Estimation of inflow uncertainties in laminar hypersonic double-cone experiments

J. Ray*
Sandia National Laboratories, Livermore CA 94550-0969

S. Kieweg†, D. Dinzl‡, B. Carnes§, V. G. Weirs¶, B. Freno‖, M. Howard**, and T. Smith††
Sandia National Laboratories, Albuquerque, NM 87185

I. Nompelis‡‡ and G. V. Candler§§
University of Minnesota, Minneapolis MN 55455, USA

We propose a probabilistic framework for assessing the consistency of an experimental dataset, i.e., whether the stated experimental conditions are consistent with the measurements provided. In case the dataset is inconsistent, our framework allows one to hypothesize and test sources of inconsistencies. This is crucial in model validation efforts. The framework relies on statistical inference to estimate experimental settings deemed untrustworthy, from measurements deemed accurate. The quality of the inferred variables is gauged by its ability to reproduce held-out experimental measurements; if the new predictions are closer to measurements than before, the cause of the discrepancy is deemed to have been found. The framework brings together recent advances in the use of Bayesian inference and statistical emulators in fluid dynamics with similarity measures for random variables to construct the hypothesis testing approach. We test the framework on two double-cone experiments executed in the LENS-XX wind tunnel and one in the LENS-I tunnel; all three have encountered difficulties when used in model validation exercises. However, the cause behind the difficulties with the LENS-I experiment is known, and our inferential framework recovers it. We also detect an inconsistency with one of the LENS-XX experiments, and hypothesize three causes for it. We check two of the hypotheses using our framework, and we find evidence that rejects them. We end by proposing that uncertainty quantification methods be used more widely to understand experiments and characterize facilities, and we cite three different methods to do so, the third of which we present in this paper.

Nomenclature

CRPS_j	=	Continuous ranked probability score for a prediction at probe j
d_S	=	Sorensen distance between a PDF of predictions and observations at a probe
h'_{0_∞}	=	non-dimensional stagnation enthalpy of the flow
P'_{Pitot}	=	non-dimensional stagnation pressure of the flow
$p(x)$	=	dimensional pressure on the double cone
$q(x)$	=	heat flux on the double cone
y_1	=	Non-dimensional surface pressure on the double cone
y_2	=	Non-dimensional surface heat flux on the double cone

*Research Staff, Data Science and Analytics, MS 9152; AIAA member
†Research Staff, V&V, UQ, and Credibility Processes, MS 0828
‡Research Staff, Aerosciences, MS 0825; AIAA Member
§Research Staff, V&V, UQ, and Credibility Processes, MS 0828
¶Research Staff, Computational Science, MS 1320; AIAA Member
‖Research Staff, V&V, UQ, and Credibility Processes, MS 0828
**Research Staff, Aerosciences, MS 0825; AIAA Member
††Research Staff, Computational Science, MS 1320
‡‡Research Staff, Akerman Hall, Room 230; AIAA senior member
§§Professor, Akerman Hall, Room 119; Fellow AIAA

$\mathbf{Y}$ = Prediction of non-dimensional heat flux, pressure and the total enthalpy and total pressure
$\mathbf{Y}^{\text{obs}}$ = Experimentally measured value of $\mathbf{Y}$
Θ = The freestream conditions $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$
σ = Model–data mismatch, a composite of measurement error and model-form error

I. Introduction

Models are typically validated by comparing to experiments. Implicit in this approach is the premise that the experimental data are more trustworthy than the model predictions. This may not always be the case. When simulating extreme environments in wind tunnels, e.g., high-enthalpy flows in large test sections, we often take recourse to novel apparatus, experimental techniques, and instrumentation whose interactions may not be fully understood or which may be required to function at the limits of their operational envelope. While experimental data and their uncertainties are usually specified in a dataset used to validate models, they are difficult to estimate in such extreme environments and when using a novel apparatus. In such a case, if model predictions fail to agree with experimental data, it is difficult to isolate the cause — it could be due to the shortcomings of the model or a mis-specification of the experimental conditions, which form inputs to the model. Ref. [1] describes one such case, when model predictions of laminar hypersonic flow over a $25^{\text{o}}/55^{\text{o}}$ double cone failed to match experimental data from Calspan-University at Buffalo Research Center's (CUBRC's) Large Energy National Shock Tunnel (LENS-I) wind tunnel. Some of the authors of this paper simulated the flow in the entire wind tunnel (the design details were made available to them) and discovered that the flow entering the test section was in vibrational non-equilibrium [1]. This was not known to the experimentalists. When the correct, i.e., non-equilibrium inflow conditions were specified to the model, it had far fewer problems matching the experimental data.

We encountered a similar problem (see Ref. [2]) when validating SPARC (Sandia Parallel Aerodynamics and Reentry Code), a hypersonic flow simulator, using data from CUBRC's LENS-XX wind tunnel. The same $25^{\text{o}}/55^{\text{o}}$ double cone was used in this case. We created an ensemble of inflow conditions, generated from the uncertainties provided by CUBRC for the test section's environment. This ensemble was used to seed SPARC simulations (as the inflow boundary condition) to generate an ensemble of surface-pressure and heat-flux predictions on the double cone. We found that the ensemble was not contained within the uncertainty bounds specified for heat-flux measurements, indicating, superficially, that SPARC could not simulate some of the physics. An obvious approach to check this conclusion would be to simulate the flow in the entire wind tunnel, as done in Ref. [1], and predict the test-section inflow. However, the design details of LENS-XX are unknown to us. In addition, its operation is far more complex than that of LENS-I; see Refs. [3, 4] for a description of how it works. There is no equivalent of a doubly shocked gas at stagnation conditions being expanded via a converging–diverging nozzle. Ref. [5] simulates the operation on an expansion tunnel (the LENS-X, the predecessor of LENS-XX) and finds regions of steady and unsteady flows inside the facility. Thus, computing a test section's inflow conditions, by simulating the entire facility, is not a feasible path forward for validating models using LENS-XX data.

In this paper, we devise an alternative method to reconstruct inflow boundary conditions for SPARC. The method is based on statistical inference and uses experimental quantities that are trusted more, e.g., actual measurements like pressure on the double cone, to estimate quantities that are typically derived conditionally on assumptions, e.g., freestream Mach number. Many of these derived quantities form the input into CFD simulators. The method is agnostic of the experimental facility, i.e., the wind-tunnel design, though it requires some knowledge of the test section to impose far-field and outflow boundary conditions on the CFD model. The method marshals some recent developments in Bayesian inference and statistical emulators *, demonstrated by the authors in climate [6] and turbulence modeling [7], to provide a methodology to estimate uncertain experimental conditions as a joint probability density. Thus, the uncertainty in the estimate, due to sparse data and SPARC's inability to match experimental measurements exactly, is also captured. It is, therefore, a probabilistic framework for assessing the consistency between the stated experimental environment and the measurements gathered in the experiment. This is novel, and it is the first contribution of this paper.

We test our method on one LENS-I case (run35 [1]) and two LENS-XX cases (the low enthalpy Case 1 and high-enthalpy Case 4 [8]), simulating flow over the same $25^{\text{o}}/55^{\text{o}}$ double cone. The reason is two-fold. First, we illustrate how a probabilistic framework can be used to assess fairly complex experiments, i.e., we are not limited to simple canonical flows. Secondly, we show how the probabilistic framework can be used to check hypotheses regarding the cause of the mismatch between SPARC's predictions and experimental data. It is here that similarity measures for uncertain quantities, i.e., random variables characterized by their probability densities, are introduced. The notion of

*A statistical emulator is a fast running proxy or "curve-fit" of a physics model, SPARC in our case

testing competing hypotheses using an inference technique, within the context of hypersonic flow experiments, is a new development and is the second contribution of this work.

Using this framework, we find that, for Case 4, the stated experimental conditions are somewhat inconsistent with the measurements of pressure and heat flux over the double cone. The most probable value of our inferred freestream conditions, computed under an assumption of axisymmetry, lie outside the error bounds prescribed for the stated inflow conditions. We also find that two of the proposed causes of the discrepancy between model predictions and measured data do not explain them. The framework thus rejects the hypotheses that (1) a slight restatement of the inflow conditions and (2) the shortcomings of the sub-models in SPARC can explain the disagreement between measurements from the two experiments and SPARC's predictions, even when we account for measurement errors and inflow boundary condition uncertainties.

More broadly, our other two papers [1, 2] and this one, present three different ways of assessing *experimental* uncertainties. The first way [1] assesses uncertainties in the experimental environment by simulating the entire facility. The second method [2] uses sensitivity analysis and forward propagation of uncertain quantities to decide the rigor with which various physical processes need to be simulated (in models) or controlled (in experiments). Many of these processes may not affect the observables being collected. Finally, this paper closes the loop by providing a consistency check — given experimental observables, it estimates the experimental environment and the estimation uncertainty. If the experimental uncertainty in the environment is computed correctly, the inference should be properly bracketed by it. Note that when performing the inference, one should use a flow model that has been validated using experiments where errors are well-controlled and properly estimated.

The paper is organized as follows. In Sec. II, we review literature on Bayesian inference and statistical emulators, as applied to fluid mechanics. We also describe the literature on double-cone experiments in LENS-I and LENS-XX. In Sec. III, we formulate the estimation problem and describe the various components of the probabilistic framework. In Sec. IV, we test the inference technique on a LENS-I experiment where inflow quantities consistent with the experimental data are known, and then apply it to two LENS-XX experiments. In Sec. V, we discuss our results and assess the hypothesized causes of experiment–model mismatch with our similarity measures. In Sec. VI, we summarize our findings and conclude.

II. Literature review

A. Bayesian inference in fluid mechanics

Bayesian inference [9] is a statistical estimation technique that infers the quantities of interest (QoI) e.g., model parameters, as probability density functions (PDFs). A common situation is when Bayesian inference is used to estimate model parameters from observational data. Bayesian inference is useful when the QoIs cannot be estimated with much certainty due to sparse and/or noisy data or due to model shortcomings that prevent it from reproducing observed/measured data accurately. In such a case, the PDF concisely captures the QoI's estimation error. Unlike traditional (deterministic) estimation techniques [10], Bayesian techniques require one to specify a model for the model–data mismatch.

Unless one assumes a functional form for the PDF, Bayesian inverse problems are solved using a class of methods called Markov chain Monte Carlo (MCMC; Ref [11, 12]). MCMC describes a random walk through the space of QoIs (model parameters); each step generates a new parameter combination called a "proposal". The model (called the "forward problem") is run with the proposal to generate a prediction of the model outputs for which measurements exist. The measurements and predictions are compared, and, based on a set of rules specific to MCMC, are accepted or rejected. Note that poor proposals are *not* always rejected, unlike in a deterministic method. The random walk yields a collection of QoI combinations, which are deemed draws from the PDF that we seek. Simple histograms of the samples are used to visualize the PDF. Apart from the ability to generate PDFs, the key difference between statistical and deterministic inversion techniques is the number of forward-model evaluations required; in the case of Bayesian inference, the difference is generally a factor of a hundred or so more. This makes its use with CFD models very challenging.

Bayesian inference has been used to estimate turbulence model parameters in canonical flow configurations, e.g., channel flows [13] and flat-plates [14]. It is sometimes possible to run a CFD model offline, for a number of QoI/parameter combinations, create a database of the outputs, and construct a statistical map ("curve-fit") between model inputs (QoIs) and outputs for which we have measurements. In such a case, this statistical emulator ("curve-fit") can be used in lieu of the CFD model in a Bayesian-inference setting. Such an approach has been used in a small

handful of recent studies to estimate turbulence model parameters in 3D flow configuration [7, 15]. There exists a single instance of its use in a hypersonic boundary layer transition study [16]. There are no instances of the use of Bayesian inference and statistical measures of similarity to check the consistency of experimental fluid dynamics datasets.

The key to using Bayesian inference with CFD models is the construction of accurate statistical emulators. These "curve-fits" may be low-order, multi-dimensional polynomials, Gaussian processes or splines; Ref. [17] is a review paper on the topic. There are specialized methods for fitting each type of model to simulation data and for checking for over-fitting and accuracy. We use polynomial emulators in this paper; details on how the emulators are fitted can be found in Ref. [6]. A short summary of the exact techniques used to construct emulators and the MCMC algorithm used in this work are provided in Sec. III.

B. Numerical simulations of double-cone experiments

In this paper, we concentrate on checking the consistency of two double-cone experiments conducted in the LENS-XX wind tunnel [8]. LENS-XX is an *expansion* tunnel where temperatures do not exceed about 3000 K [3], implying that the inflow into the test section does not suffer from complications induced by thermochemical non-equilibrium. This considerably eases boundary-condition specifications when simulating these experiments. Consequently, there have been prior attempts to simulate the LENS-XX experiments (there are six of them [18, 19]), including a "blind" code comparison [20]. The six LENS-XX experiments were also used in validation tests for SPARC [2], our compressible flow solver; see the appendix for a summary of SPARC. As described above, an ensemble of simulations (commensurate with the stated uncertainties in the inflow conditions) was performed; they failed to bracket the experimental measurements.

Ref. [18] provides a good description and plots of the shock-wave/laminar boundary layer interaction on the double cone. The 25^{o} fore cone results in an attached shock whereas the 55^{o} aft cone results in a detached bow shock. The two shocks, attached and bow, intersect, creating a triple point and a transmitted shock that impinges at a point on the aft cone. The adverse pressure gradient created by the transmitted shock feeds upstream via the (subsonic) boundary layer, causing the laminar flow to separate upstream of the impingement point. This separated region, in turn, induces a separation shock, which intersects with and modifies the transmitted shock. Plots of these complex shock interactions, and the vortex sheets they induce, can be found in Ref. [18]. CFD simulations in Ref. [18], where both perfect-gas and non-equilibrium flow models are used, show that the fore-cone *pressure* is simulated accurately (i.e., it matches LENS-XX measurements), but separation occurs too late. The simulated *heat flux* under-predicts the measurements significantly, even when measurement errors are taken into account. The authors in Ref. [19] examined whether a more sophisticated model for the coupling between vibrational non-equilibrium and dissociation could explain the mismatch between CFD and experiments. They found that Parks's simple model [21] provided results that were very similar to their more sophisticated one [22]. Furthermore, they found the same under-prediction of heat flux and separation zone size as seen in Ref. [18], and the good predictive skill of CFD for pressure on the fore cone. Ref. [23] studied the importance of catalytic effects on simulations of the six LENS-XX experiments. The authors found that, even in the highest-enthalpy case (Case 4), catalytic effects were minor in the separation zone, although wall catalysis was shown to generally increase the heat flux aft of the reattachment point.

The blind code comparison referred to in Ref. [20] is quite revealing as it shows five different CFD codes failing to agree with the LENS-XX measurements. In addition, they fail to agree with themselves, despite the fact that their constituent models were similar. The simulation results were grid converged, so the discrepancies were attributed to the differences in the grids used by the participants of the blind code comparison exercise. Again, all the models under-predicted the separation-zone size, mostly due to a delay in predicting the separation bubble, and under-predicted the heat flux. It is curious that five separate codes, with five very different grids, would suffer from the same type of error. It is even more curious that they should agree, qualitatively, with the studies in Refs. [18, 19]. One could also hypothesize that the under-prediction could be due to excessive numerical viscosity created by state or slope limiters (which are very active in the region). This issue was also partially explored in Ref. [23], which found that, while numerical issues may affect the size of the separation bubble, the effect is nowhere near large enough to explain the significant under-prediction observed in the CFD simulations. Note that all of these studies ignored the uncertainty in the freestream conditions and double-cone surface measurements, with the exception of Ref. [18], which checked whether measurement uncertainty could explain the disagreement of CFD predictions with experimental data.

