels from left to right in the x-direction (streamwise), and the geometry is symmetric about the centerline in the y-direction. We take advantage



(a) HIFiRE Flight 2 payload (b) HDCR

Fig. 1 HIFiRE Flight 2 payload [16] and HDCR cut views [18].

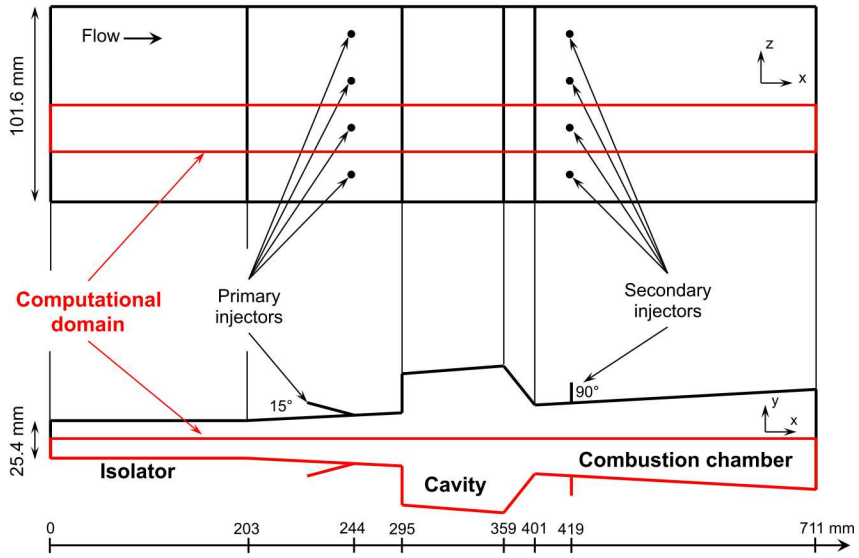


Fig. 2 The HDCR experimental setup and schematic of the full computational domain.

of this symmetry in our numerical simulations by considering a domain that covers only the bottom half of the rig. To further reduce the computational cost, we consider one set of primary/secondary injectors and impose periodic conditions in the z-direction (spanwise). The overall computational domain is highlighted by the red lines in Fig. 2.

The fuel supplied through the injectors is a gaseous methane-ethylene mixture containing 36% CH_4 and 64% C_2H_4 by volume, which acts as a surrogate with similar combustion properties as JP-7 [21]. A reduced, three-step mechanism [22, 23] is initially employed to characterize the combustion process:

$$CH_4 + 2(O_2 + f_{NO}N_2) \rightarrow CO_2 + 2H_2O + 2f_{NO}N_2 \tag{1}$$

$$C_2H_4 + 3(O_2 + f_{NO}N_2) \rightarrow 2CO_2 + 2H_2O + 3f_{NO}N_2 \tag{2}$$

$$2CO + O_2 \rightarrow 2CO_2, \tag{3}$$

where $f_{NO} = 0.79/0.21$ is the ratio between volume fractions of N_2 and O_2 in the oxidizer streams. Arrhenius formulations of the kinetic reaction rates are adopted, and the parameters are selected to retain robust/stable combustion in the current simulations.

LES calculations are then performed using the RAPTOR code framework developed by Oefelein [24, 25]. The solver has been optimized to meet the strict algorithmic requirements imposed by the LES formalism. The theoretical

framework solves the fully coupled conservation equations of mass, momentum, total-energy, and species for a chemically reacting flow, while accounting for detailed thermodynamics and transport processes at the molecular level, and under high Reynolds number, high-pressure, real-gas and/or liquid conditions over a wide Mach operating range. RAPTOR is designed specifically for LES using non-dissipative, discretely conservative, staggered, finite-volume differencing. This eliminates numerical contamination of the subfilter models due to artificial dissipation and provides discrete conservation of mass, momentum, energy, and species, which is imperative for high quality LES.

Lastly, we would like to point out some limitations of our numerical results in the current paper stemming from additional simplifications necessitated by practical considerations. In particular, constraints on computational resources both encouraged and compelled an initial UQ investigation involving simulations in a two-dimensional geometry, where we placed a single cell in the z-direction at a x-y plane intersecting the injectors. We fully acknowledge the decreased fidelity of these runs as a result of the reduced geometric description as well as the relatively simple chemical model in Eqn. (1) to (3). Indeed, certain physical features and phenomenon are eroded or otherwise not representable in a two-dimensional setting. Nonetheless, given the scale of computations demanded by any form of uncertainty assessments, performing UQ even under these emulatory settings has not been achieved previously. At the same time, fully three-dimensional simulations are computationally possible but only for relatively coarse grids and where only a small number of runs can be completed under the present computational budget; they are thus not ready to support a meaningful UQ demonstration, and are part of our future work.

IV. Conditional Karhunen-Loève Expansions for Random Fields

A. General Formulation

We seek to construct KLEs for spatially dependent field QoIs computed from the RAPTOR LES code described in Sec. III. We begin by making a brief introduction of KLEs in the general setting without making the distinction of conditioning; mathematical details can be found in references such as [111]. A field QoI, for example temperature $T(\mathbf{x},\omega)$, is treated as a RF that depends on both the spatial coordinates $\mathbf{x} \equiv (x,y,z) \in \mathcal{X}$ and sample space random event ω which induces the stochasticity of the RF (we will make this more precise when we discuss the conditioning in Sec. IV.C). Without loss of generality, we restrict ourselves to centered (zero mean) RF where, for example, the mean behavior may be subtracted out from dataset samples during preprocessing. When the RF is square integrable and continuous in the mean square sense, it can be expressed via a spectral decomposition known as the KLE:

$$T(\mathbf{x},\omega) = \sum_{k=1}^{\infty} \sqrt{\lambda_k}\, q_k(\mathbf{x})\zeta_k(\omega), \tag{4}$$

