

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/301717737>

Using field inversion to quantify functional errors in turbulence closures

Article · in Physics of Fluids · April 2016

DOI: 10.1063/1.4947045

CITATIONS

30

READS

99

2 authors, including:



Karthik Duraisamy

University of Michigan

140 PUBLICATIONS 1,021 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Physics-based Turbulence Modeling [View project](#)



Data-enabled modeling of computational physics [View project](#)



Using field inversion to quantify functional errors in turbulence closures

Anand Pratap Singh and Karthik Duraisamy

Citation: Physics of Fluids **28**, 045110 (2016); doi: 10.1063/1.4947045

View online: <http://dx.doi.org/10.1063/1.4947045>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/pof2/28/4?ver=pdfcov>

Published by the AIP Publishing

Articles you may be interested in

Large-eddy simulation of shock-cooling-film interaction at helium and hydrogen injection
Phys. Fluids **25**, 106101 (2013); 10.1063/1.4823745

Control of separated flow in a reflected shock interaction using a magnetically-accelerated surface discharge

Phys. Fluids **24**, 126102 (2012); 10.1063/1.4772197

Detailed flow physics of the supersonic jet interaction flow field

Phys. Fluids **21**, 046101 (2009); 10.1063/1.3112736

The effect of Mach number on unstable disturbances in shock/boundary-layer interactions

Phys. Fluids **19**, 054104 (2007); 10.1063/1.2720831

Direct numerical simulation of impinging shock wave/turbulent boundary layer interaction at $M = 2.25$

Phys. Fluids **18**, 065113 (2006); 10.1063/1.2216989

An advertisement for the new journal AIP APL Photonics. On the left, a thumbnail of the journal cover for January 2016, Volume 1, Number 1 is shown. The cover features a grayscale image of a laser beam profile and the text "OPEN ACCESS". To the right of the thumbnail, the text "Launching in 2016! The future of applied photonics research is here" is written in large, white, sans-serif font against a bright orange and yellow background. In the bottom right corner, the AIP APL Photonics logo is displayed.

Using field inversion to quantify functional errors in turbulence closures

Anand Pratap Singh^{a)} and Karthik Duraisamy^{b)}

Department of Aerospace Engineering, University of Michigan, Ann Arbor,
Michigan 48109, USA

(Received 10 November 2015; accepted 3 April 2016; published online 28 April 2016)

A data-informed approach is presented with the objective of quantifying errors and uncertainties in the functional forms of turbulence closure models. The approach creates modeling information from higher-fidelity simulations and experimental data. Specifically, a Bayesian formalism is adopted to infer discrepancies in the source terms of transport equations. A key enabling idea is the transformation of the functional inversion procedure (which is inherently infinite-dimensional) into a finite-dimensional problem in which the distribution of the unknown function is estimated at discrete mesh locations in the computational domain. This allows for the use of an efficient adjoint-driven inversion procedure. The output of the inversion is a full-field of discrepancy that provides hitherto inaccessible modeling information. The utility of the approach is demonstrated by applying it to a number of problems including channel flow, shock-boundary layer interactions, and flows with curvature and separation. In all these cases, the posterior model correlates well with the data. Furthermore, it is shown that even if limited data (such as surface pressures) are used, the accuracy of the inferred solution is improved over the entire computational domain. The results suggest that, by directly addressing the connection between physical data and model discrepancies, the field inversion approach materially enhances the value of computational and experimental data for model improvement. The resulting information can be used by the modeler as a guiding tool to design more accurate model forms, or serve as input to machine learning algorithms to directly replace deficient modeling terms. *Published by AIP Publishing.* [<http://dx.doi.org/10.1063/1.4947045>]

I. INTRODUCTION

Accurate simulation of turbulent flows is of critical importance to many applications in science and engineering. Despite the impressive growth in computational resources over the past two decades, modeling and simulation of many practical turbulent flows — to the desired level of accuracy — have remained challenging, and in some cases, inaccessible. Engineering prediction of turbulent flow has relied on closure models for the Reynolds Averaged Navier-Stokes equations (RANS). Though Direct Numerical Simulations (DNS) and Large Eddy Simulations (LES) have provided previously unattainable insights and predictive capabilities in many flows, these techniques continue to be infeasible for high Reynolds number wall-bounded flows. This situation is unlikely to change within the next few decades unless significant advances are made in hybrid techniques that employ a near-wall model in conjunction with an outer-layer LES. Near-wall models in both a LES and RANS context will be the pacing item in applied computational fluid dynamics.¹

Despite the simplifying assumptions, RANS models have found utility in several engineering problems. However, it is well-recognized that they lack accuracy for flows with adverse pressure

^{a)}Electronic mail: anandps@umich.edu

^{b)}Electronic mail: kdur@umich.edu

gradients, separation, curvature/rotation, secondary flows, etc. Turbulence closure models usually involve a number of additional transport variables and many adjustable constants selected by the engineering judgment of the modeler. These constants are calibrated by a small number of canonical test cases such as homogeneous turbulence and thin-shear flows. Given this development process, it is unsurprising that accuracy diminishes as the model is applied to problems which deviate from the calibrated cases. In attempts to increase accuracy, some work has shifted toward more complex turbulence models. As a recent example, Gerolymos *et al.*² attempted to model the various terms in differential Reynolds stress closures in great detail. Particular care is devoted to modeling the redistribution terms. Additional tensorial bases are introduced to account for inhomogeneity, wall echo, anisotropies, etc. Recent efforts have involved the exploration of third³ and fourth⁴ moment closures for complex flows. While these efforts are encouraging and lead to improved understanding of turbulence modeling, potential benefits are often obscured by the need to determine a number of free parameters from a small set of often idealized test cases.

Since RANS models have always been empirical to a large degree, it can be well-argued that these models have always been data-driven. At the same time, tremendous advances in computing power and measurement resolution have provided datasets containing complete information in problems of reduced scope or essential information in regimes of interest. Despite the availability of such high quality DNS, LES, and PIV data, the improvement in closure models has not been significant.

The key impediment to improving closure models from higher fidelity data is the issue of model discrepancy. Approaches to address model discrepancy and calibration are inherently philosophical in nature, and excellent papers on these topics exist.^{5,6} For the purposes of this work, however, the role of model discrepancy can be illustrated by the example of a typical turbulence closure model. Since turbulence models are typically formulated to provide first and second moments, secondary variables such as the dissipation rate typically assume different values in closure models compared to their values in the real-world or in higher-fidelity simulations. Thus, approaches that attempt to extract *a priori* information from DNS may not be very useful in a predictive setting, since the *extracted information will be inconsistent with the information processed by the model*. Further, even if individual terms in the closure equations are modeled accurately, small discrepancies can lead to poor results. For instance, Poroseva and Murman⁷ observed that even if some terms in a second moment closure are extracted directly from DNS, the overall prediction was not satisfactory. It is thus the *balance* between every term in the model equation that is important. To obtain useful data-informed models, one should pursue methods that produce *information* within the context of the model.