These difficulties with simulating LENS-XX data parallel those experienced with double-cone experiments in the LENS-I *shock* tunnel [24] in 2000. Navier–Stokes simulations over-predicted the heat flux on the double cone, though the pressure profile and the separation zone were well predicted [25]. Note that only one experiment, called run35, was simulated. However, it was later found that the extreme conditions in the shock tunnel lead to an inflow into the

test section that was in strong vibrational non-equilibrium [1]. When the inflow conditions were adjusted accordingly, CFD simulations matched the experimental data quite well. This approach was repeated for a number of LENS-I experiments, and, after accounting for vibrational non-equilibrium in the inflow, CFD predictions matched experiments quite well [26]. Again, as with Ref. [19] and LENS-XX, the choice of vibrational non-equilibrium models had little impact on the predictive skill of CFD on LENS-I experiments [26].

The Navier–Stokes simulators used in all these studies are different, use different sub-models and grids, and yet seem to experience the same difficulty. It would seem to imply that the cause of the disagreement between CFD and experiments is common to all of the studies and the experience with LENS-I and Ref. [1] lead us to believe that it could again be the freestream conditions. While the ensemble of simulations in Ref. [2] failed to bracket the experimental measurements, a ±15% perturbation to the inflow conditions of two LENS-XX cases (Case 1 and 4; see Table 1 for the experimental conditions) did bracket the experimental data. This success implies that a modified inflow condition could, perhaps, lead to model predictions close to measurements. This motivated us to develop the inference framework for checking the consistency of experimental measurements against the stated experimental conditions.

III. The estimation problem

In this section, we develop a probabilistic framework for checking the consistency of an experimental dataset. Its main component is a Bayesian-inference step that uses experimental data deemed trustworthy, e.g., actual measurements, to estimate experimental data/conditions deemed less trustworthy, e.g., those that are computed from more fundamental measurements, but, nevertheless, accompany experimental datasets and are inputs into SPARC. Both inferences and experimental measurements are uncertain (they are described by PDFs), and we use similarity metrics to compare them. This framework requires us to carefully demarcate between trustworthy and untrustworthy data. We do so by proposing hypothesized causes of the disagreement between model predictions and experimental data, as described in Sec. II. We will consider three causes:

1) *Cause 1*: The freestream conditions are mis-specified and, therefore, untrustworthy. They will be inferred from measurements on the double cone, the Pitot pressure, and total enthalpy, which are easily measured/computed. The true freestream conditions are expected to be outside the measurement error bounds in Table 1.
2) *Cause 2*: The freestream has spatial variations, i.e., it is not radially uniform, and an assumption of axisymmetry (which we, along with all previous studies, have used) is not warranted. Consequently, the CFD model will continue to disagree with measurements, even when using inferred freestream conditions (as they were obtained under axisymmetry).
3) *Cause 3*: The sub-models included in the CFD model, e.g., chemical mechanisms, transport properties, etc. are deficient. Thus, even when using our inferred freestream conditions, we expect our predictions to degrade in accuracy as we increase the enthalpy of the flow, and thermochemical processes become significant. This cause has been partially investigated before for LENS-XX experiments [19].

In Sec. III.A, we characterize the experimental dataset to determine, qualitatively, what might be inferred from it. Specifically, we perform a global sensitivity analysis to gauge the influence of $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$ on the pressure and heat-flux measurements on the double cone. If any of the freestream variables are found not to impact the measurements substantially, the measurements, in turn, will be of little use in constraining the freestream variables during the inference process. We thus establish expectations and design our inference method commensurate with the information content of the dataset.

In Sec.III.B, we develop the inference scheme. The fundamental idea is as follows. There exists a "forward" process by which the freestream conditions, input into a flow simulator, result in pressure and heat-flux measurements on the double cone, as well as Pitot-pressure and total-enthalpy measurements. The measured quantities are deemed trustworthy. Therefore, starting from the trustworthy measurements, it should be possible to infer freestream conditions that play a dominant part in deciding the value of the measurements. This is performed by posing a Bayesian inverse problem, which will yield a four-dimensional joint PDF for $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$. We will only use data gathered by the sensors on the fore cone, where the flow is attached. The sensor locations on the double cone can be found in the illustrations in Ref. [18]. The flow on the fore cone is free of complex shock interactions (see the figures in Ref. [18]). Not having to capture and resolve the separation point and bubble reduces the otherwise stringent grid-resolution requirements when solving the inverse problem. In addition, the fore-cone flow does not contain appreciable levels of dissociation, even for the highest enthalpy case and, consequently, issues with gas chemistry do not affect the inference process. The inference scheme will use SPARC as its forward process. Our previous experience with inverse problems with three parameters [6, 7] lead us to believe that $O(10^4)$ SPARC invocations might be needed. Consequently, a statistical

Table 1 Freestream conditions, their uncertainties ("Error"), and the references from which they were obtained. run35 was conducted in the LENS-I tunnel, and is described in Ref. [1], i.e., for an inflow freestream in vibrational non-equilibrium. Cases 1 and 4 were conducted in LENS-XX and are described in Ref. [18].

Test	U_∞	ρ_∞	T_∞	T_{v_∞}	Re	M_∞	h_{0_∞}	P_{Pitot}	Gas
	(m/s)	(g/m^3)	(K)	(K)	(m^{-1})		(MJ/kg)	(kPa)	
Error	(±3%)	(±7%)	(±3%)			(±3%)	(±5%)	(±5%)	
run35	2545	0.5848	98.27	2562	14.3×10^4	12.59	3.71	3.55	N_2
Case 1	3246	0.499	175	175	0.14×10^6	12.2	5.44	5.1	Air
Case 4	6479	0.984	652	652	0.20×10^6	12.82	21.77	39.5	Air

emulator for SPARC is needed.

The development of statistical emulators, guarding against over-fitting, as well as means of assessing accuracy, are described in Sec. III.C and Sec. III.D. These emulators are actually used to solve the inference problem in Sec. IV. Because they serve as a proxy for SPARC, the accuracy requirements for the emulator are stringent.

The PDF yielded by the inference process, in the $\left(\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\right)$ space, is called the *posterior PDF*. The *prior belief* regarding $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$ is specified as ±15% error bounds about the conditions in Table 1, which we, adopting a maximum-entropy assumption, interpret as a uniform, four-dimensional PDF. We sample $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$ from the posterior PDF and seed each SPARC simulation with it. This is performed using a fine mesh that resolves the complex shock interactions downstream of the separation point. This process is called a *push-forward posterior* or PFP, and results in an ensemble of predictions (i.e., predictive distribution) at the sensors. In Sec. III.E, we describe the similarity measures used to gauge the quality of the predictions, i.e., their agreement with measurements. We consider two cases: (1) where we ignore the measurement error bounds, and (2) where the error bounds are used in computing the agreement. Similarity measures computed using both the prior and posterior PDFs are compared in Sec. V, where we check the hypothesized causes listed above.

A. Characteristics of the Flow

In Table 1, we list the double-cone experiments that were used in this study, along with the references from which we obtained the experimental conditions. U_∞, ρ_∞, T_∞, and T_{v_∞} are the freestream velocity, density, temperature, and vibrational temperature. Re is the unit Reynolds number, whereas M_∞ is the freestream Mach number. h_{0_∞} is the stagnation enthalpy and P_{Pitot} the Pitot pressure. The uncertainties in these measurements [27], modeled as a uniform distribution, are also provided. We assume that the flow over the double cone is axisymmetric and that the inflow is spatially uniform. No-slip boundary conditions are imposed on the double-cone surface, which is modeled as non-catalytic and is held at a constant temperature of 300K (unless mentioned otherwise). These flows are simulated on a stretched mesh (see Fig. 1), using a 5-species, 17-reaction chemical mechanism for air (containing O_2, O, N_2, N, NO) [18, 28]. Simulations are run to steady state by either reducing the residual by 6 orders of magnitude or ensuring that 100 flow-through times have been simulated.

As a first test, for Case 1, we check whether these uncertainty bounds can bracket observations. We use Dakota [29] to sample 100 freestream conditions independently from the uniform distribution centered around their "nominal" values (as in Table 1) and conduct SPARC simulations on a 256×128 mesh (see Ref. [30] for grid convergence studies). This resolution is sufficient for the attached portion of the flow. In Fig. 2, we plot the pressure (left) and heat flux (right) on the double cone. The dotted lines are the 5$^{\text{th}}$ and 95$^{\text{th}}$ percentile limits of the ensemble of simulations, whereas the solid line is the median prediction. The green line is the experimentally observed separation point. There are three pressure sensors and 17 heat-flux sensors upstream of the green line. Clearly, the pressure measurements ahead of the separation point are bracketed by simulations that honor the experimental uncertainties, but the heat flux is under-predicted, similarly to Ref. [18]. Finding the cause of this discrepancy is the aim of this work. Our hypothesis is that a corrected freestream condition may be the explanation.

Note that the surface measurements on the fore cone admit a self-similar simplification. The following expressions hold:

$$y_1 = \frac{p(x)}{\rho_\infty U_\infty^2} \qquad \text{and} \qquad y_2 = \text{StRe}_x^{1/2} = \frac{q(x)\sqrt{x}}{\rho_\infty U_\infty^3 \sqrt{L}}, \tag{1}$$

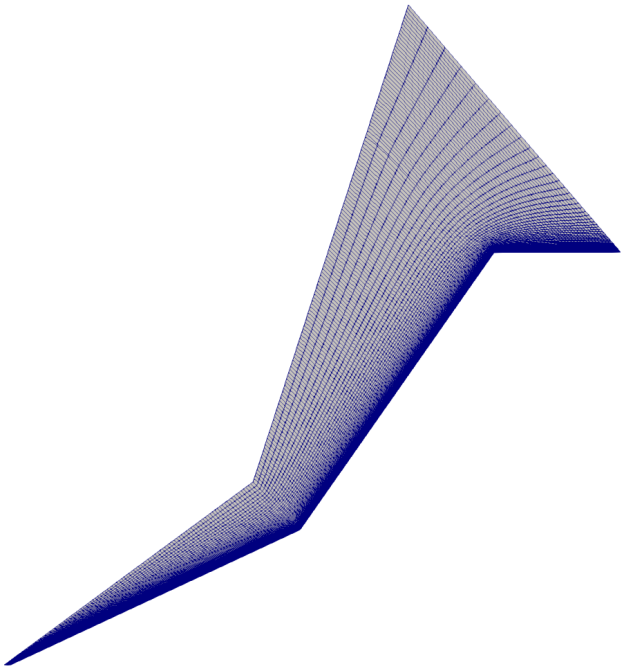

Fig. 1 A 256×128 mesh, somewhat coarsened in the boundary layer region for clarity.

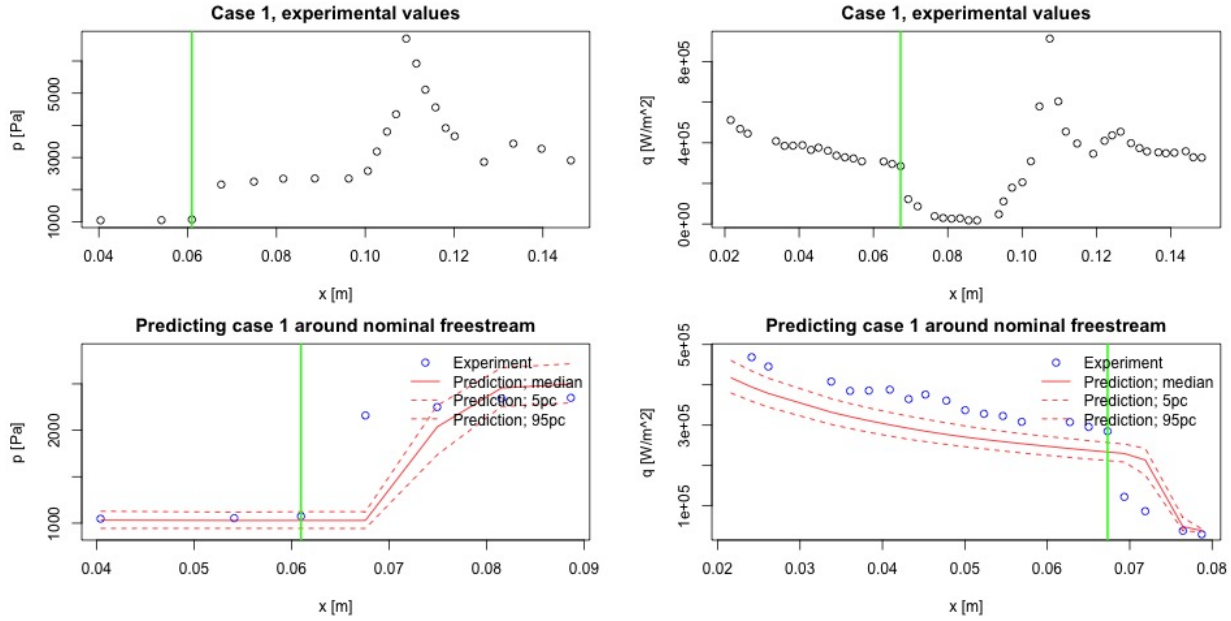

Fig. 2 Top: Experimental measurements of pressure (left) and heat flux (right) for LENS-XX Case 1. The green line denotes the sensor closest to the separation point. Bottom left: 5^{th} and 95^{th} percentiles of the predicted pressure, as well as the median prediction, compared to experimental measurements. Uncertainties arise due to inexact experimental measurements (see Table 1). The experimental measurements of surface pressure are bracketed successfully. The green line is the separation point. Bottom right: The corresponding plot for the heat flux. The modeled heat flux under-predicts the experimental measurements.

where $p(x)$ is the surface pressure, $q(x)$ is the heat flux, x is the streamwise direction, and $L = \sqrt{\nu_\infty/U_\infty}$ = constant is a length scale and a weak function of $(\rho_\infty,\ U_\infty,\ T_\infty,\ T_{v_\infty})$. ν_∞ is the kinematic viscosity of the freestream. In our work, we set $L = 1$ (instead of $\sqrt{\nu_\infty/U_\infty}$); therefore, the values of y_2 are not $O(1)$. Our motivation for the expression for y_2 is derived from flat-plate thermal-boundary-layer scaling $\mathrm{StRe}_x^{1/2} \propto \mathrm{Pr}^{1/2}$ (Chapter 9 in Ref. [31]). The experimental measures, scaled per Eq. 1, are plotted in Fig. 3 (left), and the scalings hold. This collapse of the experimental data in the attached region is quite critical as it implies that, while we have 20 sensors' worth of measurements on the double-cone surface (3 pressure and 17 heat fluxes), the information content of the measurements is quite meager — they can be reduced to two numbers, one for surface pressure and one for heat flux.