where $\zeta_k(\omega)$ are mutually uncorrelated random variables with zero mean and unit variance. $\lambda_k \geq 0$ and $q_k(\mathbf{x})$ are eigenvalues and eigenfunctions, respectively, of the RF covariance function between two locations $\mathbf{x}_i$ and $\mathbf{x}_j$:

$$K(\mathbf{x}_i,\mathbf{x}_j) \equiv \mathbb{E}\left[T(\mathbf{x}_i,\omega),T(\mathbf{x}_j,\omega)\right], \tag{5}$$

and they are solutions to the homogeneous Fredholm equation of the second kind:

$$\int_{\mathcal{X}} K(\mathbf{x}_i,\mathbf{x}_j) q_k(\mathbf{x}_j)\, d\mathbf{x}_j = \lambda_k q_k(\mathbf{x}_i), \qquad k = 1,2,\ldots. \tag{6}$$

The eigenfunctions, when properly normalized, satisfy

$$\int_{\mathcal{X}} q_k(\mathbf{x}) q_l(\mathbf{x})\, d\mathbf{x} = \delta_{kl}, \qquad k,l = 1,2,\ldots \tag{7}$$

where δ_{kl} is the Kronecker delta. The KLE can be shown to be optimal in the sense that a finite truncation under decreasing eigenvalues yields the minimum mean square error equaling to the sum of truncated eigenvalues. Thus, a k_t-term truncated KLE can provide an accurate representation of the RF using a finite number of random variables $\zeta_k(\omega), k = 1,\ldots,k_t$, where k_t is often much smaller than the number of grid points (i.e., the dimension of a grid discretized RF) especially when rapid spectrum decay takes place (which occurs under high correlation settings).

The construction of a RF thus involves characterizing λ_k, q_k, and ζ_k, which we will compute numerically from a dataset of samples of the targeted field QoI.

B. Computing Eigenvalues λ_k and Eigenvectors q_k

Let $\{\mathbf{x}_j\}_{j=1}^{n_g}$ represents the grid locations where the field variable is stored. Associated with these points are also quadrature weights $w_j > 0$ that reflects the local cell area (for example, a uniform grid would have $w_j = 1/n_g$). We begin by establishing the spatially discretized version of Eqn. (6) and Eqn. (7):

$$\int_{\mathcal{X}} K(\mathbf{x}_i, \mathbf{x}_j) q_k(\mathbf{x}_j)\, d\mathbf{x}_j = \lambda_k q_k(\mathbf{x}_i) \qquad \longrightarrow \qquad \sum_{j=1}^{n_g} w_j K_{ij} q_{k,j} = \lambda_k q_{k,i}, \qquad i = 1, \ldots, n_g \tag{8}$$

$$\int_{\mathcal{X}} q_k(\mathbf{x}) q_l(\mathbf{x})\, d\mathbf{x} = \delta_{kl} \qquad \longrightarrow \qquad \sum_{j=1}^{n_g} w_j q_{k,j} q_{l,j} = \delta_{kl}, \tag{9}$$

for $k, l = 1, 2, \ldots$, and where we adopted notations $K_{ij} \equiv K(\mathbf{x}_i, \mathbf{x}_j)$ and $q_{k,j} \equiv q_k(\mathbf{x}_j)$. With a total of $N < \infty$ samples, the covariance matrix can be forged from sample covariance of our dataset via the formula

$$K_{ij} = \frac{1}{N-1} \sum_{n=1}^{N} T(\mathbf{x}_i, \omega^{(n)}) T(\mathbf{x}_j, \omega^{(n)}), \qquad i, j = 1, \ldots, n_g, \tag{10}$$

where the RF is assumed to be centered in preprocessing and so the mean subtrahend is zero, and the superscript $(\cdot)^{(n)}$ represents the nth sample from the dataset. Equation (8) in matrix form is then

$$(KW)Q = Q\Lambda, \tag{11}$$

where $K \in \mathbb{R}^{(n_g \times n_g)}$, $W \in \mathbb{R}^{(n_g \times n_g)}$ is the diagonal quadrature weight matrix with $W_{jj} = w_j$, and Q and Λ are the right eigenvector matrix and the eigenvalue matrix, respectively. Two comments are warranted at this point. First, proper scaling of the eigenvectors is required in order to attain orthonormality with respect to W, as required by Eqn. (9). Second, while $K \in \mathbb{R}^{(n_g \times n_g)}$, the number of nonzero eigenvalues (assuming all samples are unique) equals its rank, $\min(n_g, N-1)$, as a consequence of discretization and finite sampling. In practice, good accuracy can often be retained while truncating the expansion further with

$$T(\mathbf{x}, \omega) \approx \sum_{k=1}^{k_t} \sqrt{\lambda_k} q_k(\mathbf{x}) \zeta_k(\omega). \tag{12}$$

One practical rule of thumb is to select the smallest k_t such that λ_{k_t} has decayed to some fraction (e.g., e^{-1}) of the largest eigenvalue λ_1. One may also choose the smallest k_t so that the retained energy $\sum_{k=1}^{k_t} \lambda_k$ achieves some large percentage (e.g., 90%) of the total energy $\sum_{k=1}^{n_g} \lambda_k$, but this requires computing all λ_k. When the spectrum decay is rapid, k_t is often much smaller than either n_g or N.