At this stage, it would be convenient to categorize modeling errors into those resulting from (a) the model structure (such as the choice of an algebraic Reynolds stress model or $k-\epsilon$ model), (b) the chosen functional forms within the model structure, and (c) parameters in the chosen functional form. Much of the data-based approaches that have been applied in turbulence modeling have involved the calibration of a set of model parameters. Yarlanki *et al.*,⁸ Kato and Obayashi,⁹ and Edeling *et al.*¹⁰ apply modern data-assimilation techniques to infer model parameters. These techniques provide insight into the variability of model outputs, but are limited in scope since the functional form of the model is frozen, thus restricting the scope and generality of the calibration. Other attempts have approached the problem from the viewpoint of *characterizing* structural errors. Dow and Wang¹¹ infer the error in the coefficient of the eddy viscosity relationship by assuming it to be a Gaussian random field. Cheung *et al.*¹² employ Bayesian model averaging to calibrate turbulence models. Emory *et al.*¹³ inject realizable perturbations to the Reynolds stresses computed by eddy viscosity models to obtain uncertainty bounds on the predictions. Ling and Templeton¹⁴ employ machine learning techniques directly on high and low-fidelity data to identify regions in which high turbulence modeling uncertainty may be present.

In this work, a more comprehensive approach towards extracting modeling information from data is pursued, with the goal of ultimately improving our qualitative and quantitative understanding of model errors. In contrast to inferring calibration parameters, the functional form of the model discrepancy is targeted. The full field of modeling terms provides hitherto inaccessible information within the context of the model. This information can be used by the modeler as a guiding tool

to design parametric model forms, or serve as input to machine learning algorithms to supplement models.^{15–17}

The rest of the paper is organized as follows: Section II introduces the inversion framework and methodology for functional error and uncertainty quantification and Sections III–VI present the application of the inversion framework to four different turbulence modeling scenarios. The summary and future perspectives are provided in Section VII.

II. METHODOLOGY

The proposed method can be adapted for near-wall modeling in LES, or for improving closures in general. Our focus in this paper is on a subset of RANS-based turbulence models which are based on the Boussinesq approximation. While this class of models suffer from inherent model inadequacies, the examples are intended to demonstrate how useful modeling information can be derived from data. More general improvements may be achievable by applying the procedure to second moment closures.

In linear eddy viscosity closures, transport equations are formulated for surrogate variables which are then used to define the eddy viscosity (ν_t). For example, in the one-equation Spalart–Allmaras (SA) model, the eddy viscosity is calculated using a surrogate viscosity $\tilde{\nu}_t$; which in turn satisfies the following transport equation:

$$\frac{D\tilde{\nu}_t}{Dt} = P(\tilde{\nu}_t, \mathbf{U}) - D(\tilde{\nu}_t, \mathbf{U}) + T(\tilde{\nu}_t, \mathbf{U}), \quad (1)$$

where \mathbf{U} contains the Reynolds averaged flow variables, $P(\tilde{\nu}_t, \mathbf{U})$, $D(\tilde{\nu}_t, \mathbf{U})$, and $T(\tilde{\nu}_t, \mathbf{U})$ are the production, destruction, and transport terms, respectively. Similarly, the $k-\omega$ model uses transport equations for turbulent kinetic energy k and specific dissipation rate ω . For such models, a key challenge is the accurate formulation of source terms in the transport equations as they have to be modified to represent important phenomena such as streamline curvature and relaminarization. In this work, the focus is on using data to comprehensively inform the functional form of the source term. It is pertinent to recognize that modifying the source terms does indeed change the balance of the entire model. Further, because of the underlying transport equation, a change in the source term affects the solution domain in a non-local sense. At this stage, it is noted that the Boussinesq approximation may also introduce an unrecoverable discrepancy.¹⁸ As mentioned earlier, it is not the object of this work to propose radically improved turbulent models, but rather to highlight the potential of data to inform turbulence models.

Without a loss of generality, the applications considered in this work involve steady flows, and thus the inferred functional forms are of a spatial — instead of spatio-temporal — nature. To estimate the defect in the source term balance, a spatially varying term $\beta(\mathbf{x})$ is introduced as a multiplier of a source term $S(U(\mathbf{x}))$, typically the production term, the destruction term, or the balance between the two terms. The multiplication by $\beta(\mathbf{x})$ is equivalent to adding a source term $\delta(\mathbf{x}) = (\beta(\mathbf{x}) - 1)S(U(\mathbf{x}))$. For example, in the context of Spalart–Allmaras model, $S(U(\mathbf{x}))$ can be the production term. Then the modified transport equation will be given by

$$\frac{D\tilde{\nu}_t}{Dt} = \beta(\mathbf{x})P(\tilde{\nu}_t, \mathbf{U}) - D(\tilde{\nu}_t, \mathbf{U}) + T(\tilde{\nu}_t, \mathbf{U}). \quad (2)$$

The multiplication to the source rather than a stand-alone correction function, $\delta(\mathbf{x})$ presents several advantages to the inversion process as $\beta(\mathbf{x})$ is, by definition, non-dimensional and has a simple prior value of unity. This structure is also favorable to the conditioning of the inverse problem. Note that the inference of $\beta(\mathbf{x})$ is not physically restrictive. As an example, if $\beta(\mathbf{x})$ is a multiplier of the production term and the production is zero, the turbulence model is not relevant in the region.

Once $\beta(\mathbf{x})$ is introduced into the solution, the next step involves inference. Given data from a higher fidelity simulation or experiment, $\beta(\mathbf{x})$ can be estimated by solving an inverse problem as described in Sec. II A.

Solution of individual inverse problems can be considered to be a step towards the larger goal of improving predictive capabilities using a data–driven framework.¹⁶ The inferred function

$\beta(\mathbf{x})$ can be mapped to $\beta(\eta)$, where η may include quantities such as mean velocity gradients and turbulent time scales. By considering a sufficiently diverse set of inverse problems, the form of $\beta(\eta)$ can be generalized, for instance, using a machine learning algorithm. The inverse solution can, by itself, be used by a modeler to understand and correct for deficiencies in existing models.

A. Bayesian inversion

Inverse problems can be used to extract causal relationships from data. While inverse problem can be formulated in a deterministic setting, we are also interested in characterizing the impact of uncertainties in the data, model inadequacies and existing knowledge, on the inferred outputs. These statistical measures can be used to quantify the confidence on the inferred model, and by extension, can be further propagated in predictive modeling.

The Bayesian approach¹⁹ provides a formalism to quantify the posterior probability distribution $q(\beta|\mathbf{d})$, given a prior distribution ($p(\beta)$), data vector, \mathbf{d} , and a likelihood function, $h(\mathbf{d}|\beta)$. The Bayes theorem states that,

$$q(\beta|d) = \frac{h(\mathbf{d}|\beta)p(\beta)}{c}, \quad (3)$$

where $c = \int h(\mathbf{d}|\beta)p(\beta)d\beta$. In the current form, the solution to Equation (3) is intractable using a sampling based method because of the infinite dimensional nature of β . The problem is made finite-dimensional by re-defining it such that value of β is inferred at every point in the computational domain of the RANS grid.⁵²

In this work, the prior probability distribution $p(\beta)$ is assumed to be Gaussian. The assumption of a Gaussian distribution implies that β can take negative values, whereas in a turbulence model formulation, such as in Eq. (2), $P(\tilde{v}_t, \mathbf{U})$ is generally prescribed to be positive. Within the context of the present application, however, it is not necessary to restrict β to be positive as the intended correction is to the entire model (i.e., the right hand side of Eq. (2)) and not just to $P(\tilde{v}_t, \mathbf{U})$. In other words, the modification is not restricted to just one term, but rather to the entire model.