Secondly, we check the relative importance of the four freestream conditions $\Theta = \{U_\infty,\ \rho_\infty,\ T_\infty,\ T_{v_\infty}\}$ in impacting the surface quantities $\{p(x),\ q(x)\}$ predicted by SPARC, i.e., we conduct a global sensitivity analysis (GSA), conditioned on the uncertainties in Θ tabulated in Table 1. Specifically, we compute the *total effects* sensitivity indices (Sobol' indices [32]) of $p(x)$ and $q(x)$ at the 3 and 17 sensors (respectively) in the attached portion of the flow. Sobol' indices are a measure of the sensitivity of $p(x)$ and $q(x)$ at the sensors on the fore cone when Θ is varied over a *range* of values. The *total effects* Sobol' index quantifies the influence of a freestream quantity as it interacts with the other quantities, as they too are varied over a range. In contrast, gradients provide *local* sensitivity at a specified Θ and interactions are ignored. In our case, the GSA is performed over the prior belief of Θ. The Sobol' indices are computed as follows. 209 SPARC simulations (about 150 CPU-hours per run) are performed on a 3-level hierarchical mesh in the Θ-space. Each simulation generates $\{p(x),\ q(x)\}$ predictions at the fore-cone sensors, as well as $\{h_{0_\infty},\ P_{\mathrm{Pitot}}\}$. The 209 simulations are stored and used to fit polynomial chaos expansions (PCE; orthogonal functions of Θ used to obtain an analytical functional map between SPARC predictions and inputs; see Ref. [32]). The coefficients of the fits yield Sobol' indices analytically. Four Sobol' indices are computed for each sensor. They represent the sensitivity of SPARC predictions at that sensor to perturbations in ρ_∞, U_∞, T_∞, and T_{v_∞}, "averaged" over the prior distribution. Sobol' indices are also computed for h_{0_∞} and P_{Pitot} (not shown in this paper). In Fig. 3 (right), we plot the *total* Sobol' indices for SPARC predictions of $p(x)$ and $q(x)$ at 7 pressure and 21 heat-flux-sensor locations. The 3 $p(x)$ and 17 $q(x)$ sensors that will be used in the inverse problem are also marked. We see that the sensitivities are dominated by ρ_∞ and U_∞, with T_{v_∞} providing a small contribution to $q(x)$.

The self-similar collapse of the surface measurements and the sensitivities have the following implications:

1) The information content of the double-cone surface measurements upstream of separation is low — while we may have multiple sensors' worth of measurements on the fore cone, they reduce to two numbers, one for pressure and one for a (scaled) heat flux. The inference scheme will thus use four observations $\{y_1,\ y_2,\ h_{0_\infty},\ P_{\mathrm{Pitot}}\}$ to infer four unknowns $\{U_\infty,\ \rho_\infty,\ T_\infty,\ T_{v_\infty}\}$, which, in a regression setting, will not allow them to be estimated with much accuracy. This is the argument for adopting a Bayesian technique, which will allow us to infer the freestream conditions as a PDF and thus capture the estimation uncertainty.
2) $p(x)$ and $q(x)$ are not affected much by T_∞ and T_{v_∞} so the fore-cone surface measurements will provide weak constraints on T_∞ and T_{v_∞}. Consequently, it may be difficult to infer the freestream temperatures. However, by the same token, the uncertainties of $\{T_\infty,\ T_{v_\infty}\}$ do not contribute much to the discrepancies seen in Fig. 2.
3) When constructing the emulator for $\{y_1,\ y_2\} = f(\rho_\infty,\ U_\infty,\ T_\infty,\ T_{v_\infty})$, a linear dependence on $\{T_\infty,\ T_{v_\infty}\}$ might be sufficient.

B. Formulation of the Bayesian Inverse Problem

In this section, we develop the inverse problem for estimating Θ as a joint, 4-dimensional PDF. First, we develop a simple formulation for the mismatch / error between observations and SPARC predictions, conditional on a specified Θ. In deterministic inversion, this mismatch is generally minimized by optimizing Θ. In contrast, in a statistical setting, we use a statistical model of the error to convert the mismatch into a statement on the likelihood of obtaining the observed data, conditional on the specified Θ. Next, we modulate the expression for the likelihood with our prior beliefs, so that we may impose prior bounds on Θ. Finally, using Bayes' rule, we obtain an expression for the PDF of Θ, conditional on measurements.

Let $\mathbf{Y} = \{\overline{y_1},\ \overline{y_2},\ h'_{0_\infty},\ P'_{\mathrm{Pitot}}\} = \mathcal{M}(\Theta)$ be the SPARC predictions of surface quantities $\{p(x),\ q(x)\}$ and the freestream stagnation enthalpy and pressure. $\{\overline{y_1},\ \overline{y_2}\}$ are averaged over the 3 and 17 sensors where $p(x)$ and $q(x)$ are measured, as y_1 and y_2 are deemed constant over them. Here $h'_{0_\infty} = h_{0_\infty}/U_\infty^2$ and $P'_{\mathrm{Pitot}} = P_{\mathrm{Pitot}}/\rho_\infty U_\infty^2$ are the non-dimensional stagnation enthalpy and pressure of the freestream flow. Let $\mathbf{Y}^{\mathrm{obs}}$ be the experimental counterpart of $\mathbf{Y}$. Assume that the two are related as $\mathbf{Y}^{\mathrm{obs}} = \mathbf{Y} + \boldsymbol{\epsilon}$, where $\boldsymbol{\epsilon} = \{\epsilon_i\}, i = 1 \ldots 4$ and $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$. Here $\mathcal{N}(0, \sigma^2)$ is a normal distribution with zero mean and an unknown standard deviation σ. $\boldsymbol{\epsilon}$ provides a crude measure of the model–data

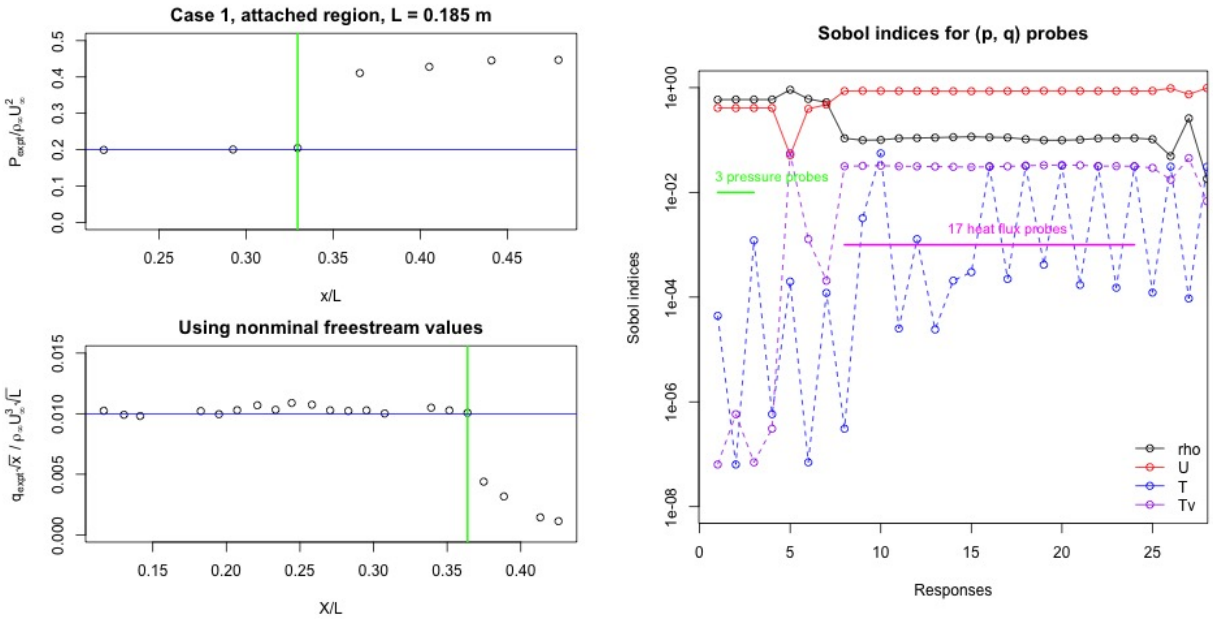

Fig. 3 Left: The experimental measurements for Case 1 scaled per Eq, 1. The green line is the point of flow separation. Right: Sobol indices for SPARC predictions at 7 $p(x)$ and 21 $q(x)$ sensor locations, due to Θ. The 3 $p(x)$ and 17 $q(x)$ probes used in the inverse problem are marked.

mismatch and is a composite of the model-form error (in this case, primarily our assumption of axisymmetry) and experimental-measurement error.

Let $P(\Theta, \sigma^2|\mathbf{Y}^{\text{obs}})$ be the five-dimensional posterior PDF over the freestream conditions and the model–data mismatch, conditional on the experimental data $\mathbf{Y}^{\text{obs}}$. Let $\Pi_1(\Theta)$ and $\Pi_2(\sigma^2)$ be our prior beliefs regarding the distribution of the freestream conditions and σ^2. Then, the likelihood of observing $\mathbf{Y}^{\text{obs}}$, given a freestream realization Θ, is:

$$\mathcal{L}(\mathbf{Y}^{\text{obs}}|\Theta, \sigma^2) \propto \frac{1}{\sigma^4} \exp\left(-\frac{||\mathbf{Y}^{\text{obs}} - \mathcal{M}(\Theta)||_2^2}{2\sigma^2}\right). \tag{2}$$

By Bayes' theorem, the calibrated value or posterior density of (Θ, σ^2) is

$$P(\Theta, \sigma^2|\mathbf{Y}^{\text{obs}}) \propto \mathcal{L}(\mathbf{Y}^{\text{obs}}|\Theta, \sigma^2)\Pi_1(\Theta)\Pi_2(\sigma^2) \propto \frac{1}{\sigma^4} \exp\left(-\frac{||\mathbf{Y}^{\text{obs}} - \mathcal{M}(\Theta)||_2^2}{2\sigma^2}\right) \Pi_1(\Theta)\Pi_2(\sigma^2). \tag{3}$$

The prior density for σ^2 is defined in terms of its reciprocal, i.e., $\Pi_2(\sigma^{-2})$ and is modeled using a Gamma prior, i.e., $\sigma^{-2} \sim \Gamma(k, \theta)$, where $k = 1, \theta = 1$. The prior density $\Pi_1(\Theta)$ is modeled as a 4-dimensional uniform distribution in $(U_\infty, \rho_\infty, T_\infty, T_{v_\infty})$-space, with lower and upper bounds described by ±15% perturbation around the "nominal" experimental values provided in Table 1.

The posterior density $P(\Theta, \sigma^2|\mathbf{Y}^{\text{obs}})$ is complicated, and can be "solved for" or constructed only via sampling. We use a Markov chain Monte Carlo (MCMC) method [11], specifically Delayed Rejection Adaptive Metropolis (DRAM, [33]), to draw the samples from $P(\Theta, \sigma^2|\mathbf{Y}^{\text{obs}})$ in Eq. 3. We solve this problem using `R` [34] with the MCMC package `FME` [35]. The samples are used to construct an approximation of $P(\Theta, \sigma^2|\mathbf{Y}^{\text{obs}})$ with kernel density estimation. The inverse Gamma prior for σ^2 is a conjugate prior, which simplifies sampling of σ^{-2} via a Gibbs sampler [11]. The Raftery–Lewis method [36], as implemented in the `R` package `mcgibbsit` [37] is used to assess convergence of the MCMC chain. The MCMC chain takes $O(10^4)$ samples to converge, each of which requires a SPARC run. Since that is unfeasible (being a Markov chain, the $O(10^4)$ samples are taken serially), we will replace SPARC with its statistical emulator, as described below.

C. Statistical Emulators of SPARC

A statistical emulator $\mathbf{Y} \approx \mathcal{M}_e(\Theta)$ is a fast-running "curve-fit" of SPARC that provides an approximate prediction of $\{p(x),\ q(x),\ h_{0_\infty},\ P_{\text{Pitot}}\}$ for different freestream settings Θ. We construct a 4-dimensional, 3-level Smolyak sparse grid in $\Theta = (\rho_\infty,\ U_\infty,\ T_\infty,\ T_{v_\infty})$-space using a tensor-product of Gauss–Patterson quadrature points. A uniform distribution centered at the conditions in Table 1 with ±15% bounds serves as our distribution for Θ. This results in 209 realizations of Θ, where SPARC is run to obtain $\{p(x),\ q(x),\ h_{0_\infty},\ P_{\text{Pitot}}\}$ predictions. We then model the predictions as:

$$\begin{aligned} \overline{p(x)} &= \rho_\infty U_\infty^2 f_1(\rho_\infty, U_\infty, T_\infty, T_{v_\infty}), & \overline{q(x)\sqrt{x}} &= \rho_\infty U_\infty^3 f_2(\rho_\infty, U_\infty, T_\infty, T_{v_\infty}), \\ h_{0_\infty} &= U_\infty^2 f_3(\rho_\infty, U_\infty, T_\infty, T_{v_\infty}), & P_{\text{Pitot}} &= \rho_\infty U_\infty^2 f_4(\rho_\infty, U_\infty, T_\infty, T_{v_\infty}). \end{aligned} \tag{4}$$

Here, $f_i(\rho_\infty, U_\infty, T_\infty, T_{v_\infty})$ are linear functions. These polynomial models are fitted to the data from the ensemble of SPARC runs via least-squares regression. In Fig. 4, we plot some of the predictions of y_1 and y_2 over all sensors on the fore cone for a few chosen Θ values. We see that these predictions are approximately constant over the fore cone and variation of Θ merely results in the change in the value of the constant, i.e., y_1 and y_2 never vary over space. Consequently, we model their averages $\overline{y_1}$ and $\overline{y_2}$. The polynomial models fitted to the SPARC ensemble are then simplified by stepwise elimination of terms and computing the Akaike Information Criterion (AIC). (AIC is used to select between models when they are fitted to the same data. AIC is a sum of a goodness-of-fit term and a model-complexity-penalty term. The goodness-of-fit term assumes a value of zero for a perfect fit. The penalty term is double the number of free parameters in a model. When a highly parameterized model is (over-)fitted to data, the goodness-of-fit term assumes a very small value, though one pays a price for the model-complexity-penalty. Model fitting and selection strives to minimize AIC, in the process achieving a balance between goodness of fit and model simplicity.) We see that the linear terms involving ρ_∞ and U_∞ in $f_i(\rho_\infty, U_\infty, T_\infty, T_{v_\infty})$ in Eq. 4 mostly do not survive this simplification procedure and f_i usually reduce to weak functions of $\{T_\infty,\ T_{v_\infty}\}$.

The form of the emulator obtained after the AIC simplification should, in principle, be immune to overfitting. As described in Ref. [7], it often is not. Consequently, we further check it using repeated random sub-sampling cross-validation. We divide the training data of 209 SPARC simulations randomly in a 75:25 proportion into a learning and testing set. We "learn" the empirical models in Eq. 4 using the learning set, and compute its prediction errors with the testing set. The process is repeated 20 times, with different learning/testing sets, and an average prediction error is computed for both the sets. We stipulate that the learning and testing set errors be similar (as a sign that the complexity of the emulators are commensurate with the training data) and the errors be less than 5%, so that they may be considered good approximations of SPARC. We find that these stipulations are met for all four predicted quantities in Eq. 4.

D. Statistical Emulators for Freestream Conditions

In Sec. III.C, we stated that we construct emulators for computing the freestream total enthalpy h_{0_∞} and the Pitot pressure P_{Pitot}. These are computed by simulating flow over a cylinder with a hemispherical end (an approximation of a Pitot tube, henceforth referred to as the *mock Pitot tube*) using SPARC. The stagnation pressure at the tip of the hemispherical nose of the cylinder was taken as the P_{Pitot} and the flow conditions at the stagnation point were used to compute h_{0_∞}. Since these conditions are available only at the grid point closest to the stagnation point and are, in any case, numerical approximations, we check their accuracy first.

The total enthalpy for run35, 3.71 MJ/kg, is sufficiently small, such that h_{0_∞} and P_{Pitot} may be calculated using Rayleigh's formula [1]. We compute the ratio of the numerical and analytical values for 100 Θ samples chosen from the prior distribution and find that the numerical value of P_{Pitot} is slightly underestimated. The mean and standard deviation of the ratio are $(0.9939, 7.982 \times 10^{-4})$. Thus the under-prediction is about 0.5%, implying that we may use predictions from the mock Pitot tube as a relatively accurate proxy for the true h_{0_∞} and P_{Pitot}.