Lastly, we note that the eigenvalue problem in Eqn. (11) can be solved through singular value decomposition (SVD), since the (weighted) covariance matrix (with zero mean) may be expressed as

$$KW = \left(\frac{1}{\sqrt{N-1}} W^{\frac{1}{2}} S \right) \left(\frac{1}{\sqrt{N-1}} S^T W^{\frac{T}{2}} \right) \tag{13}$$

where $S \equiv \left[T(\mathbf{x}, \omega^{(1)}), \ldots, T(\mathbf{x}, \omega^{(n)}) \right] \in \mathbb{R}^{(n_g \times N)}$ is the sample matrix. With the SVD: $\frac{1}{\sqrt{N-1}} W^{\frac{1}{2}} S = U \Sigma V^T$, the nonzero singular values in Σ are square root of nonzero eigenvalues in Λ, and the columns of U are eigenvectors in Q. With the truncation in Eqn. (12), iterative algorithms would be advantageous to target only the leading modes. Furthermore, since our study with scramjet computations deals with expensive simulations with large grids and few available simulation samples—that is, $n_g \gg N$—working with the S matrix of size $(n_g \times N)$ involves less intensive memory requirements than the $(n_g \times n_g)$ system in Eqn. (11).

C. Conditional Karhunen-Loève Expansions by Representing ζ_k Using Polynomial Chaos Expansions

Once λ_k and q_k are obtained, we can recover the N samples of the uncorrelated random variables $\zeta_k(\omega^{(n)})$ for each k, from our dataset through

$$\zeta_k(\omega) = \frac{1}{\lambda_k} \int_{\mathcal{X}} T(\mathbf{x}, \omega) q_k(\mathbf{x})\, d\mathbf{x} \qquad \longrightarrow \qquad \zeta_k^{(n)} = \frac{1}{\lambda_k} \sum_{j=1}^{n_g} w_j T(\mathbf{x}_j, \omega^{(n)}) q_{k,j}, \qquad n = 1, \ldots, N. \tag{14}$$

One may use these samples, for example, to perform a kernel density estimate (KDE) for ζ_k, from which new samples can be generated and Eqn. (12) evaluated for subsequent uncertainty analysis. However, such an approach would not build a relationship to the uncertain parameters and design variables, i.e., the model inputs, explicitly. Since ultimately we would like to access *conditional* KLEs that are functions of design variables and capture the stochasticity induced by uncertain parameters at given designs, we need to be more specific about the role of ω.

In the presentation so far, ω is a collective embodiment of all sources of stochasticity, and the dataset samples are perceived as different realizations of this process. To be more specific, in this work, our samples are generated by performing simulations at jointly randomized uncertain parameters as well as design variables, and so we wish to map ω into those two contributions. One possible approach is to represent $\zeta_k(\omega)$ as finitely-truncated PCEs (detailed descriptions of PCEs can be found in references such as [11, 26–28]):

$$\zeta_k(\omega) = \sum_{\beta \in \mathcal{J}} c_\beta \Psi_\beta(\xi_1, \ldots, \xi_{n_p}, \xi_{n_p+1} \ldots, \xi_{n_p+n_d}), \tag{15}$$

where c_β are the expansion coefficients, $\beta = (\beta_1, \ldots, \beta_{n_p+n_d})$, $\forall \beta_j \in \mathbb{N}_0$, is a multi-index, $\mathcal{J}$ is some finite index set, n_p is the number of uncertain parameters, n_d is the number of design variables, ξ_j are a chosen set of independent random variables, and $\Psi_\beta(\xi_1, \ldots, \xi_{n_p}, \xi_{n_p+1} \ldots, \xi_{n_p+n_d})$ are multivariate polynomials of the product form

$$\Psi_\beta(\xi_1, \ldots, \xi_{n_p}, \xi_{n_p+1} \ldots, \xi_{n_p+n_d}) = \prod_{i=1}^{n_p} \psi_{\beta_i}(\xi_i) \prod_{j=n_p+1}^{n_p+n_d} \psi_{\beta_j}(\xi_j), \tag{16}$$

with ψ_{β_i} being degree-β_i polynomials orthonormal with respect to the probability density function of ξ_i (i.e., $p(\xi_i)$):

$$\mathbb{E}\left[\psi_k(\xi_j)\psi_n(\xi_j)\right] = \int_{\Xi_j} \psi_k\left(\xi_j\right) \psi_n\left(\xi_j\right) p\left(\xi_j\right)\, d\xi_j = \delta_{kn}. \tag{17}$$

While different choices of ξ_j and ψ_{β_j} are available under the generalized Askey family [29], our application will use only Legendre expansions with uniform $\xi_j \sim \mathcal{U}(-1, 1)$ since our parameters and design variables are endowed with uniform measures and so linear transformations between them to ξ_j can be conveniently established. Furthermore, we focus on total-order expansions of degree p with $\mathcal{J} = \{\beta : \| \beta \|_1 \leq p\}$ for simplicity. The total number of basis terms for a total-order expansion is $\frac{(n_p+n_d+p)!}{(n_p+n_d)!p!}$.

Since we have samples $\zeta_k^{(n)}, n = 1, \ldots, N$ from Eqn. (14), we may estimate PCE coefficients c_β via regression. This involves solving c in the following regression linear system $Ac = b$:

$$\underbrace{\begin{bmatrix} \Psi_{\beta^1}(\xi^{(1)}) & \cdots & \Psi_{\beta^{n_t}}(\xi^{(1)}) \\ \vdots & & \vdots \\ \Psi_{\beta^1}(\xi^{(N)}) & \cdots & \Psi_{\beta^{n_t}}(\xi^{(N)}) \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} c_{\beta^1} \\ \vdots \\ c_{\beta^{n_t}} \end{bmatrix}}_{c} = \underbrace{\begin{bmatrix} \zeta_k^{(1)} \\ \vdots \\ \zeta_k^{(N)} \end{bmatrix}}_{b}, \tag{18}$$

where the notation Ψ_{β^i} refers to the ith basis function for a total of n_t terms, c_{β^i} is the coefficient corresponding to that basis, and $\xi^{(n)}$ is the nth regression point (here ξ without subscript represents the vector of all ξ_j). The Karhunen-Loève theorem states that ζ_k are zero mean, unit variance, and mutually uncorrelated. We thus remove the constant basis term to impose zero mean. Unit variance is enforced through

$$\mathrm{Var}(\zeta_k) = \sum_{0 \neq \beta \in \mathcal{J}} c_\beta^2 = 1. \tag{19}$$