Turbulence models constitute complex non-linear systems, and thus the posterior distribution $q(\beta|d)$ and the likelihood $h(\mathbf{d}|\beta)$ cannot be expected to be Gaussian. Computing non-Gaussian posterior distributions, however, requires sampling-based methods, such as Markov chain Monte Carlo (MCMC).²⁰ These methods are extremely expensive in high-dimensional parameter estimation problems and are usually employed when the number of parameters is $O(1) - O(10)$. In the present work, since the spatially varying form of β is desired, the number of parameters to be estimated is equal to the number of grid points in the computational domain.⁵³ To promote tractability of the computation of the posterior properties, $q(\beta|d) \equiv e^{-(\beta-\beta_{prior})^T \mathbf{C}_{prior}^{-1} (\beta-\beta_{prior})}$ and $h(\mathbf{d}|\beta) \equiv e^{-\mathbf{F}^T \mathbf{C}_{obs}^{-1} \mathbf{F}}$ are approximated to be Gaussian. In these expressions, \mathbf{F} is a vector with i^{th} element and $f_i = d_{i,RANS} - d_{i,benchmark}$. d_i is the i^{th} data point and can be any flow dependent quantity such as pressure, skin-friction, and velocity. \mathbf{C}_{obs} and \mathbf{C}_{prior} are the observational and the prior covariance matrices. β_{prior} is the prior mean of the parameters (and also corresponds to the base model). Under Gaussian assumptions, the maximum a posteriori (MAP) estimate²¹ is taken to be the representative of the mean of distribution. The MAP can be computed by maximizing the numerator in Eq. (3), or equivalently by solving the following deterministic minimization problem:

$$\beta_{MAP} = \arg \min_{\beta} \mathfrak{J}(\beta) = \arg \min_{\beta} \frac{1}{2} \left[\mathbf{F}^T \mathbf{C}_{obs}^{-1} \mathbf{F} + (\beta - \beta_{prior})^T \mathbf{C}_{prior}^{-1} (\beta - \beta_{prior}) \right]. \quad (4)$$

The resulting optimization problem is still high-dimensional, but the posterior distribution is not determined by sampling methods and instead, the covariance is *approximated* by linearizing about the MAP point. This leads to a definition of the covariance in terms of the inverse of the Hessian of the objective function, $\mathfrak{J}(\beta)$,

$$\mathbf{C}_{posterior} = \left[\frac{d^2 \mathfrak{J}(\beta)}{d\beta d\beta} \right]_{\beta_{MAP}}^{-1}. \quad (5)$$

The error introduced by the Gaussian assumption is difficult to estimate without the use of sampling strategies, which are too expensive for the problems considered in this work. For a simpler problem setting, the appendix assesses the inaccuracy of the Gaussian approximation by comparing it with a sampling technique. Additional examples are provided in Ref. 15.

The posterior distribution β can be obtained by first performing a Cholesky decomposition of the form Ref. 54

$$\mathbf{R}^T \mathbf{R} = \mathbf{C}_{\text{posterior}}, \quad (6)$$

and generating realizations using

$$\boldsymbol{\beta} = \boldsymbol{\beta}_{\text{MAP}} + \mathbf{R}^T \mathbf{s}, \quad (7)$$

where \mathbf{s} is a vector of the same size as β and contains normally distributed random numbers with zero mean and unit standard deviation. The realizations of the posterior β are used to construct the posterior realizations of various flow quantities by solving the RANS equations for each realization of β . The prior realizations are constructed in a similar manner using $\mathbf{C}_{\text{prior}}$.

As mentioned previously, the posterior distribution is approximated by the inverse of the Hessian (Eq. (5)). In the examples presented in this work, computations of the exact Hessian are extremely expensive and ill-conditioned as the dimensions scale linearly with the number of the unknowns (in this case, the number of mesh points). Therefore, a Gauss-Newton approximation is employed, the details of which are presented in Appendix C.

B. Flow and adjoint solver

The unstructured compressible finite volume solver JOE^{22–24} is used to solve the RANS equations. A cell-centered, second order accurate spatial discretization is employed, with a backward Euler implicit scheme to march towards steady state. The inversion procedure requires gradients with respect to every grid point, thus necessitating a discrete adjoint²⁵ procedure (see Appendix B). The partial derivatives required in the adjoint procedure are computed efficiently using automatic differentiation via the ADOL-C framework.²⁶

Secs. III–VI demonstrate the application of the inference in a number of turbulence modeling situations.

III. TURBULENT CHANNEL FLOW

The first case to be considered is a fully developed planar turbulent channel flow (Fig. 1). This canonical problem has been heavily utilized to calibrate model parameters throughout the history of turbulence modeling, although most models do not account for the variability due to Reynolds number. The question of choosing the right combination of length and time scales is still an outstanding one in modeling. Rich Direct Numerical Simulation (DNS) datasets^{27–29} exist for such flows. These datasets are, in our opinion, under-utilized for modeling purposes. In the current work, the functional form of discrepancies in the Wilcox $k-\omega$ model³⁰ is extracted using DNS data for a wide range of Reynolds numbers.

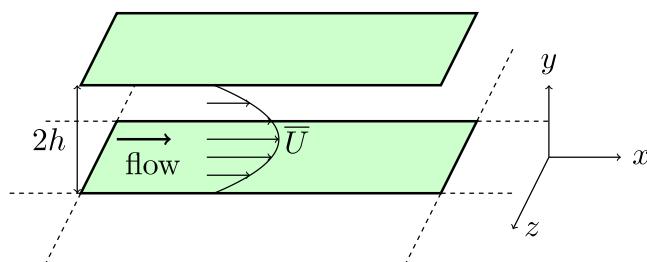


FIG. 1. Schematic of a fully developed planar channel flow.

TABLE I. Wilcox $k-\omega$ model constants.