Next we proceed to constructing the functions $f_3(\rho_\infty, U_\infty, T_\infty, T_{v_\infty})$ and $f_4(\rho_\infty, U_\infty, T_\infty, T_{v_\infty})$ in Eq. 4. 512 samples of Θ were obtained from the prior distribution via Latin hypercube sampling (LHS; [32]) corresponding to Case 1 (see Table 1) and flow over the mock Pitot tube simulated using SPARC. 512 corresponding realizations of h_{0_∞}, P_{Pitot} were recorded to construct a training dataset $\{\Theta,\ h_{0_\infty},\ P_{\text{Pitot}}\}$. The surrogate model, which predicts $\{h^*_{0_\infty}, P^*_{\text{Pitot}}\}$ for any arbitrary Θ^* sampled from the prior distribution, is constructed using local regression, as implemented in `R` by the function `loess`; the principles are described in [38]. Fundamentally, the space of Θ, as realized by the 512 samples, is divided into 4-dimensional rectangles with approximately the same number of samples. These rectangles are stored in a 4-dimensional tree. When a prediction at Θ^* is desired, the hyper-rectangle containing Θ^* is identified, and a sphere, centered at Θ^* and covering a fraction s of the Θ-space, is constructed. The samples that lie inside this sphere (the neighborhood of Θ^*) are gathered, a process that requires rapid access of the neighboring hyper-rectangles (and hence

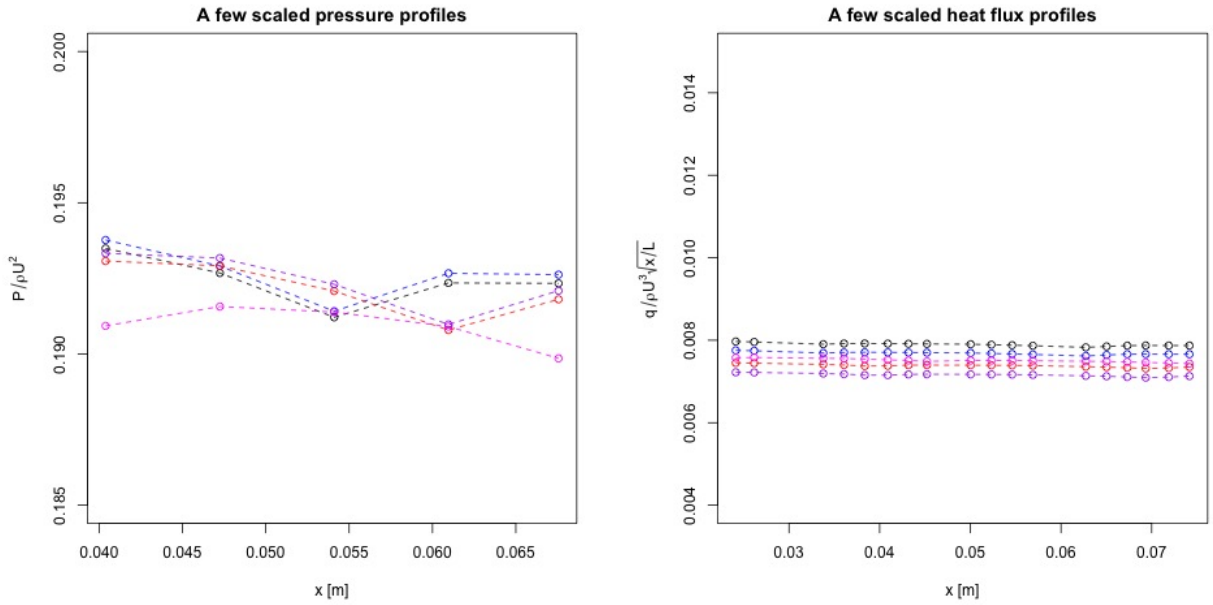

Fig. 4 Predictions of $p(x)/\rho_\infty U_\infty^2$ and $q(x)\sqrt{x}/\rho_\infty U_\infty^3\sqrt{L}$ for a few randomly chosen Θ, from an SPARC ensemble simulation corresponding to Case 4 in Table 1. Left: pressure distribution on the fore cone. Right: Distribution of scaled heat flux on the fore cone.

the need to store them in a tree). A quadratic surface in $(\rho_\infty, U_\infty, T_\infty, T_{v_\infty}, h_{0_\infty})$-space, is fitted to SPARC-predicted h_{0_∞} data and $h^*_{0_\infty}$ interpolated out. An identical procedure is repeated for P_{Pitot}. In our model, we set $s = 0.75$ i.e., 75% of the samples are used to perform each interpolation. During interpolation, the samples are weighted unequally depending on their distance from Θ^*. The weight for a sample Θ is given by $w(\Theta) = (1 - d^3_{\text{norm}})^3$, where d_{norm}, the normalized Euclidian distance between Θ and Θ^* is given by

$$d_{\text{norm}}(\Theta, \Theta) = \frac{d(\Theta, \Theta^*)}{d_{\max}}.$$

Computations are performed after normalizing the variables in the training data to $[-1, 1]$. $d(\Theta, \Theta^*)$ is the Euclidean distance in the normalized Θ-space and $d_{\max}$ is the maximum value realized within the neighboring samples.

The particular choices, i.e., $s = 0.75$ and quadratic surfaces, could result in overfitting depending upon the smoothness of $f_3(:)$ and $f_4(:)$. We check their suitability via repeated random subsampling validation. We divide our 512 records in the training data into a Learning Set (LS) and a Training Set (TS), in a 75:25 ratio. The local regression model (henceforth `loess`) is trained on the LS, used to predict the h_{0_∞} values there and to compute a root-mean-square relative error $\epsilon_{rmse,LS}$. The same `loess` model is used to predict at the Θ values in the TS and compute the corresponding $\epsilon_{rmse,TS}$. This is repeated 40 times, and the mean errors $\overline{\epsilon}_{LS}$ and $\overline{\epsilon}_{TS}$, over 40 rounds of cross-validation, computed. We require that both $\overline{\epsilon}_{LS}$ and $\overline{\epsilon}_{TS}$ be small (so that the interpolations are accurate) and that they are both similar in magnitude (a requirement that is violated if the model overfits the data, in which case $\overline{\epsilon}_{TS} \gg \overline{\epsilon}_{LS}$).

In Table 2, we tabulate $\overline{\epsilon}_{LS}$ and $\overline{\epsilon}_{TS}$ for the LENS-XX cases in Table 1, for both h_{0_∞} and P_{Pitot}. We see that the cross-validation errors $\overline{\epsilon}_{LS}$ and $\overline{\epsilon}_{TS}$ are small and of similar magnitude regardless of the variable being predicted. In the next section, we will use these emulators to estimate the freestream quantities $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$ from experimental measurements.

E. Gauging quality

As part of the evaluation of freestream conditions Θ, we will have to compare its predictions $\mathbf{z}$, as generated by SPARC, with observations $\mathbf{z}_{\text{obs}}$. Here $\mathbf{z} = \{z_j\}, j = 1 \ldots J$, is a vector of predictions, e.g., pressures at a set of J probes. If we are provided with a single Θ, we can compute the root-mean-square error (RMSE) between $\mathbf{z}$ and $\mathbf{z}_{\text{obs}}$ to obtain a simple figure of merit for the quality of Θ. However, under an uncertain characterization of freestream conditions, we

Table 2 Cross-validation errors for h_{0_∞} and P_{Pitot}. Both $\overline{\epsilon}_{LS}$ and $\overline{\epsilon}_{TS}$ are small and similar in magnitude.

	P_{Pitot}		h_{0_∞}	
Experiment	$\overline{\epsilon}_{LS}$	$\overline{\epsilon}_{TS}$	$\overline{\epsilon}_{LS}$	$\overline{\epsilon}_{TS}$
Case 1	4.6×10^{-4}	5.5×10^{-4}	9.89×10^{-8}	1.12×10^{-7}
Case 4	4.25×10^{-4}	5.1×10^{-4}	1.67×10^{-7}	1.88×10^{-8}

have a probability density function (PDF) $P(\Theta)$, which does not lend itself to a single figure of merit using RMSE. More accurately, our observations $\mathbf{z}_{\text{obs}}$ are also uncertain (measurement errors are specified using percentages) and we model the uncertain observations as a uniform distribution $Q(\mathbf{z}_{\text{obs}})$, with the bounds computed using the error percentages. For convenience, we also require a single figure of merit to compare the quality of $P(\Theta)$ given $Q(\mathbf{z}_{\text{obs}})$.

We will use continuous ranked probability score (CRPS, [39, 40]) as the figure of merit for $P(\Theta)$ given $\mathbf{z}_{\text{obs}}$ (note, *not* $Q(\mathbf{z}_{\text{obs}})$). The CRPS is computed for each probe, and a single figure of merit is obtained by taking the average over all J probes. We take a set of Θ samples, $\Theta^{(m)}, m = 1 \ldots M$, and compute predictions $\mathbf{z}^{(m)}$. Thus, at every probe j, we have the predictions $z_j^{(m)}$, which allow us to construct a predicted distribution $Q'(z_j)$ to be compared with $z_{\text{obs},j}$. Let $Q'_{c,j}(z)$ be the cumulative distribution function (CDF) corresponding to the probability density $Q'(z_j)$. The CDF of the perfect observation $z_{\text{obs},j}$ is the Heaviside function $\mathbb{H}(z_{\text{obs},j})$. The CRPS of the prediction at the j^{th} probe, CRPS_j, is defined as

$$\text{CRPS}_j = \int_\infty^\infty \left(Q'_{c,j}(y) - \mathbb{H}(y - z_{\text{obs},j}) \right)^2 dy. \tag{5}$$

Thus, CRPS_j has units of the observed quantity $z_{\text{obs},j}$. Small values of CRPS_j denote a good match between predictions and observations. We will denote the mean of CRPS_j as CRPS.

CRPS, as defined above, will be used to gauge the quality of $P(\Theta)$ when we are not quite sure of our uniform distribution model for $Q(\mathbf{z}_{\text{obs}})$. In cases where we decide to proceed with the uniform distribution model for measurement errors, we will gauge the quality of $P(\Theta)$ using the Sorensen distance d_S [41]. Let $Q(y)$ denote the PDF of the observed quantity $z_{\text{obs},j}$. Let $Q'(y)$ be the PDF of the predicted quantity z_j at probe j. Let $Q_k(y)$ and $Q'(y)$ be their discretized form on a 1-D discretization of the $y-$space. Then d_S between $Q(y)$ and $Q'(y)$ is defined as

$$d_S = \frac{\sum_k |Q_k - Q'_k|}{\sum_k (Q_k + Q'_k)}. \tag{6}$$

$d_S = 0$ indicates a perfect match. $d_S = 1$ corresponds to $Q(y)$ and $Q'(y)$ being disjoint.

IV. Results

In this section, we check the usefulness of Bayesian estimation/calibration of freestream conditions and exercise the metrics that can be used to judge the quality of a calibration. This is performed with run35, for which we have a good estimate of the freestream conditions from [1]. Thereafter, we proceed with estimation of freestream conditions for Case 1 and Case 4.

A. Verifying the Bayesian calibration approach

In this section, we investigate whether $\mathbf{Y}^{\text{obs}}$ can be used to infer freestream conditions which, though different from those listed in Table 1, are nevertheless plausible. First, however, we test the ability of the inversion procedure described in Sec. III.B to infer freestream conditions. We do this with run35. run35 is being used because (1) its freestream conditions, taken from Ref. [1], can reproduce experimental data quite well and thus allows us to check our inverse-problem formulation and (2) the test is conducted with N_2, which, along with its low enthalpy, ensures that complications from dissociated and reacting flows do not arise.

run35, as listed in Table 1, is a low-enthalpy case tested in the LENS-I tunnel. The freestream conditions in Table 1 are at variance with the original paper, where it is described [24]. The reason for this discrepancy is as follows. The freestream conditions for run35 (Case D for the $25^{\text{o}}/55^{\text{o}}$ cone in Ref. [24]) were computed from wind-tunnel measurements, assuming that the flow was in equilibrium, i.e., $T_\infty = T_{v_\infty}$. The authors of Ref. [1] simulated the flow in the LENS-I nozzle, upstream of the test section, and found that the flow entering the test section was in non-equilibrium

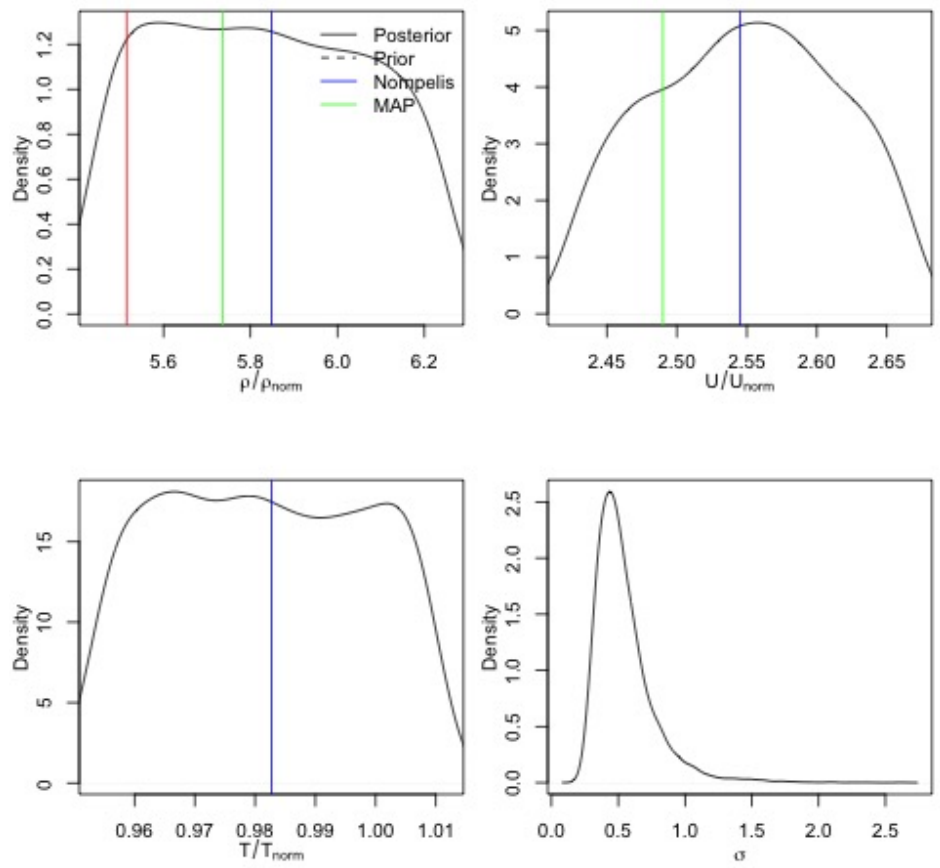

Fig. 5 Marginalized PDFs of ρ_∞, U_∞, T_∞, and σ, inferred from run35 data (Table 1). The variables have been normalized to ease interpretation. $\{\rho_{\text{norm}}, u_{\text{norm}}, T_{\text{norm}}\} = \{0.1\,\text{g/m}^3, 1000\,\text{m/s}, 100\,\text{K}\}$. The green line indicates the MAP value, as inferred. The blue line is the "nominal" freestream value computed in Ref. [1]. The red line is the freestream value from the original description of the experiments in Ref. [24]. They are missing in most of the plots as they fall outside the range of the horizontal axis. The PDF for T_∞ lacks a MAP value as it is practically flat.