Zero mutual correlation is more difficult to achieve, as it requires

$$\mathbb{E}\left[\zeta_k \zeta_l\right] = \sum_{0 \neq \beta \in \mathcal{J}} c_{k,\beta} c_{l,\beta} = \mathbb{E}\left[\zeta_k\right] \mathbb{E}\left[\zeta_l\right] = 0, \qquad k, l = 1, \ldots, k_t \text{ and } k \neq l, \tag{20}$$

where the first equality can be obtained by applying Eqn. (2.11) and (2.22) in [30], the second equality is the condition for no correlation, and the last equality is due to the zero-mean properties of these random variables. A is thus the

regression matrix where each column corresponds to a basis (except the constant basis term) and each row corresponds to a regression sample point. The number of columns n can easily become quite large in high-dimensional settings; for example, a total-order expansion of degree 3 in 16 dimensions contains $n_t = \frac{(3+16)!}{3!16!} - 1 = 969$ terms. In fact, we often encounter situations where the number of samples N may be much smaller than n_t. We thus employ compressive sensing [31, 32] to discover sparse structures in PCEs by finding a sparse solution for the (potentially underdetermined) system in Eqn. (18). Specifically, we target the unconstrained least absolute shrinkage and selection operator (LASSO) form:

$$\min_{c} \| Ac - b \|_2^2 + \mu \| c \|_1 \qquad \text{subject to Eqn. (19) and (20),} \tag{21}$$

where μ is a regularization constant. We solve the LASSO problem numerically using the alternating direction method of multipliers (ADMM) [33, 34] with modification to include constraints Eqn. (19) and (20), and select μ based on cross-validation techniques (such as in the manner utilized by [35]).

V. Numerical Results

In our numerical studies, we designate 11 parameters to bear uncertainty (uncertain parameters), and allow a separate set of 5 parameters to vary and reflect the scramjet design configuration (design variables). The total model input space is thus 16 dimensional, and tabulated in Table 1. The uncertain parameters reflect uncertainty in inlet and fuel inflow boundary conditions as well as turbulence model parameters for the Smagorinsky model. With lower and upper bounds suggested by domain experts, we invoke the maximum entropy principle [36, 37] and endowed the parameters with uninformative uniform "prior" distributions across the ranges indicated in the table. While there are no intrinsic distributions associated with the design variables, we view them as random variables in the exploratory stage. Without any prior information to favor particular regions of the design space, we also sample the design variables in accordance with uniform measures[*].

We will perform our investigations on multiple field (spatial) QoIs, such as temperature (T), Mach number (M), mass fractions of chemical species (e.g., Y_{CO}), etc. All QoIs are time-averaged variables unless indicated otherwise. The simulation data utilized in the current analysis are from two-dimensional scramjet computations, employing grid resolutions where cell sizes are $1/8$ (coarse), $1/16$ (medium), and $1/32$ (fine) of the injector diameter $d = 3.175$ mm. The run lengths at different grid resolutions are selected to try to keep the total time in the problem context roughly equal, while considering practical constraints of computational resources. While the timestep sizes are determined adaptively based on guidance from the Courant-Friedrichs-Lewy (CFL) condition, their values for the next refined grid resolution roughly decrease by a factor of two. The simulations are thus performed for 2×10^5 iterations for the coarse grid, 4×10^5 for medium, and 8×10^5 for fine, from their respective warm-start solutions that were engineered from a quasi-steady state nominal condition simulation, and the last halves of these iterations are used for time-averaging. The total number of simulations performed in establishing our database, as well as their average central processing unit (CPU) times, are shown in Table 2.

Figures 3 to 5 present the sample mean and variance fields for T [K], M, and Y_{CO}, and the three figures correspond to the coarse ($N = 1053$ samples), medium ($N = 222$ samples), and fine ($N = 23$ samples) grids, respectively.

The mean and variance are computed from our database of runs in Table 2, which consists of simulations conducted at input samples drawn *jointly* from the parameter and design spaces in accordance to Table 1. Together with effects of time-averaging, these plots thus are not meant to display detailed time-dependent LES flow features, but rather only a summary of statistical behaviors. We emphasize that working directly with these samples cannot produce *conditional* statistics for any given design variables, which would be needed to perform subsequent design optimization. We thus produce conditional KLE using the tools introduced in this paper.

For the full version of this paper, we will include the main KLE results, including examples of KLE modes, eigenvalue decay, generated sample fields, and also conditional statistics at select design values.

[*] A more appropriate measure for the design space would be one proportional to the frequency or probability of where the optimizer would visit, since we would like more samples (accuracy) in areas evaluated more often. However, it is heavily dependent on the problem and the numerical methods employed, and certainly difficult to assess *a priori*.

Parameter	Range	Description
Inlet boundary conditions:		
p_0	$[1.406, 1.554] \times 10^6$ Pa	Stagnation pressure
T_0	$[1472.5, 1627.5]$ K	Stagnation temperature
M_0	$[2.259, 2.761]$	Mach number
I_i	$[0, 0.05]$	Turbulence intensity horizontal component
R_i	$[0.8, 1.2]$	Ratio of turbulence intensity vertical to horizontal components
L_i	$[0, 8] \times 10^{-3}$ m	Turbulence length scale
Fuel inflow boundary conditions:		
I_f	$[0, 0.05]$	Turbulence intensity magnitude
L_f	$[0, 1] \times 10^{-3}$ m	Turbulence length scale
Turbulence model parameters:		
C_R	$[0.01, 0.06]$	Modified Smagorinsky constant
Pr_t	$[0.5, 1.7]$	Turbulent Prandtl number
Sc_t	$[0.5, 1.7]$	Turbulent Schmidt number
Design variables:		
ϕ_G	$[0.5, 0.8]$	Global equivalence ratio (total of primary and secondary)
ϕ_R	$[0.25, 0.35]$	Ratio of primary to secondary injector equivalence ratios
x_1	$[0.231, 0.2564]$ m	Primary injector x-location
x_2	$[0.40755, 0.43295]$ m	Secondary injector x-location
a_1	$[5, 25]^\circ$	Primary injector angle

Table 1 Uncertain model parameters and design variables. The distributions are assumed uniform across the ranges shown.