α^*	σ^*	γ	α	σ
0.09	0.6	13/25	0.09	0.5

A. Bayesian inversion setup

The Reynolds-averaged momentum equation for a fully developed, incompressible and steady channel flow is given by,

$$\frac{\partial}{\partial y} \left[\mu \frac{\partial u}{\partial y} - \rho \bar{u}' \bar{v}' \right] - \frac{\partial p}{\partial x} = 0, \quad (8)$$

where, u represents the mean streamwise velocity and the Reynolds stress $-\bar{u}' \bar{v}'$ is modelled as $\nu_t \frac{\partial u}{\partial y}$ in eddy viscosity models. In the Wilcox $k-\omega$ model, $\nu_t = \frac{k}{\omega}$, where k and ω satisfy the following transport equations:

$$\nu_t \left(\frac{\partial U}{\partial y} \right)^2 - \alpha^* k \omega + \frac{\partial}{\partial y} \left[\left(\nu + \sigma^* \frac{k}{\omega} \right) \frac{\partial k}{\partial y} \right] = 0, \quad (9)$$

$$\gamma \left(\frac{\partial U}{\partial y} \right)^2 - \alpha \omega^2 + \frac{\partial}{\partial y} \left[\left(\nu + \sigma^* \frac{k}{\omega} \right) \frac{\partial \omega}{\partial y} \right] = 0. \quad (10)$$

Table I contains the values of the model constants. In this work, a correction function $\beta(\mathbf{y})$ is introduced as a multiplier to the production term in the ω equation (Equation (10)). The system is solved on a geometrically graded mesh consisting of 201 points with the first grid point placed well into the viscous sublayer at $y^+ \approx 0.05$. Inversion is performed for a set of five different Reynolds numbers, $Re_\tau \in [180, 550, 950, 2000, 4200]$. The ensemble averaged velocity—at every grid point—from the DNS solution²⁸ is used as the benchmark data. A diagonal observational covariance with standard deviation, σ_{obs} , of 10^{-8} is used. The very low value represents a high degree of confidence on the DNS data. A diagonal prior covariance with $\sigma_{prior} = 0.5$ is used, which corresponds to a nearly uninformative prior. The MAP solution is obtained as described in Sec. II A. The posterior and prior distributions are constructed using 10 000 realizations.

B. Results

Figure 2 shows the prior, posterior, and DNS values for flow properties at $Re_\tau = 550$. The posterior velocity profile is almost identical to the DNS profile with the standard deviation collapsing to a very low value. The accuracy of the posterior can be further verified from Figure 2(d), which shows the variation of $y^+ du^+/dy^+$, and further highlights the differences. As turbulent production is negligible for $y^+ < 1$, the posterior solution is not affected by $\beta(\mathbf{y})$ and hence $\beta(\mathbf{y})$ remains same as the prior. Correspondingly, there is also no reduction in uncertainty in the viscous sublayer for $y^+ < 2$. Outside of the sublayer, $\beta(\mathbf{y})$ changes to match the specified benchmark velocity profile and a drastic reduction in the posterior standard deviation is noticeable.

Figure 3 shows the inferred function, β , for all Reynolds numbers. In every case, β in the overlap and the log-layer scales with wall units. In the outer region, β scales with wall height. This is a physically reasonable variation and represents the systematic nature of modeling deficiency. The large variation in β around $y^+ = 25$ shows the inability of the underlying model to provide the correct time scale near the interface of the overlap region and the log-layer. It is particularly revealing that the nature of the required correction is almost universal (with second-order viscous corrections) for $y^+ < 50$, a fact that can be used to develop improved models.

IV. TURBULENT FLOW WITH CURVATURE

Turbulent flows over curved surfaces are subjected to streamwise pressure gradients and additional mean strain. Curvature acts to stabilize (or destabilize) the flow via the diverging (or

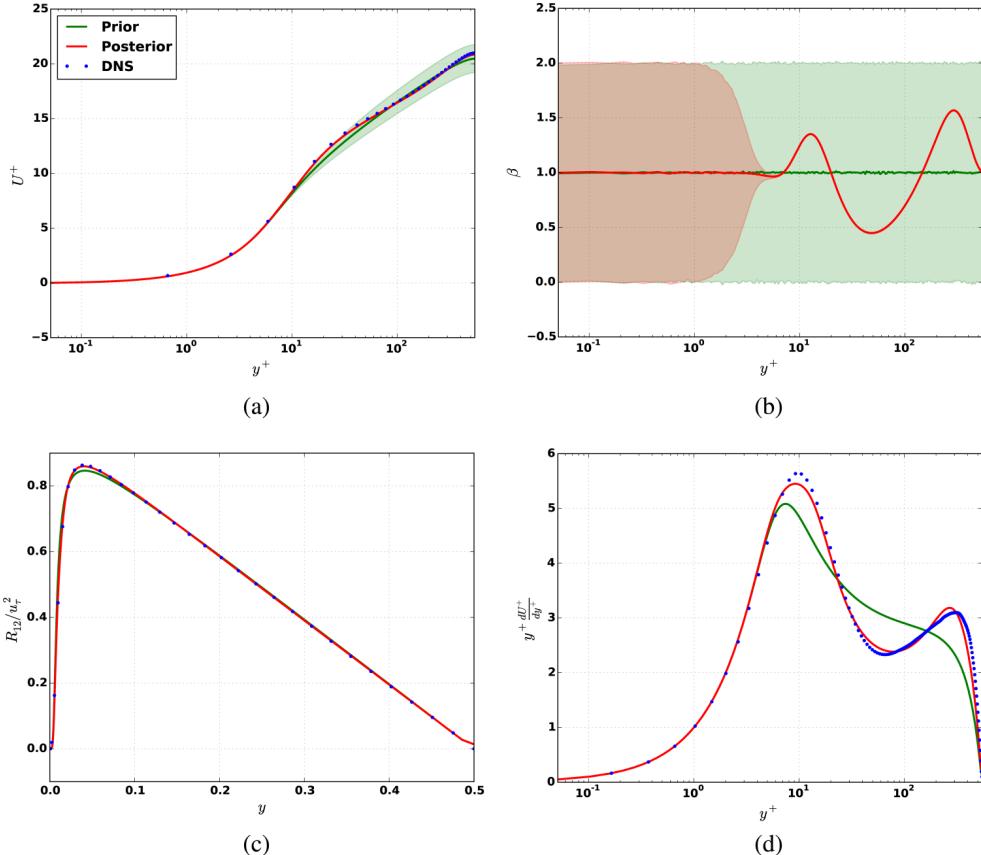


FIG. 2. The prior, the posterior and the DNS solution for $Re_\tau = 550$. The prior solution is represented in green, red represents the posterior and blue represents the DNS solution. (a) Non-dimensionalized velocity, U^+ . Mean is shown with the solid line; shaded region represents the 95% confidence interval. (b) Correction function, β . Mean is shown with the solid line; shaded region represents the 95% confidence interval. (c) Non-dimensionalised Reynolds stress. (d) $y^+ \frac{dU^+}{dy^+}$.

impinging) inviscid core.^{31–33} The flow field is complicated by the presence of complex small-scale vortical structures.³⁴ Linear eddy viscosity closures have to be explicitly sensitized to capture curvature effects.^{35–37} In this section, the inversion procedure is applied to boundary layer flows subject to (a) convex and (b) concave curvature. A comparison of the posterior solution with the analytically sensitized rotational correction (SARC) model of Spalart³⁸ is also presented.