($T_\infty \neq T_{v_\infty}$) and "frozen," i.e., the vibrational energy was partitioned out and did not play much of a role in the interaction with the double cone. Consequently, T_{v_∞} was dropped from the inference problem described in Sec. III.B and Sec. III.C, and the flow was computed under a perfect-gas assumption, e.g., one could use Rayleigh's formula to compute the stagnation pressure (see Ref. [1]). In addition, the double cone was maintained at a temperature of 296.11 K, instead of 300 K for Cases 1 and 4 (Table 1).

In Fig. 5, we plot the one-dimensional projections (called "marginalizations") of the four-dimensional PDF in $\{\rho_\infty, U_\infty, T_\infty, \sigma^2\}$ obtained from an MCMC solution of Eq. 3. The green line is the MAP (maximum *a posteriori*) value of the inferred quantities, whereas the blue line is the value computed in Ref. [1]. The red line is the value as reported in the original paper describing the experiment [24]. We see that the red line is missing from the plots for U_∞ and T_∞ because it lies outside the range of the horizontal axis. We see that the PDFs for ρ_∞ and T_∞ are wide, indicating that they cannot be estimated very well, i.e., there is a lot of uncertainty in the inferred quantity. This is far from being the case for U_∞, which has a sharp PDF. Also, the MAP values, as inferred, are very close to the non-equilibrium freestream conditions (within 5%) computed in Ref. [1]. It would seem that the freestream conditions that the Bayesian method (see Sec. III.B) infers from the data are quite close to those computed in Ref. [1] by simulating the flow in the wind tunnel.

Next we check the quality of the estimated solution/PDF. We take 100 random samples from the three-dimensional PDF over $\{\rho_\infty, U_\infty, T_\infty\}$ and simulate flow over the double cone for these realizations of the freestream conditions

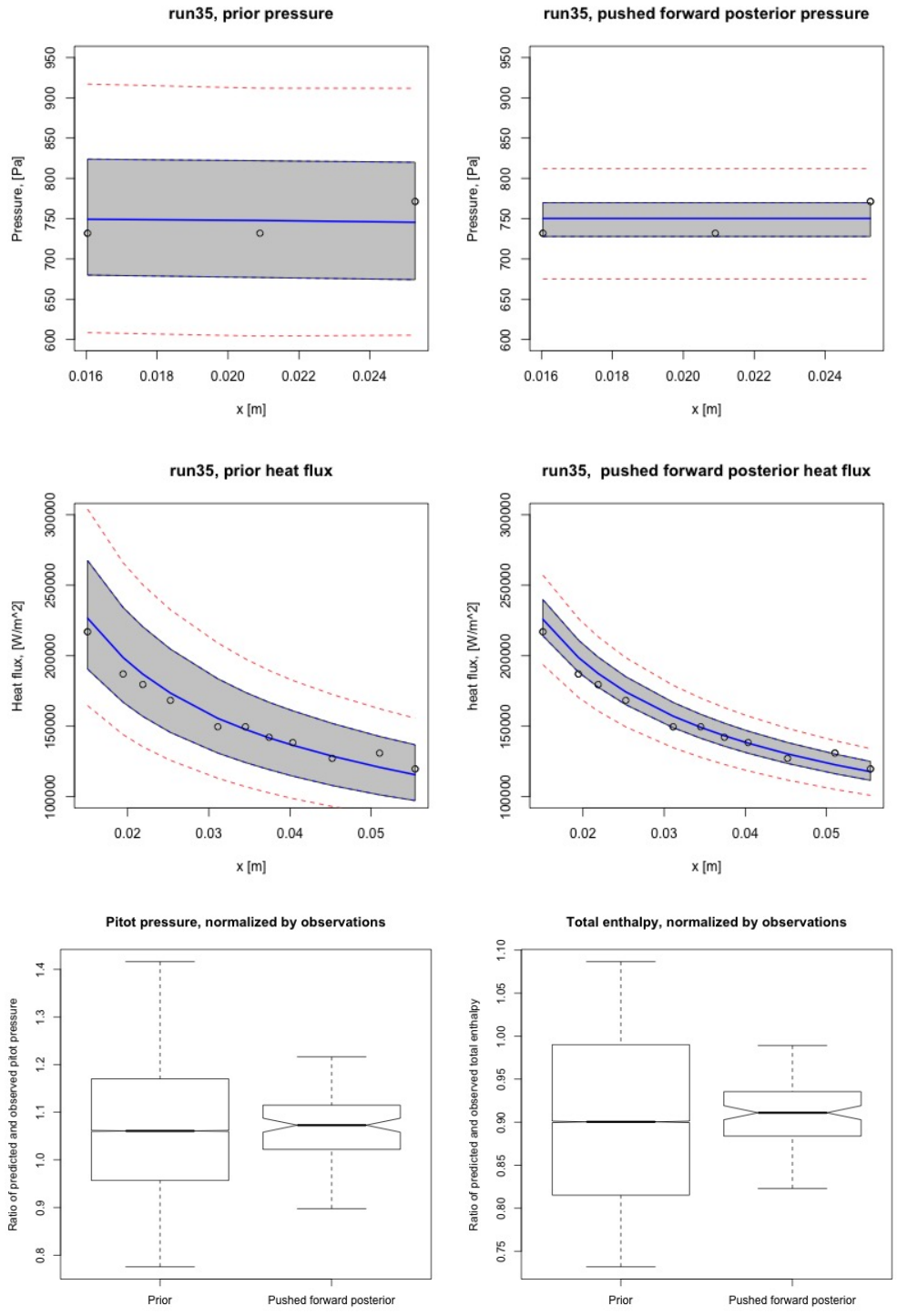

Fig. 6 Top left: Predictions of pressure on the fore cone using Θ samples drawn from the prior. The grey region is the IQR of predictions and the red dotted lines the 5th and 95th percentile predictions. The blue line is the median prediction. Top right: Pressure predictions from the PFP. Middle row: Same as the top row, but for heat flux on the fore cone. Bottom left: Box-and-whisker plot of predictions of Pitot pressure using Θ sampled from the prior and PFP and normalized by the measured value. Bottom right: Similar plots for h_{0_∞}.

using our statistical emulators. These predictions are called the "push-forward posterior" (PFP) and, due to the use of emulators, are limited to the region of the fore cone, before the separation point. We repeat the same exercise using samples taken from the prior distribution, with the aim of verifying whether the assimilation of observational data improved predictions over what was possible with the prior. In Fig. 6 (top row), we plot the predictions of the pressure and heat flux at the probes due to samples from the prior and the posterior PDFs. The region between the first and third quartile (the inter-quartile range or IQR) is plotted in grey, whereas the 5^{th}- and 95^{th}-percentile predictions are plotted in red. The blue line in the middle is the median prediction. Clearly, the PFP predictions are narrower, demonstrating the effectiveness of assimilating experimental data. A similar behavior is observed in the middle row, which plots the heat flux measured on the fore cone. Finally, in the bottom row, we plot the Pitot-pressure predictions from the prior and posterior densities, normalized by the measured Pitot pressure. A value of unity implies a perfect match between the experiment and predictions. The bold line in the middle is the ratio of the median prediction to the experimental value. We see that the median prediction is over-predicted by about 5%. While the median predictions before and after calibration are not very different, the uncertainty in the prediction (vertical size of the box) is far smaller after calibration, i.e., in the PFP. In the bottom row (right), we plot the total enthalpy. Qualitatively, we see the same trend as for Pitot pressure, except the median predicted enthalpy is under-predicted by about 10%. The reason is as follows. As described in Ref. [1], a part of the stagnation enthalpy in run35 was "frozen out" as vibrational energy upstream of the test section and is unavailable to the flow, i.e., the time-scale to equilibrate was too slow for the vibrational energy to contribute to the measured heat flux. Since our inference uses the heat flux to estimate the flow h_{0_∞}, the "frozen out" fraction of the enthalpy could not be accounted for by the inference. Thus, the basic requirements of calibration, i.e., being able to estimate a tight posterior density for Θ seems to be met. We conclude that the Bayesian inference procedure has been implemented correctly.

However, this does not automatically imply a high-quality estimation This is deduced from the PFP plots for pressure and heat flux. We see that, even after calibration, the IQR contains almost all of the experimental measurements; in case of an ideal calibration, only about 50% of the measurements would fall within the IQR region. Consequently, our calibration is *over-dispersed*. This is caused by the residual errors, called model-form errors, of the emulator. They are due to its approximate nature, errors in the experimental data, and the inherent shortcomings of the models in SPARC, including the boundary conditions (e.g., the assumption of uniform inflow) and the assumption of axisymmetry.

Finally, we check whether calibration using measurements from the laminar fore-cone region improves flow predictions over the entire double cone. To do so, we take 100 random samples from the three-dimensional PDF over $\{\rho_\infty, U_\infty, T_\infty\}$ and simulate flow over the double cone for these realizations of the freestream conditions. These simulations are conducted on a 1024×512 mesh and are, in essence, PFPs computed using SPARC. The grid resolution was chosen to resolve the shock–boundary layer interaction in the separation bubble. In Fig. 7, we plot the median predictions from this ensemble of runs, along with a few quantiles against the experimental data. Predictions using the "nominal" freestream conditions from Table 1 are also plotted. We see that the 5^{th} and 95^{th} quantiles of the prediction bracket the experimental data on the fore cone, and, for pressure, also in the separation bubble. However, in the post-attachment region, experimental measurements of both $p(x)$ and $q(x)$ tend to be near the 95^{th} percentile of the predictions, i.e., the calibrated freestream under-predicts the flow in those regions.

B. Investigating Case 1

Having verified in Sec. IV.A that Bayesian estimation can yield useful PDFs of Θ, we proceed to Case 1. Table 1 provides the nominal experimental conditions and the uncertainties in them. The case is of moderate enthalpy, and the simulation accounts for vibrational non-equilibrium, i.e., we assumed that $T \neq T_v$, though, nominally, the freestream is in vibrational equilibrium. The flow over the double cone does not have an appreciable degree of dissociation or reaction between oxygen and nitrogen. We assume that freestream uncertainties mentioned in Table 1 are untrustworthy (as we will see later, this assumption is not met) and instead use a $\pm 15\%$ uniform distribution as our prior probability density for $\{\rho_\infty, U_\infty, T_\infty, T_{v_\infty}\}$. We expect that this density might be excessively wide, but the posterior density would be comfortably included inside it.

In Fig. 8, we plot the posterior distribution for Θ obtained using measurements of h_{0_∞}, P_{Pitot}, pressure, and heat flux on the attached-flow region of the fore cone. As in Fig. 5, T_∞ and T_{v_∞} cannot be estimated (they resemble uniform distributions similar to the priors). The posterior density for freestream density ρ_∞ is marginal, at best. The velocity U_∞ can be estimated quite well. The most probable estimate (or maximum *a posterior* (MAP) estimate) is plotted using the vertical green line, and the nominal values from Table 1 with a red line. We see that the difference between the nominal and inferred quantities are quite low and well within the measurement error bounds stated in Table 1. Thus, our

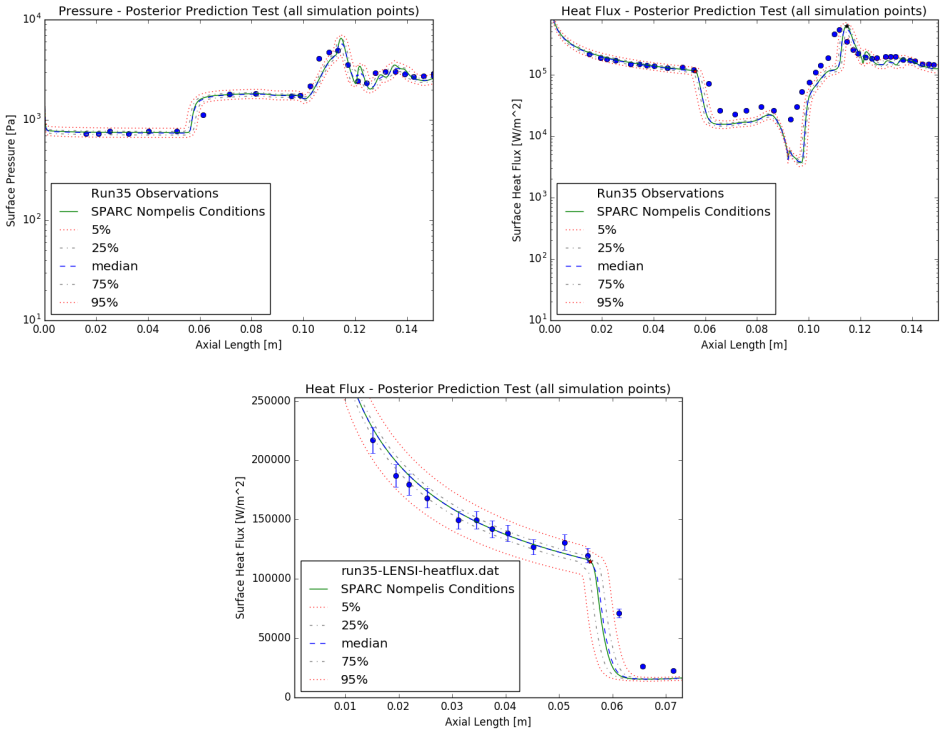

Fig. 7 Push-forward posterior distributions, corresponding to the posterior density in Fig. 5, as computed using 100 realizations of Θ. The top row of plots contains the pressure (left) and heat-flux (right) distributions. The bottom plot zooms into the fore-cone area.

assumption of ignoring the measurement error bounds and using ±15% bounds was erroneous, though conservative.

Next, we gauge the quality of the posterior PDFs plotted in Fig. 8. In Fig. 9, we plot results from push-forward-prior and PFP predictions for pressure and heat flux on the fore cone, as well as h_{0_∞} and P_{Pitot}. These predictions were performed using the statistical emulators seeded with 100 samples of Θ extracted from the posterior PDFs plotted in Fig. 8 and are limited to the fore cone. The convention for plotting is the same as in Fig. 6. We see that the prior predictions are excessively wide (not a surprise given ±15% bounds) but seem to under-predict the heat flux, h_{0_∞}, and P_{Pitot} (see the median predictions and the tendency of the observations to cluster around the third quartile of the predictions). After Bayesian calibration, the spread in the predictions is much reduced, as is the bias between the median prediction and the observations. Thus, Bayesian calibration of Θ seems to have improved predictive skill of emulator-based simulations and could, potentially, do the same for SPARC.

In Fig. 10, we perform simulations with SPARC, using the same Θ samples that were used to generate Fig. 9. On the left, we plot 5$^{\text{th}}$- and 95$^{\text{th}}$-percentile predictions, the IQR and the experimental data in red (with measurement error bounds in red dashed lines). Predictions are made for the entire double cone, not just the fore cone with the attached flow. We see that, while the agreement between the predictions and observations is excellent in the attached-flow region, there is hardly any improvement in the detached and reattached section of the flow. SPARC's prediction of the separation point is too far downstream, and then the flow re-attaches too quickly. This error is more evident in the pressure predictions than in the heat-flux predictions. On the right column of the plots, we summarize the (dis-)agreement between predictions and observations using CRPS_j and d_S, which are plotted for all of the pressure and heat-flux probes. Since CRPS_j has the same units as the quantity being plotted ($p(x)$ or $q(x)$), it should be interpreted using the vertical axis on the left. The vertical axis on the right is used for d_S. The poor agreement is evident in the CRPS_j, which assumes very large values, as well as in d_S, which seems to vary between 0.8 and 1.0. Thus, the physical processes that control the essential dynamics in the post-detachment region are not captured by SPARC. A simple mis-specification of freestream conditions is not the cause of this lack of predictive skill.