VI. Conclusions

Acknowledgments

Support for this research was provided by the Defense Advanced Research Projects Agency (DARPA) program on Enabling Quantification of Uncertainty in Physical Systems (EQUiPS). This research used resources of the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy Office of Science User Facility operated under Contract No. DE-AC02-05CH11231. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department Of Energy or the United States Government.

References

[1] Schmisseur, J. D., "Hypersonics into the 21st century: A perspective on AFOSR-sponsored research in aerothermodynamics," *Progress in Aerospace Sciences*, Vol. 72, 2015, pp. 3–16. doi:10.1016/j.paerosci.2014.09.009.

[2] Urzay, J., "Supersonic Combustion in Air-Breathing Propulsion Systems for Hypersonic Flight," *Annual Review of Fluid Mechanics*, Vol. 50, No. 1, 2018, pp. 593–627. doi:10.1146/annurev-fluid-122316-045217.

Grid resolution	Run length (iterations)	Approximate CPU hours per run	Total number of runs
Coarse	2×10^5	1.7×10^3	1053
Medium	4×10^5	1.1×10^4	222
Fine	8×10^5	3.9×10^4	23

Table 2 Run lengths, CPU hours per run, and total number of runs performed in establishing our database.

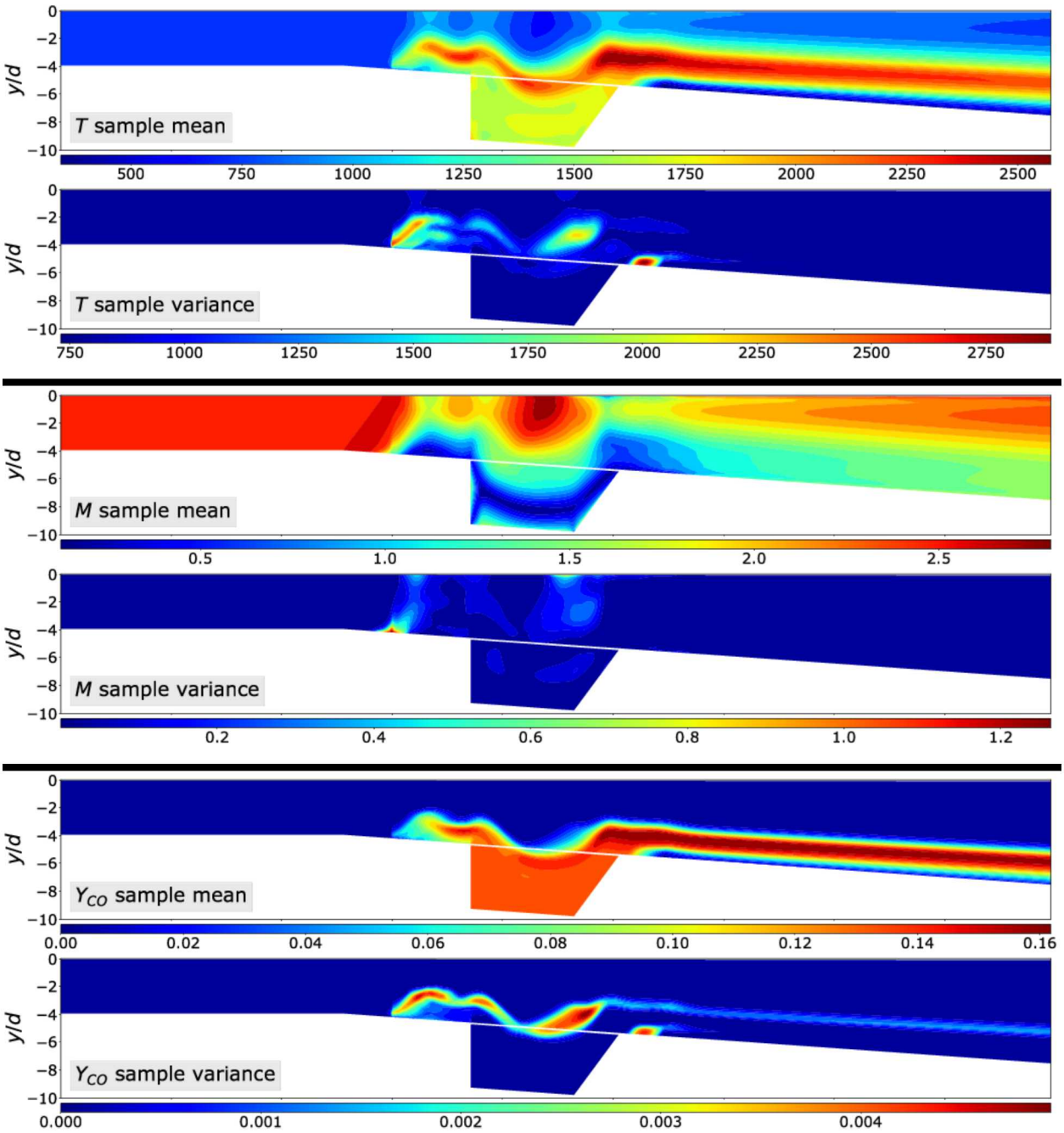


Fig. 3 Sample mean and variance fields for T [K], M, and Y_{CO} from $N = 1053$ coarse grid database samples.