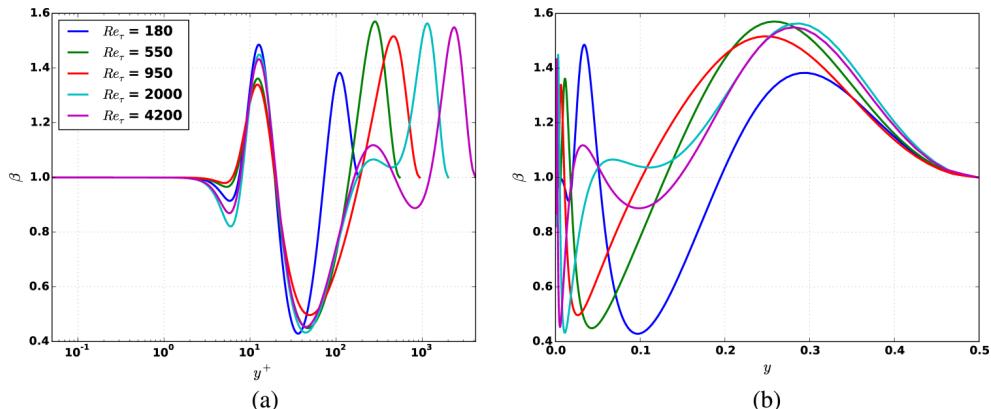


FIG. 3. The inferred correction function, β_{MAP} , for $Re_\tau \in [180, 550, 950, 2000, 4200]$.

A. Flow conditions

The benchmark data are obtained from a large eddy simulation.^{39,40} In the LES, an inlet boundary layer with Re_θ of 2000 is generated using the recycling-rescaling procedure.³⁹ The inlet boundary layer thickness, δ_0 , is 0.04 grid units and the inlet size is $4\delta_0$. The recycling and outflow sections were of length $5\delta_0$ and $7\delta_0$, respectively. The LES solution is performed over 13 flow-through times and statistics are computed from the last 9 flow-through times. The grid inlet in the RANS is extended to minimize the effects of the inlet profile, such that $Re_\theta = 1500$ at the new grid entrance and it ensures $Re_\theta = 2000$ near the original grid entrance. The extended mesh consists of 473 points in the streamwise direction and 111 points in the wall-normal direction. The upper wall is assigned an inviscid slip boundary condition while the lower wall is given an adiabatic no-slip boundary condition. The inlet velocity, density, and eddy viscosity profiles are obtained from a precursor zero pressure gradient flat plate simulation. The pressure at the inlet is extrapolated from the interior. At the outlet, pressure is interpolated from the exterior and the rest of the quantities are extrapolated.

B. Bayesian inversion setup

A correction function $\beta(x, y)$ is multiplied to the production term of the Spalart–Allmaras (SA) model in the form

$$\frac{D\tilde{v}_t}{Dt} = \beta(x, y)P(\tilde{v}_t, \mathbf{U}) - D(\tilde{v}_t, \mathbf{U}) + T(\tilde{v}_t, \mathbf{U}). \quad (11)$$

The coefficient of friction (C_f) at the lower wall, from the LES solution, is used as the benchmark data for inversion. The variation in pressure for this case is small, therefore comparing pressure from our compressible RANS to the incompressible LES is prone to error. Hence, pressure-related quantities are not used for benchmarking. A diagonal prior covariance with a standard deviation, $\sigma_{prior} = 1$ is used. A diagonal observation covariance with standard deviation, σ_{obs} , of 10^{-6} is used. This corresponds to a standard deviation of around 0.03% on LES C_f , and represents a high degree of confidence in the LES. The posterior covariance and realizations are computed as detailed in Section II A. The flow solver is evaluated for 4000 realizations of β to obtain both the prior and the posterior realizations.

C. Results

Figure 4 shows the prior and the posterior C_f with 95% confidence intervals alongside LES and SARC results for the convex channel. The LES data points lie within the bounds or close to the

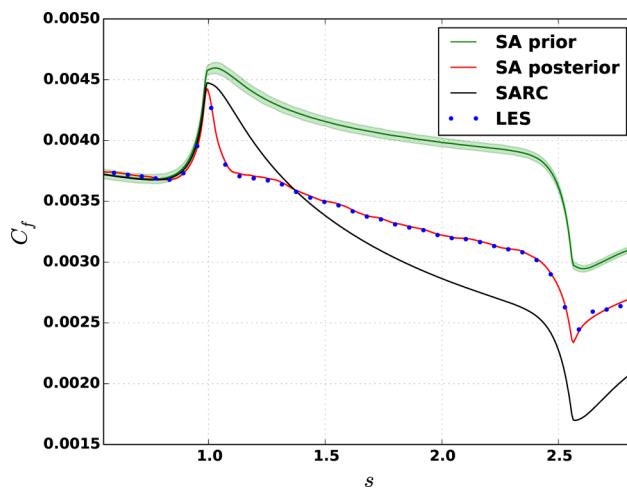


FIG. 4. The prior, posterior, SARC, and LES coefficient of friction (C_f) at the lower wall for the convex channel. Shaded region represents the 95% confidence interval.

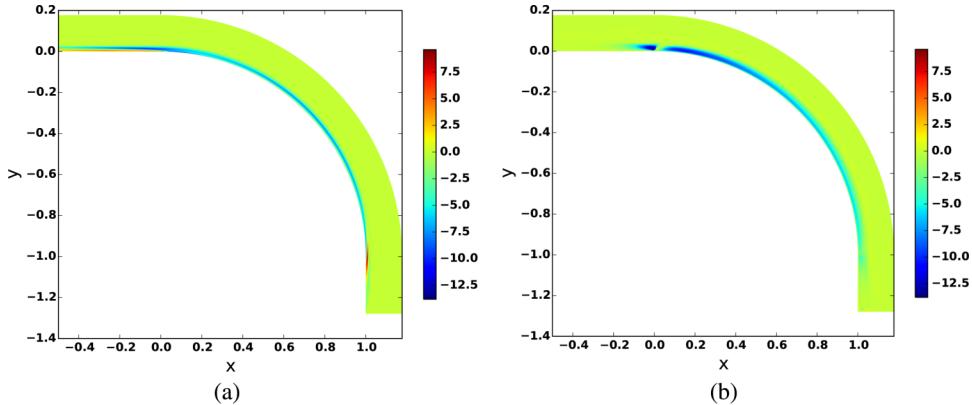


FIG. 5. Contour plot showing the inferred correction term (δ_{MAP}) and the SARC correction term for the convex channel. (a) Inferred Correction Term(δ_{MAP}). (b) SARC.

posterior C_f , while the SARC model is not very accurate for this case. The confidence level in the posterior solution is seen to be very high. Figure 5 shows the inferred MAP correction term, δ_{MAP} , and the analytically defined correction term from the corresponding SARC model. Qualitatively, the correction term is similar for both models, but the magnitudes are locally different. The trend in the correction is consistent with the expectation that the convex curvature reduces the turbulence intensity. The quantitative nature of the results should provide valuable information to the modeler.

Figure 6 shows the variation of the streamwise velocity with respect to the distance from the wall at various streamwise locations. The posterior velocity is seen to correlate well with the LES solution compared to the prior and the SARC prediction. The results suggest that the SARC model requires improvements in the log layer.