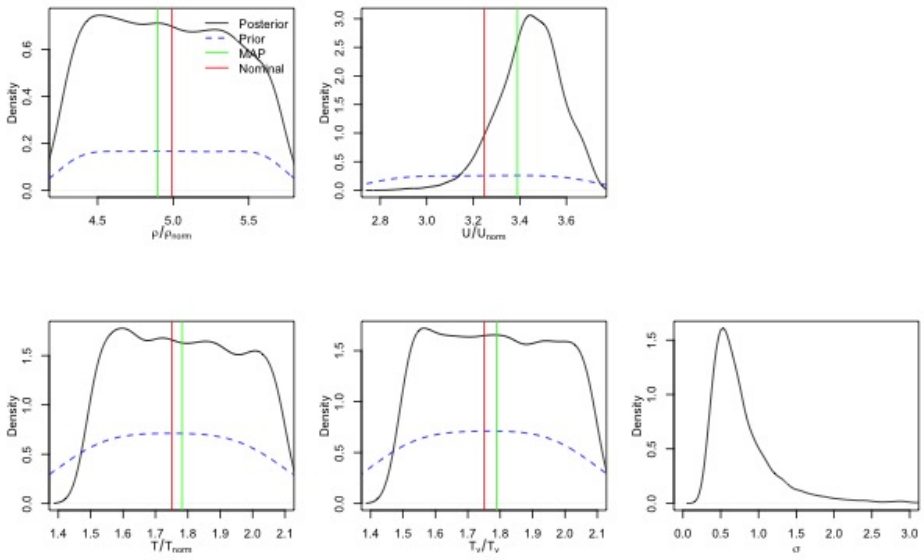

Fig. 8 Marginalized PDFs of ρ_∞, U_∞, T_∞, T_{v_∞}, and σ, inferred from Case 1 data (Table 1). The variables have been normalized to ease interpretation. $\{\rho_{\text{norm}}, u_{\text{norm}}, T_{\text{norm}}, T_{v_{\text{norm}}}\} = \{0.1\,\text{g/m}^3, 1000\,\text{m/s}, 100\,\text{K}, 100\,\text{K}\}$. The green line indicates the MAP value, as inferred. The red line is the "nominal" freestream values from Table 1. The dotted lines are the prior distribution (uniform distributions); their curved nature is an artifact of constructing them using kernel density estimation.

C. Investigating Case 4

In Sec. IV.B, we saw that, despite adjusting for errors in specifications of freestream conditions (which were not large, as we discovered after our analysis), our predictions over the entire double cone were not markedly improved. We next investigate whether the cause may be related to flow physics that are linked to the energetics of the flow, e.g., thermal and chemical non-equilibrium. To that end, we investigate Case 4, which has a total enthalpy four times larger than Case 1 (see Table 1).

First, we investigate whether the nominal freestream, as specified in Table 1, has minor deviations, as we discovered in Sec. IV.B. We perform Bayesian calibration for the freestream conditions Θ and plot the posterior PDFs in Fig. 11. Again, we find that we cannot infer T_∞ and T_{v_∞} (broad flat posterior PDFs), which is not surprising given their small Sobol' indices (see Fig. 3). Freestream density ρ_∞ is marginally inferred, whereas velocity U_∞ is inferred with quite some certainty (a clean peak in the posterior PDF). The MAP values of ρ_∞ and U_∞ are plotted with green lines and show a substantial deviation from the nominal specifications in Table 1 — around 7% – 10% and certainly far larger than their stated measurement errors. Thus, the difficulties in matching experimental data, especially in the attached, laminar fore-cone region, as seen in Ref. [2], could be due to *both* a mis-specification of freestream conditions *and* our assumption of uniformly distributed measurement errors. However, after calibration, these difficulties should be removed. We check this next.

In Fig. 12, we plot predictions from the push-forward prior and PFP simulations for pressure and heat flux on the fore cone, as well as h_{0_∞} and P_{Pitot}. These predictions were performed using the statistical emulators seeded with 100 samples of Θ extracted from the posterior PDFs plotted in Fig. 11 and are limited to the fore cone. The prior distribution is again uniform and spans ±15% around the nominal conditions in Table 1. We see that, post-calibration, the spread in the predictions has narrowed and the IQR contains about half the predictions. However, the calibration is still biased (the heat-flux measurements are greater than the median prediction in the blue line), though the bias is small for $p(x)$, h_{0_∞}, and P_{Pitot}. Thus, the posterior PDF should result in a better match with the measurements in the attached, laminar fore-cone flow region, when simulated using SPARC. The more important question is whether it improves flow predictions over the entire double cone.

In Fig. 13, we plot predictions of $p(x)$ and $q(x)$ over the entire double cone, obtained by seeding SPARC with the same 100 Θ samples used in Fig. 12. The plotting convention is the same as for Fig. 9. The sub-figures on the left in

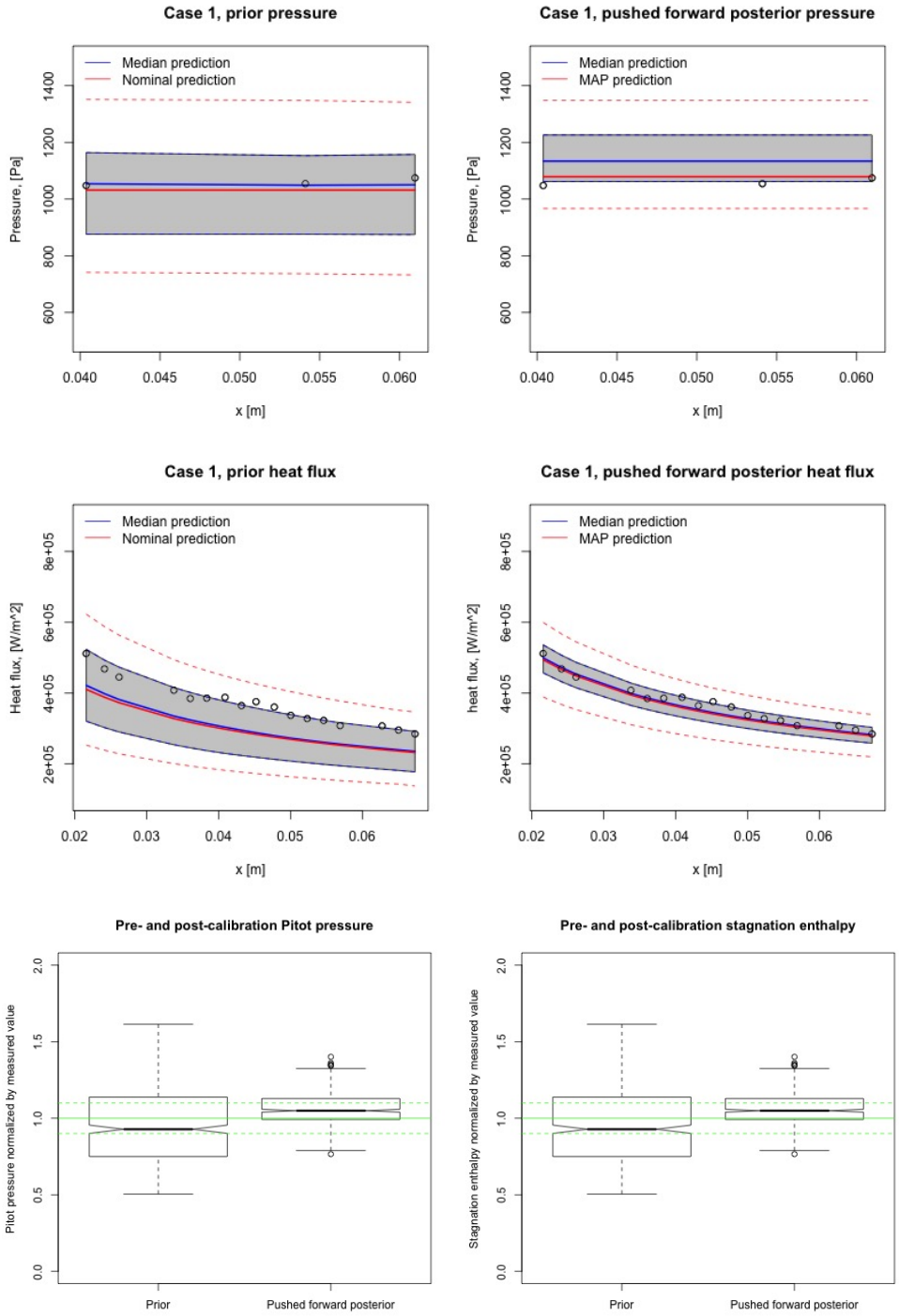

Fig. 9 Top left: Predictions of pressure on the fore cone using Θ samples drawn from the prior ($\pm 15\%$ bounds) for Case 1. The grey region is the prediction IQR and the red dotted lines the 5^{th}- and 95^{th}-percentile predictions. The blue line is the median prediction and solid red line the prediction using Θ_{MAP}. Top right: Pressure predictions from the PFP. Middle row: Same as the top row, but for heat flux. Bottom left: Box-and-whisker plot of predictions of Pitot pressure using Θ sampled from the prior and PFP and normalized by the measured value. Bottom right: Similar plots for h_{0_∞}. The solid green line denotes a perfect match between predictions and observations. The dotted green lines are the $\pm 10\%$ limits.

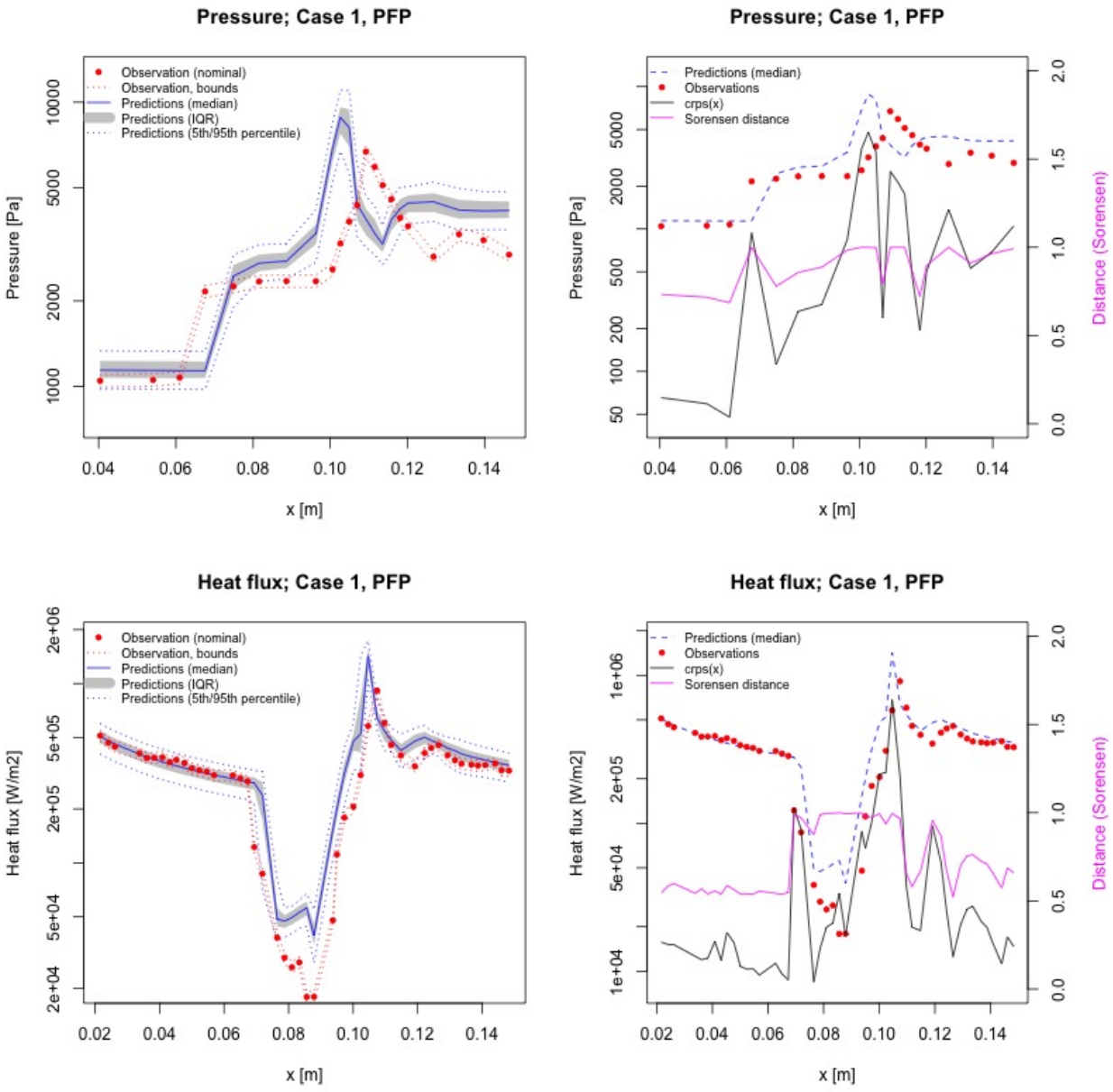

Fig. 10 Top left: Push-forward-posterior (PFP) predictions for pressure for Case 1, showing the 5th and 95th percentiles (blue dashed lines), the IQR (grey region), and median prediction (solid blue line). Observations are plotted in red symbols, with the bounds denoted by measurement errors plotted in red dashed lines. Top right: We plot CRPS$_j$ for the probes using the black line and d_S, the Sorensen distance, using the magenta line. The vertical axis for d_S is on the right. Bottom: Similar plots for the heat flux. The measurement errors for pressure and heat flux are 5% and 7% respectively.

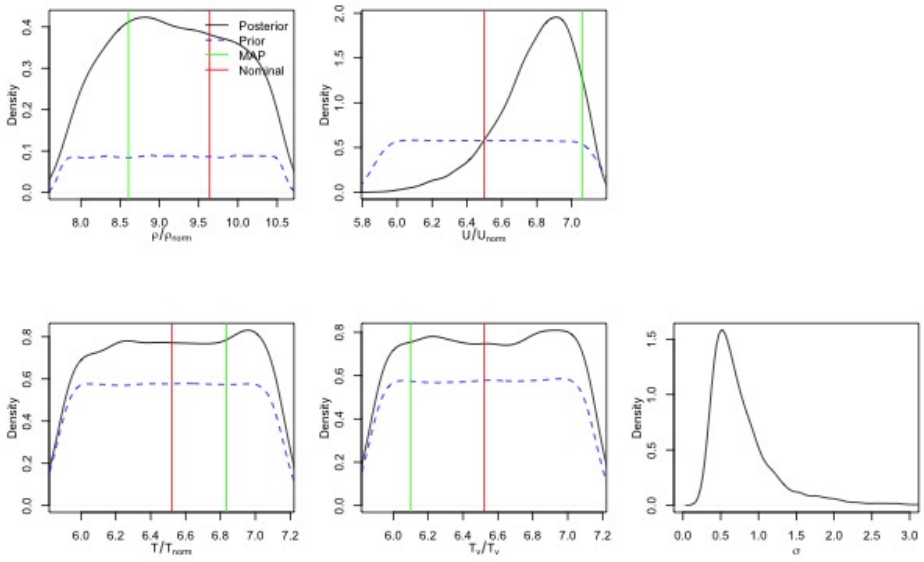

Fig. 11 Marginalized PDFs of ρ_∞, U_∞, T_∞, T_{v_∞}, and σ, inferred from Case 4 data (Table 1). The variables have been normalized to ease interpretation. $\{\rho_{\text{norm}}, u_{\text{norm}}, T_{\text{norm}}, T_{v_{\text{norm}}}\} = \{0.1\,\text{g/m}^3, 1000\,\text{m/s}, 100\,\text{K}, 100\,\text{K}\}$. The green line indicates the MAP value, as inferred. The red line is the "nominal" freestream values from Table 1. The dotted lines are the prior distribution (uniform distributions); their curved nature is an artifact of constructing them using kernel density estimation.