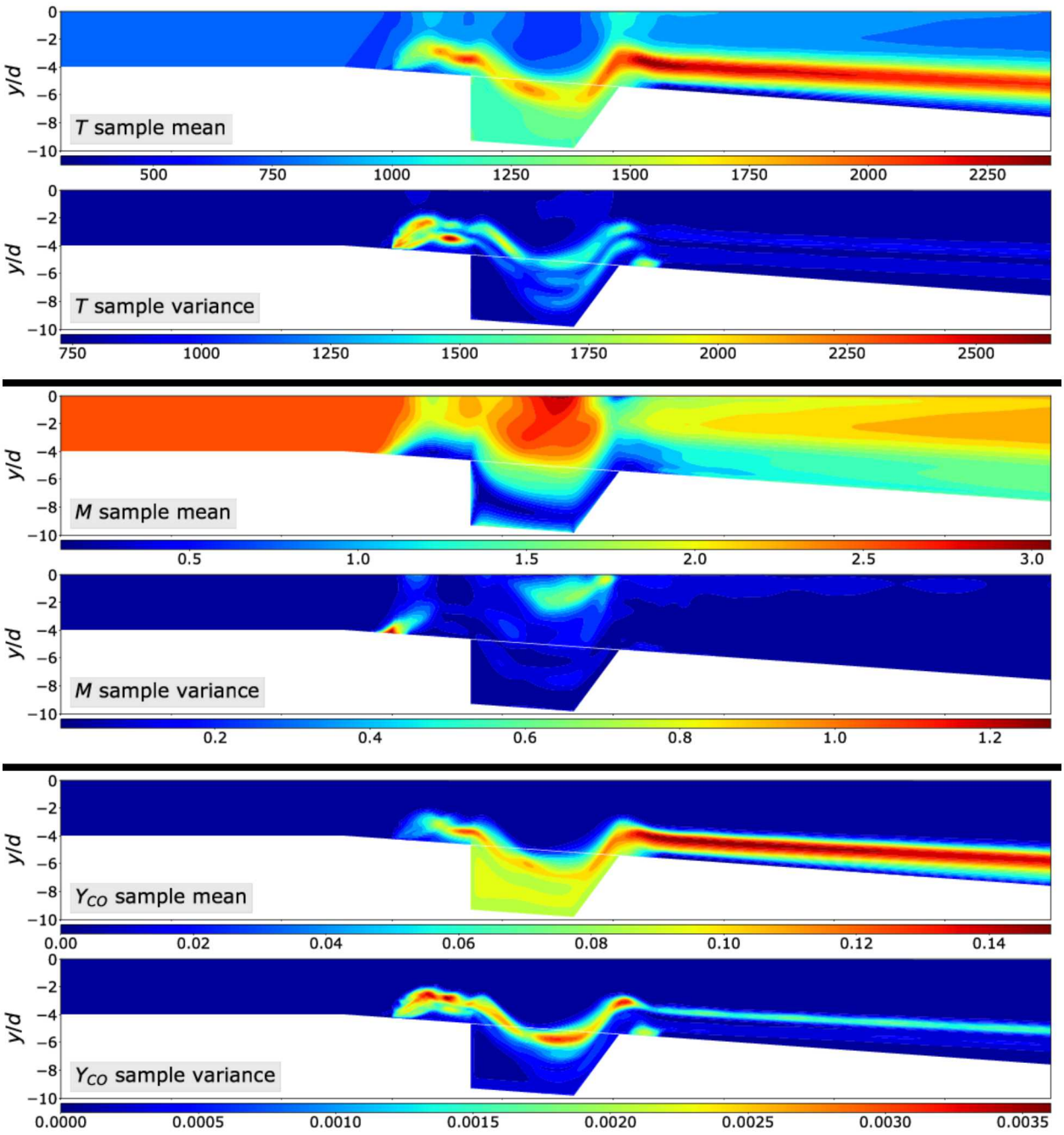


Fig. 4 Sample mean and variance fields for T [K], M, and Y_{CO} from $N = 222$ medium grid database samples.

[3] Mantis, G. C., and Mavris, D. N., "A Bayesian Approach to Non-Deterministic Hypersonic Vehicle Design," *SAE Aerospace Congress and Exhibition*, Technical Paper 2001-01-3033, Seattle, WA, 2001. doi:10.4271/2001-01-3033.

[4] Witteveen, J., Duraisamy, K., and Iaccarino, G., "Uncertainty Quantification and Error Estimation in Scramjet Simulation," *17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA Paper 2011-2283, San Francisco, CA, 2011. doi:10.2514/6.2011-2283.

[5] Constantine, P. G., Emory, M., Larsson, J., and Iaccarino, G., "Exploiting active subspaces to quantify uncertainty in the numerical simulation of the HyShot II scramjet," *Journal of Computational Physics*, Vol. 302, 2015, pp. 1–20. doi: 10.1016/j.jcp.2015.09.001.