Figure 7 shows the skin-friction predictions for a channel with concave curvature. The posterior C_f is close to the LES solution except in a small region downstream. It is to be noted that linear eddy viscosity-based models may be deficient in their ability to perform better. The correction term in Figure 8 is consistent with the increase in turbulent production in a concave boundary layer.

Figure 9 shows the streamwise velocity profiles. Clearly, the posterior velocity prediction is closer to the LES velocity profile compared to the prior. The SARC model appears to perform reasonably well in this case.

As observed in both the concave and convex cases, while the posterior C_f matches the LES extremely well, the velocity profiles show some degree of variability. This is because C_f is a local quantity and a small local change in the velocity profile may be enough to match C_f . Nevertheless, the improvement in the posterior velocity profiles adds more credibility to the inversion procedure.

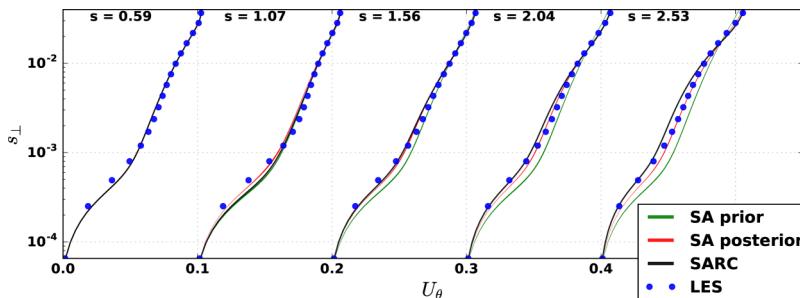


FIG. 6. The prior, posterior, SARC, and LES stream-wise velocity at various locations for the convex channel. s_\perp refers to the perpendicular distance from the lower wall.

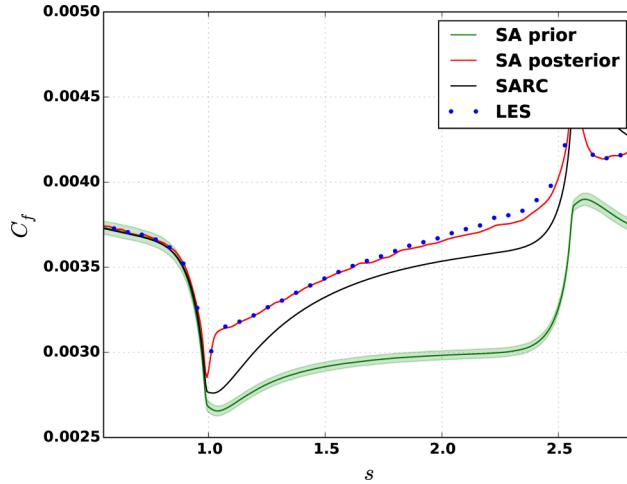


FIG. 7. The prior, posterior, SARC, and LES coefficient of friction (C_f) at the lower wall for the concave channel. Shaded region represents the 95% confidence interval.

V. OBLIQUE SHOCK TURBULENT BOUNDARY LAYER INTERACTION (OSTBLI)

The next test problem considers the interaction of an oblique shock with a spatially evolving turbulent boundary layer. Shock boundary layer interactions have been studied extensively^{41–43} over the years. Knight and Degrez⁴⁴ present an extensive survey for this case and found RANS to be inaccurate, except for very weak interactions. This behavior is attributed to the inability of RANS models to respond to a sharp pressure rise and the use of a single turbulent scale in an inherently multi-scale environment. We use the Bayesian formulation to infer an improved time scale estimate for use within a $k-\omega$ model solution.

The flow conditions attempt to emulate the benchmark setup in the work of Morgan *et al.*,⁴⁵ which employs a well-resolved Large Eddy Simulation. The incoming boundary layer has a free-stream Mach number of 2.28, an inlet momentum thickness Reynolds number of 1500, and a thickness of 0.87 grid units. A grid of extent 10 units in the streamwise direction and 15 units in the wall-normal direction is discretized using a cartesian mesh of 201 uniform points in the streamwise direction and 151 points in the wall-normal direction. The undisturbed shock is aligned at an angle of 8° to the freestream and its reflection point is at the center point of the wall domain. The inlet boundary layer is generated using a precursor flat plate simulation. Rankine-Hugoniot conditions are used to obtain the shock jump conditions, which are then forced at the inlet plane.

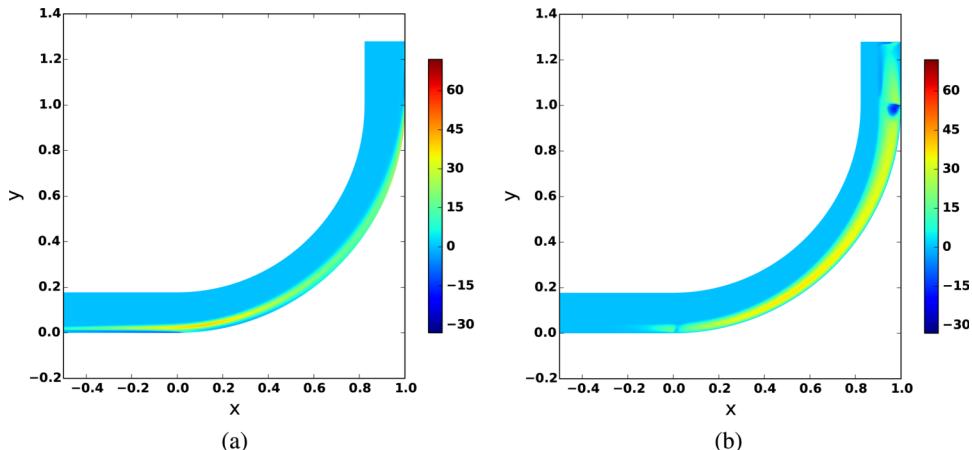


FIG. 8. Contour plot showing the inferred correction term (δ_{MAP}) and the SARC correction term for the concave channel. (a) Inferred Correction Term(δ_{MAP}). (b) SARC Correction Term.

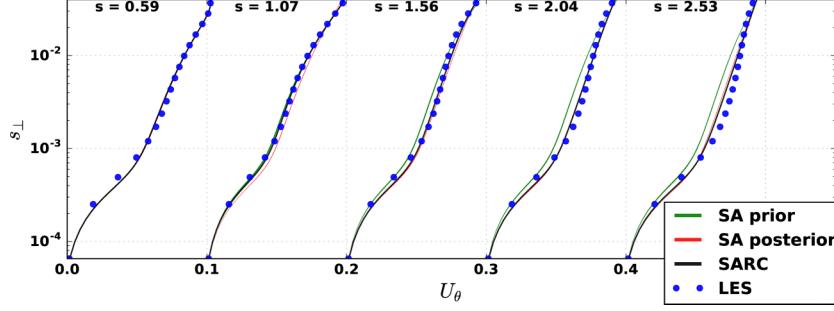


FIG. 9. The prior, posterior, SARC, and LES wall parallel velocity (U_θ) at various stream wise locations for the concave channel. s_\perp refers to the perpendicular distance from the lower wall.