Fig. 13 show that the agreement between experiments and predictions of $p(x)$ and $q(x)$ is quite good in the attached-flow region. However, as in Fig. 10, the predicted separation is delayed and the simulated flow re-attaches too early. While the predicted pressure and heat flux at the reattachment point are larger than the observed values (similar to Fig. 10), the difference in their locations is smaller than in Fig. 10. However, this observation should be treated with caution — the accuracy with which we can infer the reattachment point is governed by the coarse spacing of the pressure and heat-flux probes. The sub-figures on the right plot the CRPS_j and d_S for the pressure and heat flux and here the disagreement between predictions and observations, both probabilistic, becomes evident. The CRPS_j are large in the separation region and at the reattachment point, and expectedly so, as our simulations are deficient. However, *after* reattachment, CFD predictions are closer to the experimental measurements. This improvement in pressure predictions is not seen in Case 1 (Fig. 10). In contrast, in the case of heat flux, Case 1 shows an improvement in heat-flux predictions after the reattachment point, whereas, in Case 4, we see a disagreement as CRPS_j remain high. Note that the pressure and heat flux experienced in Case 4 are higher than in Case 1 and have correspondingly wider measurement error bars (as they are specified as a percentage). The plots with d_S show that these expanded error bars do not lead to an improvement in the agreement between predictions and observations. Comparing Fig. 10 and Fig. 13, d_S saturates at 1.0 in the separation bubble, showing an utter inability to predict the flow there.

V. Discussion

Collating the findings of Secs. IV.B and IV.C, we see that the physics that govern flow separation, reattachment, and the separation bubble are not captured well by SPARC simulations. The cause is only weakly related to the total enthalpy of the flow since predictions for both Case 1 and 4 suffer from it in approximately equal measure, despite the energy in the flow differing by a factor of four. The clearest manifestation of the difference in h_{0_∞} is the degree of dissociation in the flow, which is extensive in Case 4 and rather limited for Case 1. However, this does not seem to have a marked effect on prediction errors, indicating that the 5-species Park reaction mechanism for air and other thermodynamic/thermochemical models may not be the root cause of the difficulties in reproducing the flow at the junction of the two cones. Thus, *Cause 3* seems unlikely to be the cause of our lack of predictive skill. This conclusion is not entirely unexpected. In Ref. [42], the authors investigate double-cone experiments conducted in the LENS-I tunnel

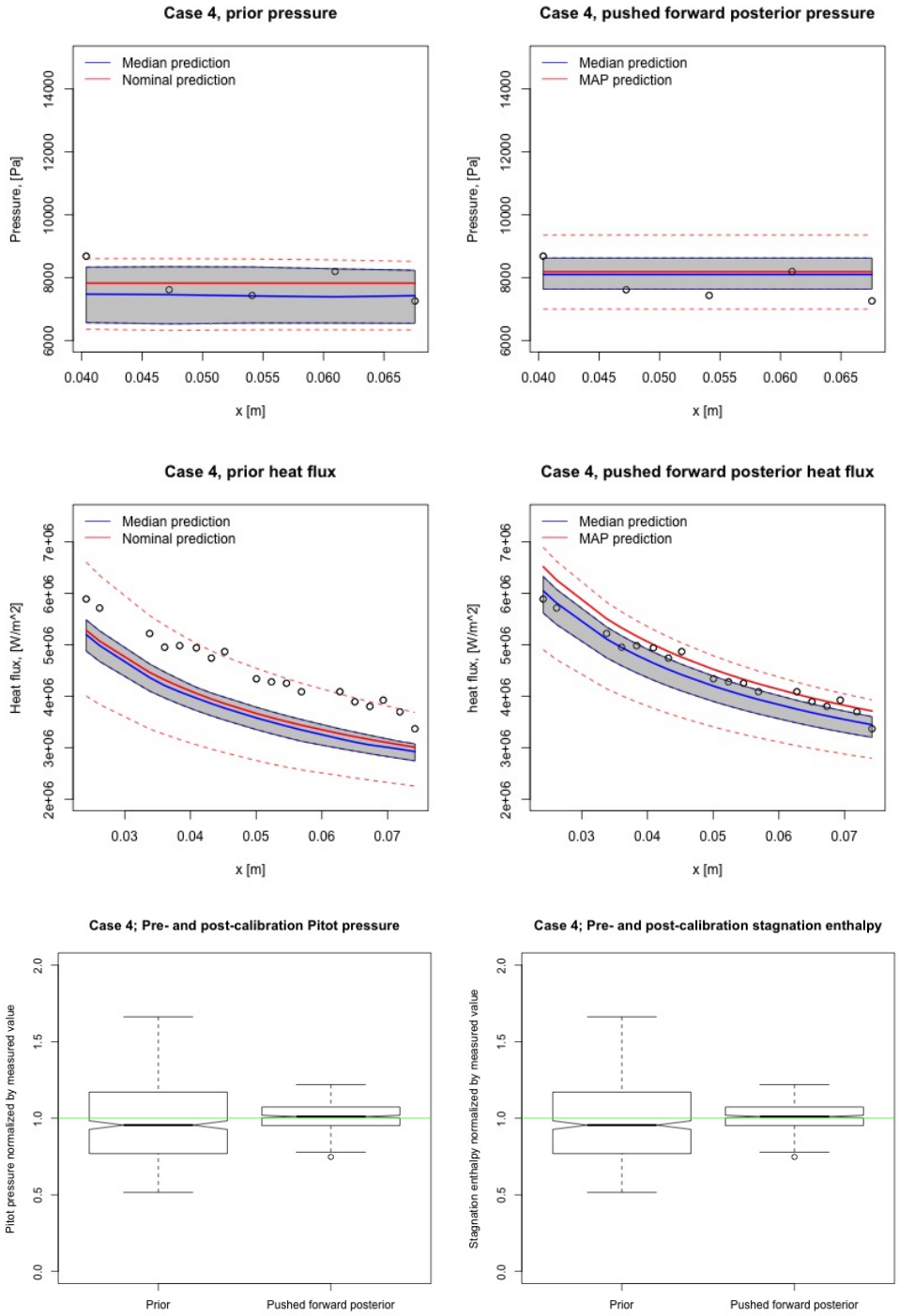

Fig. 12 Top left: Predictions of pressure on the fore cone using Θ samples drawn from the prior ($\pm 15\%$ bounds) for Case 4. The grey region is the prediction IQR and the red dotted lines the 5^{th}-and 95^{th}-percentile predictions. The blue line is the median prediction and solid red line the prediction using Θ_{MAP}. Top right: Pressure predictions from the PFP. Middle row: Same as the top row, but for heat flux. Bottom left: Box-and-whisker plot of predictions of Pitot pressure using Θ sampled from the prior and PFP and normalized by the measured value. Bottom right: Similar plots for h_{0_∞}. The solid green line denotes a perfect match between predictions and observations. The dotted green lines are the $\pm 10\%$ limits.

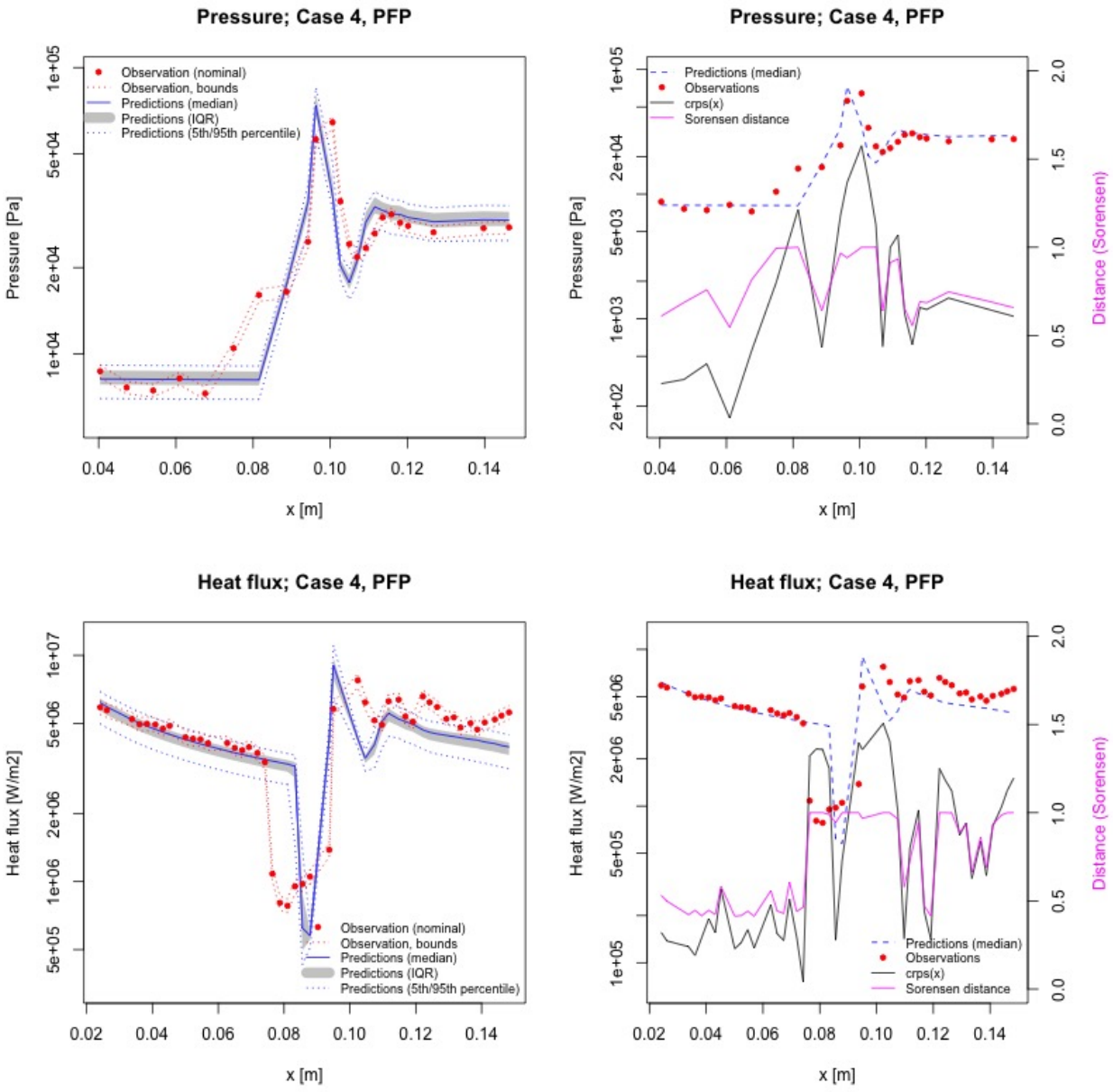

Fig. 13 Top left: Push-forward-posterior (PFP) predictions for pressure for Case 4, showing the 5th and 95th percentiles (blue dashed lines), the IQR (grey region), and median prediction (solid blue line). Observations are plotted in red symbols, with the bounds denoted by measurement errors plotted in red dashed lines. Top right: We plot CRPS$_j$ for the probes using the black line and d_S, the Sorensen distance, using the magenta line. The vertical axis for d_S is on the right. Bottom: Similar plots for the heat flux. The measurement errors for pressure and heat flux are 5% and 7% respectively.

Table 3 Comparison of mean CRPS$_j$ and d_S (computed over the entire double cone) before and after calibration, for $p(x)$ and $q(x)$. The table entries denote (mean CRPS$_j$, mean d_S). For both the cases, post-calibration, the predictive skill of the heat flux increases (lower mean CRPS and d_S) at the expense of the pressure predictions.

	$p(x)$		$q(x)$	
Experiment	Pre-calibration	Post-calibration	Pre-calibration	Post-calibration
Case 1	(1136, 0.7742)	(1183, 0.899)	$(6.23 \times 10^4, 0.87)$	$(5.36 \times 10^4, 0.734)$
Case 4	(2895, 0.6756)	(3954, 0.7882)	$(1.13 \times 10^6, 0.955)$	$(8.11 \times 10^5, 0.7248)$

and found that models for transport coefficients had limited effect on the predictions. They also found that forward reaction rates had to be "tinkered with significantly" to explain the discrepancies between LENS-I measurements and CFD predictions. In addition, the study in Ref. [19], performed for LENS-XX experiments, found that the different models for vibrational non-equilibrium provided similar results.

Next, we check whether calibration improved predictions of pressure and heat flux on the entire double cone. We do so by comparing the mean CRPS and Sorensen distances between the PFP and push-forward prior, where the prior is defined by the measurement errors (i.e., the ones in Table 1). We tabulate these figures of merit in Table 3. Note that these prediction errors are computed over the entire double cone and are dominated by the errors incurred inside the separation bubble and behind the reattachment point where most of the probes lie. We find that, for both Case 1 and Case 4, post-calibration, the agreement of the predicted heat flux with the experimental values improves, as depicted by a smaller mean CRPS and d_S. This comes at the expense of the predictive skill of the pressure field, where the mean CRPS and d_S increase, post-calibration. Thus, it does not seem that a simple mis-specification of an incoming, axisymmetric freestream is the dominant cause of the lack of predictive skill seen in Ref. [2]; had that been the case, both heat-flux and pressure predictions would have improved, post-calibration. Therefore, it seems that *Cause 1* may not be the primary reason behind our lack of predictive skill.

VI. Conclusions

In this paper, we have developed and tested a probabilistic framework for assessing the consistency of an experimental dataset, i.e., whether the stated experimental conditions are consistent with the measurements. In case they are not, our method allows one to hypothesize and test causes behind the inconsistency. While such a hypothesis-testing method may not necessarily lead to the isolation of the true cause, it can be used to eliminate candidates. The probabilistic framework synthesizes the use of similarity measures, e.g., Sorensen distances, with recent advances in the application of Bayesian inference and statistical emulators in fluid mechanics, to assess experimental datasets. Such a framework has not been demonstrated previously. The framework is tested on two experiments of flows over double cones conducted in LENS-XX.

In Ref. [2], we attempted to reproduce the experimental measurements of $p(x)$ and $q(x)$, gathered from the two LENS-XX experiments. We assumed that freestream uncertainties were such that they could be modeled as a uniform distribution around the nominal conditions specified in Table 1. The bounds of the uniform distribution are specified by the measurement errors. We failed to reproduce the observations well. In the case of Case 1, we under-predicted the heat flux somewhat, and, in Case 4, our predictions were grossly biased. We hypothesized that this prediction failure could be due to three causes: (1) a mis-specification of the freestream measurement errors, both in magnitude and type of distribution; (2) lack of axisymmetry in the flow and (3) shortcomings of the thermochemical models in SPARC. In this paper, we used the probabilistic framework to explore the hypothesized causes.

In order to address the (potential) problem of mis-specification of the freestream conditions, we performed a Bayesian inference study of the freestream conditions, conditional on observations collected from the attached, laminar flow on the fore cone of the test article. The inferred freestream conditions showed that Case 1 was not mis-specified at all, though Case 4 clearly was. We also found that calibration did indeed improve flow predictions on the fore cone significantly, but the predictive accuracy for the rest of the flow leaves much to be desired. In fact, we utterly fail to predict the separation bubble; even allowing for uncertainty in the predictions *and* observations, the two distributions do not overlap much. For *both* Case 1 and Case 4, we see that the improvement in heat-flux predictions over the entire double cone comes at the expense of the fit for pressure. It seems that there might be some mis-specification of the freestream, especially for the high-enthalpy Case 4, but it does not seem to be a major explanation for the lack of

predictive skill.