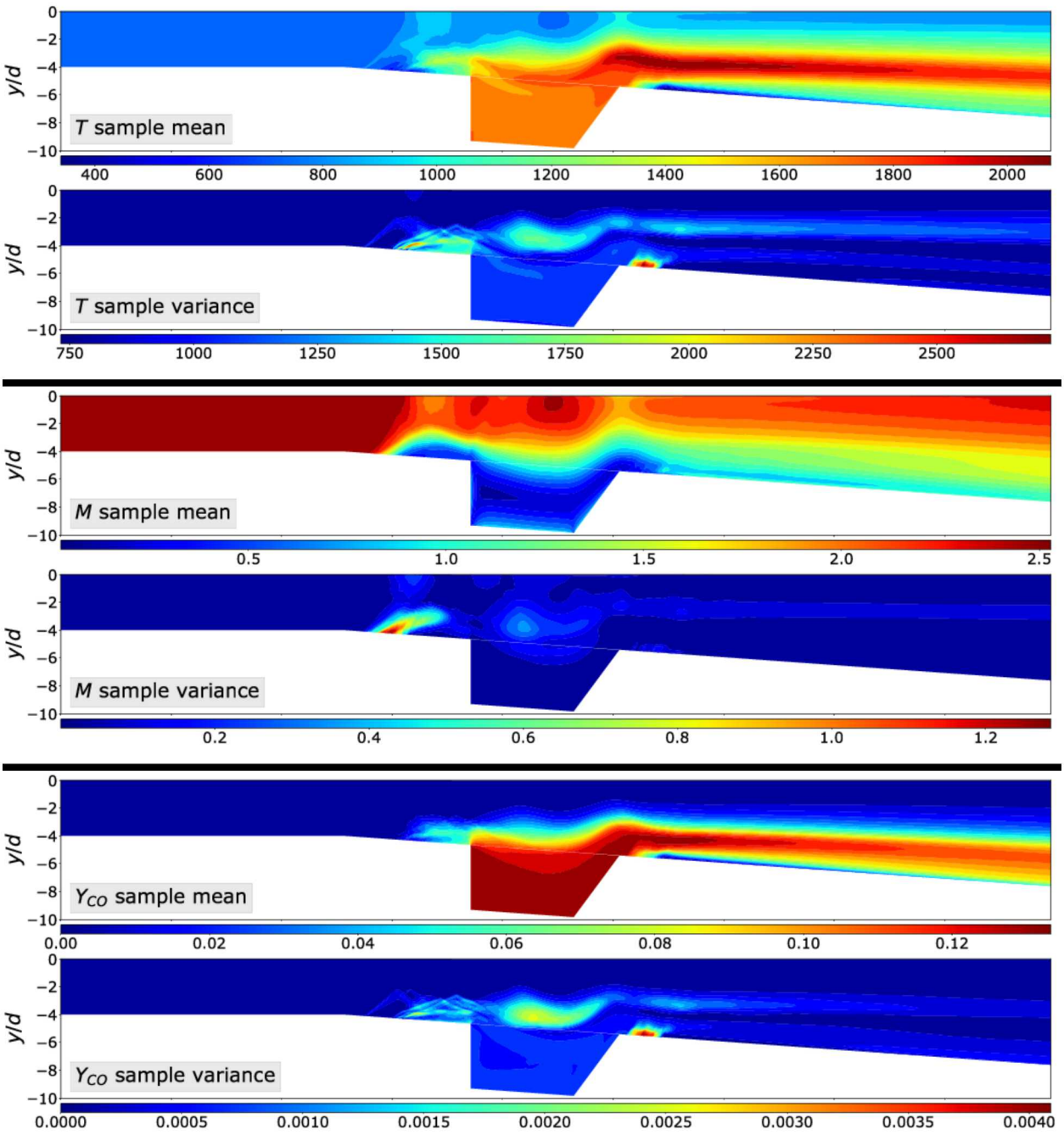


Fig. 5 Sample mean and variance fields for T [K], M, and Y_{CO} from $N = 23$ fine grid database samples.

[6] Geraci, G., Eldred, M. S., and Iaccarino, G., "A multifidelity multilevel Monte Carlo method for uncertainty propagation in aerospace applications," *19th AIAA Non-Deterministic Approaches Conference*, AIAA Paper 2017-1951, Grapevine, TX, 2017. doi:10.2514/6.2017-1951.

[7] Huan, X., Geraci, G., Safta, C., Eldred, M. S., Sargsyan, K., Vane, Z. P., Oefelein, J. C., and Najm, H. N., "Multifidelity Statistical Analysis of Large Eddy Simulations in Scramjet Computations," *20th AIAA Non-Deterministic Approaches Conference*, AIAA Paper 2018-1180, Kissimmee, FL, 2018. doi:10.2514/6.2018-1180.

[8] Huan, X., Safta, C., Sargsyan, K., Geraci, G., Eldred, M. S., Vane, Z. P., Lacaze, G., Oefelein, J. C., and Najm, H. N., "Global

Sensitivity Analysis and Estimation of Model Error, Toward Uncertainty Quantification in Scramjet Computations," *AIAA Journal*, Vol. 56, No. 3, 2018, pp. 1170–1184. doi:10.2514/1.J056278.

[9] Feil, M., and Staudacher, S., "Uncertainty quantification of a generic scramjet engine using a probabilistic collocation and a hybrid approach," *CEAS Aeronautical Journal*, 2018. doi:10.1007/s13272-018-0303-6.

[10] Soize, C., Ghanem, R. G., Safta, C., Huan, X., Vane, Z. P., Oefelein, J. C., Lacaze, G., and Najm, H. N., "Enhancing Model Predictability for a Scramjet Using Probabilistic Learning on Manifolds," *Submitted*, 2018.

[11] Le Maître, O. P., and Knio, O. M., *Spectral Methods for Uncertainty Quantification: with Applications to Computational Fluid Dynamics*, Springer Netherlands, Houten, Netherlands, 2010. doi:10.1007/978-90-481-3520-2.

[12] Sargsyan, K., Debusschere, B., Najm, H., and Maître, O. L., "Spectral Representation and Reduced Order Modeling of the Dynamics of Stochastic Reaction Networks via Adaptive Data Partitioning," *SIAM Journal on Scientific Computing*, Vol. 31, No. 6, 2010, pp. 4395–4421. doi:10.1137/090747932.

[13] Dolvin, D. J., "Hypersonic International Flight Research and Experimentation (HIFiRE)," *15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA Paper 2008-2581, Dayton, OH, 2008. doi:10.2514/6.2008-2581.