Similar to the previous case, a correction function $\beta(\mathbf{x}, \mathbf{y})$ is multiplied to the production term of the ω equation. The coefficient of pressure (C_p) at the lower wall, from the LES solution,⁴⁵ is used as the benchmark data for inversion. A diagonal prior covariance with a standard deviation, $\sigma_{prior} = 1$ is used. This corresponds to a nearly uninformative prior. The selection of the prior can indeed be guided by additional insights gained during the inversion process. A diagonal observation covariance with standard deviation, σ_{obs} , of 10^{-6} is used. The posterior covariance and realizations are computed as detailed in Section II A. The flow solver is evaluated for 1000 realizations of β to obtain both the prior and the posterior realizations.

A. Results

Fig. 10 compares the pressure and friction coefficients using the prior, the posterior model, and the LES. The posterior C_p tracks the LES consistently. As evident in the C_f plot, location, and length of the separation bubble is more accurately predicted with the posterior model. Fig. 11 shows the X-velocity contours for the prior and the posterior model. The posterior velocity field aligns well with the superimposed LES velocity. This showcases the ability of the inverse approach to predict the complete flow field using very limited information on the viscous wall. Figure 12 compares the prior and the posterior velocity correlations and the turbulent kinetic energy (TKE) at different streamwise locations. The posterior model, especially downstream of the interaction region is seen to match the LES results extremely well. The improvement in downstream TKE is especially of interest. The fact that the TKE, velocity field, pressure, density, and Reynolds stress correlate well with the LES shows that the improvements are physically realizable.

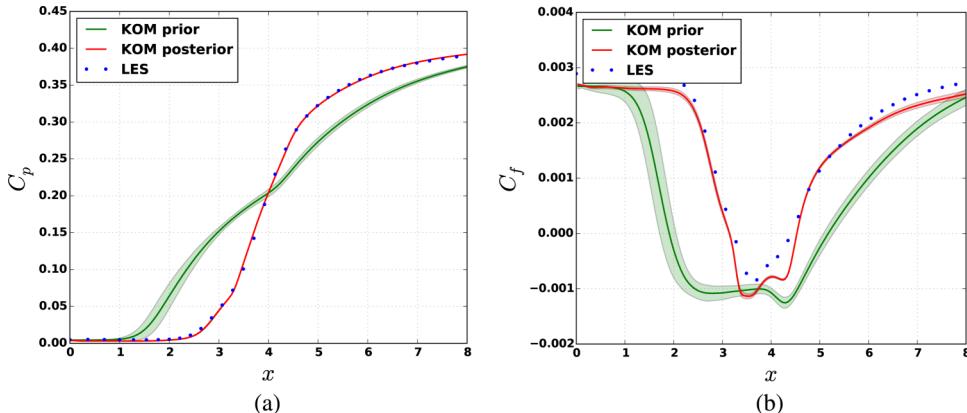


FIG. 10. Pressure and friction coefficient for the LES, prior, and the posterior $k-\omega$ model. Shaded region represents the 95% confidence interval. (a) C_p . (b) C_f .

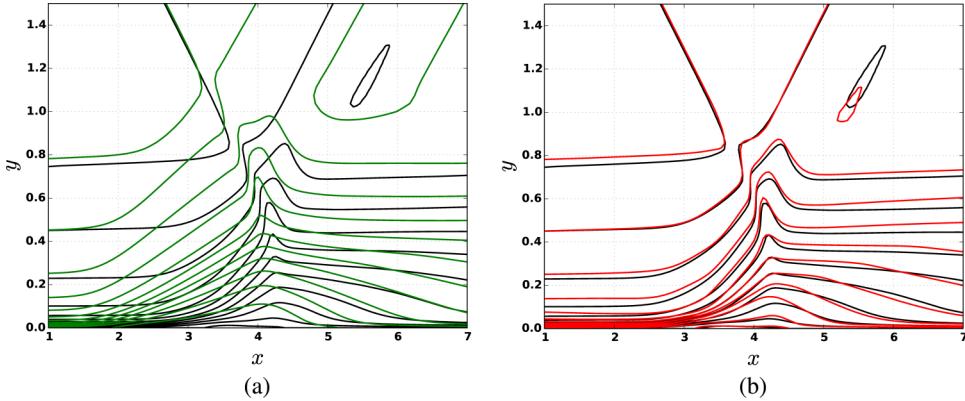


FIG. 11. X-velocity contours near the interaction region. The contours are superimposed with the LES velocity (black) for comparison. The full flow field is well predicted with the inferred MAP $k-\omega$ model. (a) Prior $k-\omega$. (b) Posterior $k-\omega$.

VI. HIGH ANGLE OF ATTACK FLOW OVER S809 AIRFOIL

The next case involves the S809 airfoil, which is extensively used in the blades of horizontal axis wind turbines (HAWTs). At a Reynolds number of 2×10^6 , Somers⁴⁶ observed a laminar separation bubble near mid-chord for angles of attack α between 0° and 5° , which leads to transition to turbulence. A further increase in the angle of attack moves the transition point towards the leading edge. At $\alpha = 9.22^\circ$, the transition point is close to the leading edge, and turbulent separation occurs near the trailing edge. Beyond this point, the flow separates near mid-chord. At the design Reynolds number, RANS based models are fairly accurate for flows at low angles of attack ($\leq 9^\circ$). Beyond this regime, the presence of a large region of turbulent separation with an adverse pressure gradient renders RANS inaccurate. Previous work^{47–49} indicates consistent failure of the RANS models beyond the point of separation. In this work, Bayesian inversion is applied to quantify discrepancies in the one equation Spalart–Allmaras (SA) model by utilizing the experimentally acquired pressure data.

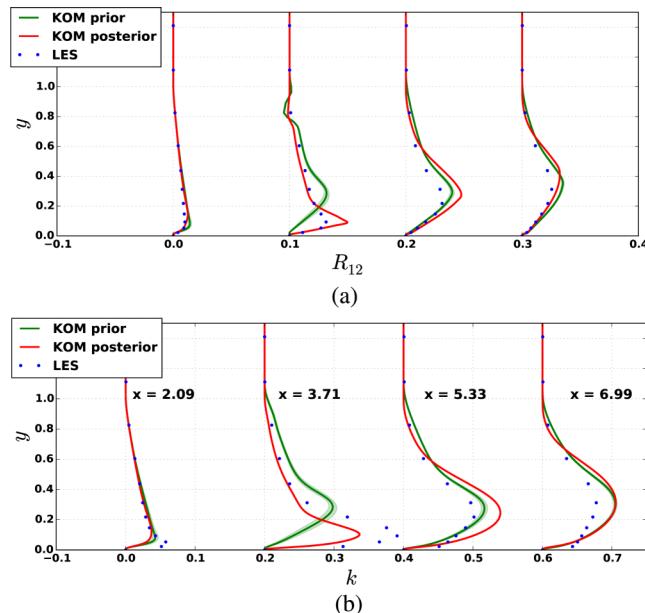


FIG. 12. Reynolds shear stress and turbulent kinetic energy (normalized by freestream velocity) at various streamwise locations. (a) Reynolds shear stress. (b) Turbulent kinetic energy (TKE).