We next consider whether the thermochemical models in SPARC could be a cause of the error in predictions. We do so by comparing the predictive skill of SPARC simulations for Case 1 and 4, seeded with the inferred freestream conditions. Case 4 is four times as energetic as Case 1. Table 3 allows a simple comparison of the post-calibration prediction errors, as captured using d_S, which is constrained to vary in [0, 1]. We see that the mean d_S assumes *smaller* values for Case 4 than for Case 1; this holds true for both $p(x)$ and $q(x)$. Thus, Case 4 is *more* predictive than Case 1, implying that errors in the thermochemical models in SPARC cannot be the reason for the poor predictive errors encountered when simulating LENS-XX experiments. Our findings parallel those of Ref. [19].

This leaves the question of whether the errors are due to flow non-uniformities, i.e., lack of axisymmetry in the freestream. A simple mis-specification of the inflow conditions does not seem to be a cause of our prediction errors. Consequently, it may be possible that the inflow itself may not be axisymmetric, and the lack of axisymmetry may be the cause of the prediction errors. However, the collapse of experimental data via a self-similar transformation (Eq. 1), as plotted in Fig. 3 for Case 1, shows that any 3D non-uniformities are small, i.e., for the attached fore-cone flow, the dominant length-scale is the downstream distance. A similar collapse is also observed for Case 4 and run35. It is unknown if sufficiently small 3D non-uniformities would explain the under-prediction of the separation bubble by our axisymmetric simulations without materially impacting the self-similar collapse. We have not performed any investigation with 3D simulations, which is left to future work. Catalytic effects, if included in SPARC simulations, could potentially improve predictions aft of the reattachment point, but the improvement would be limited to the highest-enthalpy case (Case 4).

Ideally, the process of using statistical inference to quantity uncertainties and establishing freestream quantities should be performed alongside the design of a specific experiment. In this manner, uncertainty quantification, numerical simulation, and experiments can complement each other to effectively and efficiently characterize the experiment. On the numerical side, a well-validated physical model and established numerical methodology are required (e.g., as done in Ref. [1]). On the experimental side, data from calibration runs should be used to provide more information to the uncertainty-quantification methodology (such as the one described here), so that the experimental conditions may be fully characterized, with minimal assumptions. Since calibration runs are performed ahead of the main experiment, this allows one to decide, e.g., using the methods described in Ref. [2], how rigorously various aspects of the experiment need to be controlled. These approaches are possible within existing practices in experimental facilities, such as CUBRC, where racks of probes and reference shapes are used to acquire data at nominal test conditions prior to the actual experiment.

Appendix

SPARC is a second-order-accurate finite-volume compressible-flow simulator. It solves the continuum conservation laws for mass, momentum, and energy, formulated for reacting gases in thermochemical non-equilibrium. Vibrational and translational/rotational internal energy modes are tracked by separate conservation laws. The equilibration of vibrational energy is governed by the Landau–Teller model [43], with relaxation time-scales computed using the Millikan–White expression [44]. For chemical non-equilibrium, SPARC solves a partial differential equation for the density of each species; the simulations in this paper used a 5-species, 17-reaction chemical mechanism for air [28] to model the time-dependent change in the chemical composition. For non-equilibrium flows ($T \neq T_v$), an effective temperature $T_{\text{eff}} = \sqrt{TT_v}$ is used to compute the chemical-reaction rates. Diffusion of individual species is modeled using Lewis numbers, and viscosities for the species are obtained using Blottner's model [45]. Thermal conductivities for the individual species are obtained from Eucken's relations [43], which are then assembled into the mixture's value using Wilke's model [46]. The equations are solved using a finite-volume method for the conserved variables. SPARC accommodates structured and unstructured meshes, though only the former are used in this paper. For the simulations in this paper, we use a Steger–Warming scheme for the inviscid fluxes, extended to second-order accuracy using MUSCL reconstruction. A minmod limiter is used in our simulations. Central differences are used for diffusion and viscous terms. SPARC solves the unsteady form of the conservation equations using a three-point backward-difference scheme to achieve second-order accuracy in time. However, all the simulations used in this paper are steady, and the time-integrator is run in its first-order form to accelerate convergence to a steady-state solution.

Acknowledgments

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & 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-NA0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

[1] Nompelis, I., Candler, G. V., and Holden, M. S., "Effect of vibrational nonequilibrium on hypersonic double-cone experiments," *AIAA Journal*, Vol. 41, No. 11, 2003, pp. 2162–2169. doi:10.2514/2.6834.

[2] Kieweg, S., Ray, J., Weirs, V. G., Carnes, B., Dinzl, D., Freno, B., Howard, M., Rider, W. J., and Smith, T., "Uncertainty quantification, sensitivity analysis, and validation assessment of laminar hypersonic double-cone flow simulations," *AIAA 2019 Aerosciences Conference, AIAA SciTech 2019 (AIAA-2019-2278)*, 2019.

[3] MacLean, M., Dufrene, A., Wadhams, T., and Holden, M., "Numerical and experimental characterization of high enthalpy flow in an expansion tunnel facility," *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition, Aerospace Sciences Meetings, (AIAA-2010-1562)*, 2010. doi:10.2514/6.2010-1562.

[4] Dufrene, A., MacLean, M., Parker, R., Wadhams, T., and Holden, M., "Characterization of the new LENS expansion tunnel facility," *48th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition (AIAA-2010-1564)*, 2010. doi:10.2514/6.2010-1564.

[5] Nompelis, I., Candler, G., Holden, M., and Wadhams, T., "Numerical simulation of high-enthalpy experiments in the LENS X expansion tube facility," *42nd AIAA Aerospace Sciences Meeting and Exhibit, Aerospace Sciences Meetings, (AIAA-2004-1000)*, 2004. doi:10.2514/6.2004-1000.

[6] Ray, J., Hou, Z., Huang, M., Sargsyan, K., and Swiler, L., "Bayesian calibration of the Community Land Model using surrogates," *SIAM Journal on Uncertainty Quantification*, Vol. 3, No. 1, 2015, pp. 199–233. doi:10.1137/140957998.

[7] Ray, J., Lefantzi, S., Arunajatesan, S., and Dechant, L., "Bayesian parameter estimation of a $k-\epsilon$ model for accurate jet-in-crossflow simulations," *American Institute of Aeronautics and Astronautics Journal*, Vol. 54, No. 8, 2016, pp. 2432–2448. doi:10.2514/1.J054758.

[8] MacLean, M., Holden, M. S., and Dufrene, A., "Measurements of real gas effects on regions of laminar shock wave/boundary layer interaction in hypervelocity flows," Tech. rep., CUBRC Inc., 2014. `https://www.cubrc.org/_iassets/docs/laminar-xx-paper.pdf`.

[9] Kaipio, J., and Somersalo, E., *Statistical and computational inverse problems*, Springer, 2005.

[10] Aster, R. C., Borchers, B., and Thurber, C. H., *Parameter estimation and inverse problems*, 3rd ed., Elsevier, 2018.

[11] Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (eds.), *Markov Chain Monte Carlo in Practice*, Chapman & Hall, 1996.

[12] Liang, F., Liu, C., and Carroll, R. J., *Advanced Markov chain Monte Carlo methods*, Wiley, 2010.

[13] Safta, C., Blaylock, M., Templeton, J. A., Domino, S., Sargsyan, K., and Najm, H., "Uncertainty quantification in LES of channel flow," *International Journal for Numerical Methods in Fluids*, Vol. 83, No. 4, 2017, pp. 376–401. doi:10.1002/fld.4272, URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/fld.4272`.

[14] Edeling, W. N., Cinnella, P., Dwight, R. P., and Bijl, H., "Bayesian estimates of parameter variability in $k-\epsilon$ turbulence model," *Journal of Computational Physics*, Vol. 258, 2014, pp. 73–94. doi:10.1016/j.jcp.2013.10.027.

[15] Guillas, S., Glover, N., and Malki-Epshtein, L., "Bayesian calibration of the constants of the $k-\epsilon$ turbulence model for a CFD model of street canyon flow," *Computer Methods in Applied Mechanics and Engineering*, Vol. 279, 2014, pp. 536–553. doi:10.1016/j.cma.2014.06.008.

[16] Zhang, J., and Fu, S., "An efficient Bayesian uncertainty quantification approach with application to $k-\omega-\gamma$ transition modeling," *Computers & Fluids*, Vol. 161, 2018, pp. 211 – 224. doi:10.1016/j.compfluid.2017.11.007.

[17] Viana, F. A. C., Simpson, T. W., Balabanov, V., and Toporov, V., "Metamodeling in multidisciplinary design optimization: How far have we really come?" *AIAA Journal*, Vol. 52, 2014, pp. 670–690. doi:10.2514/1.J052375.

[18] Youssefi, M. R., and Knight, D., "Assessment of CFD capability for hypersonic shock wave laminar boundary layer interactions," *Aerospace*, Vol. 4, No. 25, 2017. doi:10.3390/aerospace4020025.

[19] Hao, J., Wang, J., and Lee, C., "Numerical simulation of high-enthalpy double-cone flows," *American Institute of Aeronautics and Astronautics Journal*, Vol. 55, No. 6, 2017, pp. 2471–2475. doi:10.2514/1.J055746.

[20] Candler, G. V., "Next-generation CFD for hypersonic and aerothermal flows (Invited)," *22nd AIAA Computational Fluid Dynamics Conference, AIAA AVIATION Forum, (AIAA 2015-3048)*, 2015. doi:10.2514/6.2015-3048.

[21] Park, C., "On convergence of computation of chemically reacting flows," *23rd AIAA Aerospace Sciences Meeting (AIAA-1985-0247)*, 1985. doi:10.2514/6.1985-247.

[22] Marrone, P. V., and Treanor, C. E., "Chemical relaxation with preferential dissociation from excited vibrational levels," *Physics of Fluids*, Vol. 6, No. 9, 1963, pp. 1215–1221. doi:10.1063/1.1706888.

[23] Nompelis, I., and Candler, G. V., "US3D predictions of double-cone and hollow cylinder and cylinder-flare at high enthalpy," *44th AIAA Fluid Dynamics Conference (AIAA-2014-3366)*, 2014. doi:10.2514/6.2014-3366.

[24] Holden, M., "Experimental studies of laminar separated flows induced by shock wave/boundary layer and shock/shock interaction in hypersonic flows for CFD validation," *38th Aerospace Sciences Meeting and Exhibit, Aerospace Sciences Meetings, (AIAA-2000-0930)*, 2000. doi:10.2514/6.2000-930.

[25] Harvey, J., Holden, M., and Wadhams, T., "Code validation study laminar shock/boundary layer and shock/shock interactions in hypersonic flow Part B: Comparison with Navier–Stokes and DSMC solutions," *39th Aerospace Sciences Meeting and Exhibit, Reno, NV, U.S.A (AIAA-2001-1031)*, 2001. doi:10.2514/6.2001-1031.

[26] Nompelis, I., Candler, G., MacLean, M., Wadhams, T., and Holden, M., "Numerical investigation of high enthalpy chemistry on hypersonic double-cone experiments," *43rd AIAA Aerospace Sciences Meeting and Exhibit, Aerospace Sciences Meetings, (AIAA-2005-0584)*, 2005. doi:10.2514/6.2005-584.

[27] Wadhams, T., "Personal communications," , 2018.

[28] Park, C., "Review of chemical-kinetic problems of future NASA missions I: Earth entries," *Journal of Thermophysics and Heat Transfer*, Vol. 7, No. 3, 1990, pp. 385–398. doi:10.2514/3.431.

[29] Adams, B., Bauman, L., Bohnhoff, W., Dalbey, K., Ebeida, M., Eddy, J., Eldred, M., Hough, P., Hu, K., Jakeman, J., Stephens, J., Swiler, L., Vigil, D., , and Wildey, T., "Dakota, A multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 6.0 user's manual," Tech. Rep. Technical Report SAND2014-4633, Sandia National Laboratories, Albuquerque, NM, July 2014. Updated November 2015 (Version 6.3).

[30] Carnes, B., Weirs, V. G., and Smith, T., "Code verification and numerical error estimation with application to model validation of laminar, hypersonic flow over a double cone," *AIAA 2019 Aerosciences Conference, AIAA SciTech 2019 (AIAA-2019-2175)*, 2019.

[31] Schlichting, H., and Gersten, K., *Boundary Layer Theory*, 8$^{\text{th}}$ ed., Springer, 2000.

[32] Smith, R. C., *Uncertainty quantification: Theory, implementation and applications*, Society for Industrial and Applied Mathematics, Computational Science and Engineering Series, 2014.

[33] Haario, H., Laine, M., Mira, A., and Saksman, E., "DRAM-Efficient adaptive MCMC," *Statistics and Computing*, Vol. 16, No. 4, 2006, pp. 339–354. doi:10.1007/s11222-006-9438-0.

[34] R Core Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2012. URL `http://www.R-project.org/`, ISBN 3-900051-07-0.

[35] Soetaert, K., and Petzoldt, T., "Inverse modelling, sensitivity and Monte Carlo analysis in R Using Package FME," *Journal of Statistical Software*, Vol. 33, No. 3, 2010, pp. 1–28. URL `http://www.jstatsoft.org/v33/i03/`.

[36] Raftery, A., and Lewis, S. M., "Implementing MCMC," *Markov Chain Monte Carlo in Practice*, edited by W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Chapman and Hall, 1996, pp. 115–130.

[37] Warnes, G. R., and with contributions by Robert Burrows, *mcgibbsit: Warnes and Raftery's MCGibbsit MCMC diagnostic*, 2011. URL `http://CRAN.R-project.org/package=mcgibbsit`, R package version 1.0.8.

[38] Cleveland, W. S., Grosse, E., and Shyu, W. M., "Local regression models," *Statistical Models in S*, edited by J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole, 1992, Chap. 8.

[39] Gneiting, T., and Raftery, A. E., "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, Vol. 102, No. 477, 2007, pp. 359–378. doi:10.1198/016214506000001437, URL http://pubs.amstat.org/doi/abs/10.1198/016214506000001437.

[40] Gneiting, T., Balabdaoui, F., and Raftery, A. E., "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Vol. 69, No. 2, 2007, pp. 243–268. doi:10.1111/j.1467-9868.2007.00587.x, URL http://dx.doi.org/10.1111/j.1467-9868.2007.00587.x.

[41] Cha, S.-H., "Comprehensive survey on distance/similarity measures between probability density functions," *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 1, No. 4, 2007.

[42] Druguet, M.-C., Candler, G. V., and Nompelis, I., "Comparison of physical models in computations of high-enthalpy double-cone flows," *9th AIAA/ASME Joint Thermophysics and Heat Transfer Conference (AIAA-2006-3419)*, 2006. doi:10.2514/6.2006-3419.

[43] Vincenti, W., and Kruger, C., *Introduction to physical gas dynamics*, Krieger Publishing Company, Malabar, FL, USA, 1965.

[44] Millikan, R., and White, D., "Systematics of vibrational relaxation," *Journal of Chemical Physics*, Vol. 39, 1963, pp. 3209–3213. doi:10.1063/1.1734182.

[45] Blottner, F. G., Johnson, M., and Ellis, M., "Chemically reacting viscous flow program for multicomponent gas mixtures," Tech. Rep. SC-RR-70-754, Sandia National Laboratories, Albuquerque, NM, 1971.

[46] Anderson, J. D., *Hypersonic and High-Temperature Gas Dynamics, Second Edition*, AIAA Education Series, 2006. doi: 10.2514/4.861956.