[14] Dolvin, D. J., "Hypersonic International Flight Research and Experimentation," *16th AIAA/DLR/DGLR International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA Paper 2009-7228, Bremen, Germany, 2009. doi:10.2514/6.2009-7228.

[15] Jackson, K. R., Gruber, M. R., and Barhorst, T. F., "The HIFiRE Flight 2 Experiment: An Overview and Status Update," *45th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, AIAA Paper 2009-5029, Denver, CO, 2009. doi:10.2514/6.2009-5029.

[16] Jackson, K. R., Gruber, M. R., and Buccellato, S., "HIFiRE Flight 2 Overview and Status Uptate 2011," *17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA Paper 2011-2202, San Francisco, CA, 2011. doi:10.2514/6.2011-2202.

[17] Jackson, K. R., Gruber, M. R., and Buccellato, S., "Mach 6–8+ Hydrocarbon-Fueled Scramjet Flight Experiment: The HIFiRE Flight 2 Project," *Journal of Propulsion and Power*, Vol. 31, No. 1, 2015, pp. 36–53. doi:10.2514/1.B35350.

[18] Hass, N. E., Cabell, K. F., and Storch, A. M., "HIFiRE Direct-Connect Rig (HDCR) Phase I Ground Test Results from the NASA Langley Arc-Heated Scramjet Test Facility," Tech. Rep. LF99-8888, NASA, 2010.

[19] Storch, A. M., Bynum, M., Liu, J., and Gruber, M., "Combustor Operability and Performance Verification for HIFiRE Flight 2," *17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA Paper 2011-2249, San Francisco, CA, 2011. doi:10.2514/6.2011-2249.

[20] Cabell, K. F., Hass, N. E., Storch, A. M., and Gruber, M., "HIFiRE Direct-Connect Rig (HDCR) Phase I Scramjet Test Results from the NASA Langley Arc-Heated Scramjet Test Facility," *17th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, AIAA Paper 2011-2248, San Francisco, CA, 2011. doi:10.2514/6.2011-2248.

[21] Pellett, G. L., Vaden, S. N., and Wilson, L. G., "Opposed Jet Burner Extinction Limits: Simple Mixed Hydrocarbon Scramjet Fuels vs Air," *43rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit*, AIAA Paper 2007-5664, Cincinnati, OH, 2007. doi:10.2514/6.2007-5664.

[22] Lu, T., and Law, C. K., "A directed relation graph method for mechanism reduction," *Proceedings of the Combustion Institute*, Vol. 30, No. 1, 2005, pp. 1333–1341. doi:10.1016/j.proci.2004.08.145.

[23] Zambon, A. C., and Chelliah, H. K., "Explicit reduced reaction models for ignition, flame propagation, and extinction of C2H4/CH4/H2 and air systems," *Combustion and Flame*, Vol. 150, No. 1-2, 2007, pp. 71–91. doi:10.1016/j.combustflame.2007.03.003.

[24] Oefelein, J. C., "Large eddy simulation of turbulent combustion processes in propulsion and power systems," *Progress in Aerospace Sciences*, Vol. 42, No. 1, 2006, pp. 2–37. doi:10.1016/j.paerosci.2006.02.001.

[25] Oefelein, J. C., "Simulation and Analysis of Turbulent Multiphase Combustion Processes at High Pressures," Ph.D. thesis, The Pennsylvania State University, 1997.

[26] Ghanem, R. G., and Spanos, P. D., *Stochastic Finite Elements: A Spectral Approach*, 1st ed., Springer New York, New York, NY, 1991. doi:10.1007/978-1-4612-3094-6.

[27] Najm, H. N., "Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics," *Annual Review of Fluid Mechanics*, Vol. 41, 2009, pp. 35–52. doi:10.1146/annurev.fluid.010908.165248.

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

[29] Xiu, D., and Karniadakis, G. E., "The Wiener-Askey Polynomial Chaos for Stochastic Differential Equations," *SIAM Journal on Scientific Computing*, Vol. 24, No. 2, 2002, pp. 619–644. doi:10.1137/S1064827501387826.

[30] 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. doi:10.1137/S1064827503427741.

[31] Candès, E. J., Romberg, J., and Tao, T., "Robust Uncertainty Principles: Exact Signal Reconstruction From Highly Incomplete Frequency Information," *IEEE Transactions on Information Theory*, Vol. 52, No. 2, 2006, pp. 489–509. doi:10.1109/TIT.2005.862083.

[32] Donoho, D. L., "Compressed sensing," *IEEE Transactions on Information Theory*, Vol. 52, No. 4, 2006, pp. 1289–1306. doi:10.1109/Tit.2006.871582.

[33] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J., "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Foundations and Trends in Machine Learning*, Vol. 3, No. 1, 2010, pp. 1–122. doi:10.1561/2200000016.

[34] Boyd, S., Parikh, N., Chu, E., Peleato, B., and Eckstein, J., "Solve LASSO via ADMM," https://web.stanford.edu/~boyd/papers/admm/lasso/lasso.html, accessed: 2017-07-02.

[35] Huan, X., Safta, C., Sargsyan, K., Vane, Z. P., Lacaze, G., Oefelein, J. C., and Najm, H. N., "Compressive Sensing with Cross-Validation and Stop-Sampling for Sparse Polynomial Chaos Expansions," *SIAM/ASA Journal on Uncertainty Quantification*, in press, 2018.

[36] Jaynes, E. T., "Information theory and statistical mechanics," *Physical Review*, Vol. 106, No. 4, 1957, pp. 620–630. doi:10.1103/PhysRev.106.620.

[37] Jaynes, E. T., "Information Theory and Statistical Mechanics. II," *Physical Review*, Vol. 108, No. 2, 1957, pp. 171–190. doi:10.1103/PhysRev.108.171.