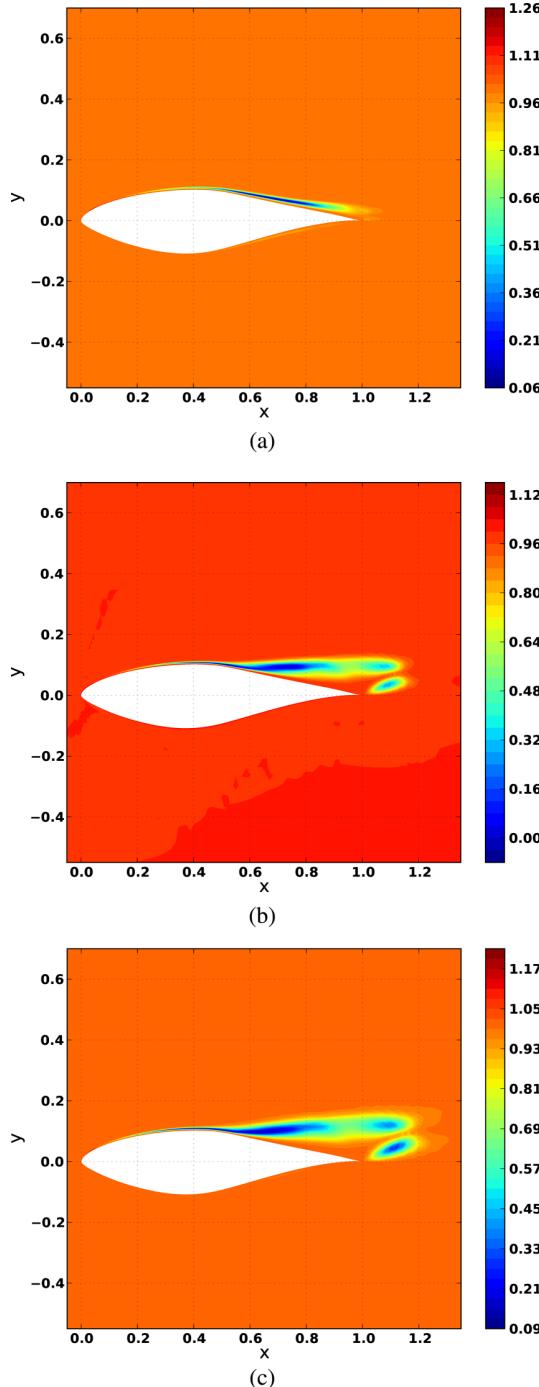


FIG. 13. The inferred correction multiplier, $\beta(\mathbf{x}, \mathbf{y})_{MAP}$. (a) $\alpha = 9.22^\circ$. (b) $\alpha = 12.20^\circ$. (c) $\alpha = 14.24^\circ$.

A. Flow conditions

A Mach number of 0.2 and a Reynolds number of 2×10^6 is used to match the experimental Reynolds number by Somers.⁴⁶ A C-grid with 291 points in the wrap-around direction and 111 points in the wall perpendicular direction is used. This corresponds to 200 grid points on the airfoil surface. The far-field boundaries are located 35 chord lengths from the airfoil surface. Characteristic based freestream boundary conditions are used for the flow variables at the far-field and the eddy viscosity is set to the fully turbulent value, $\nu_{t,\infty}/\nu_\infty = 3$.

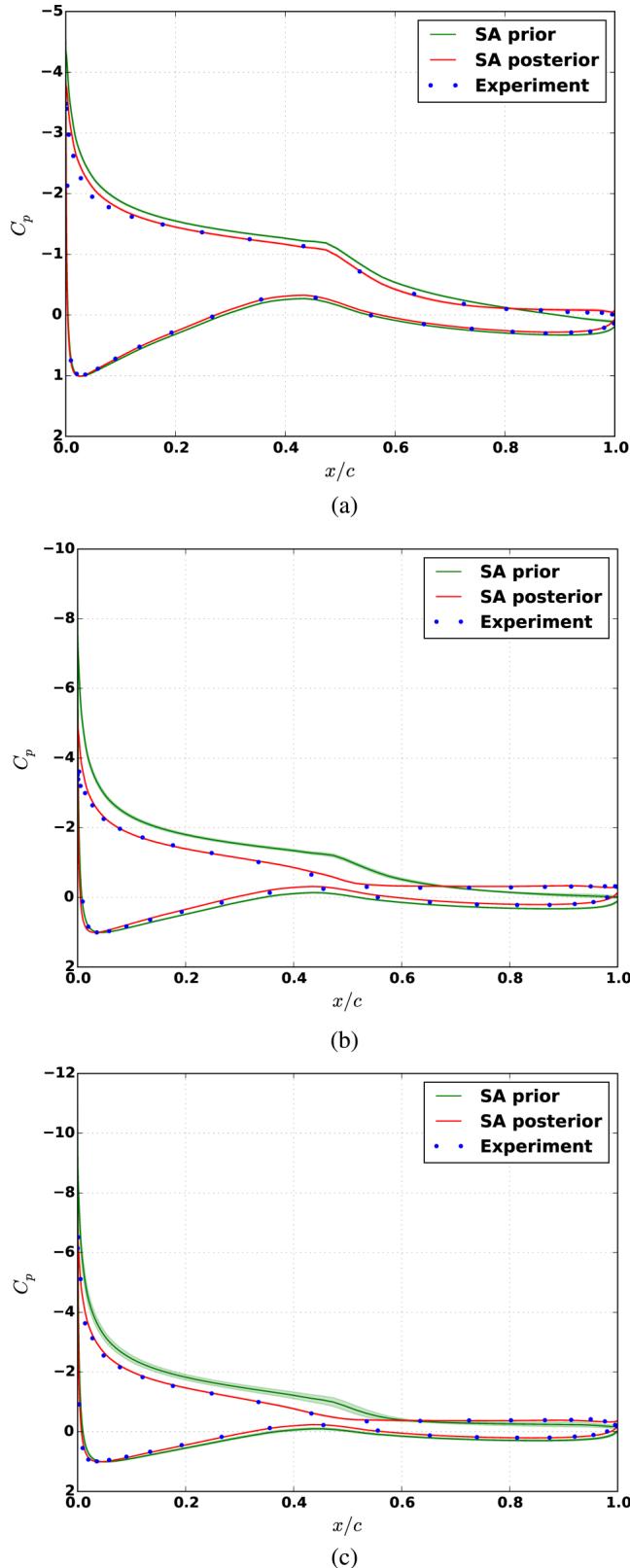


FIG. 14. The prior and the posterior coefficient of pressure (C_p) along with the experimental benchmark. Shaded region represents the 95% confidence interval. Error bars for the experimental results are not provided in the original document.⁴⁶
 (a) $\alpha = 9.22^\circ$. (b) $\alpha = 12.20^\circ$. (c) $\alpha = 14.24^\circ$.

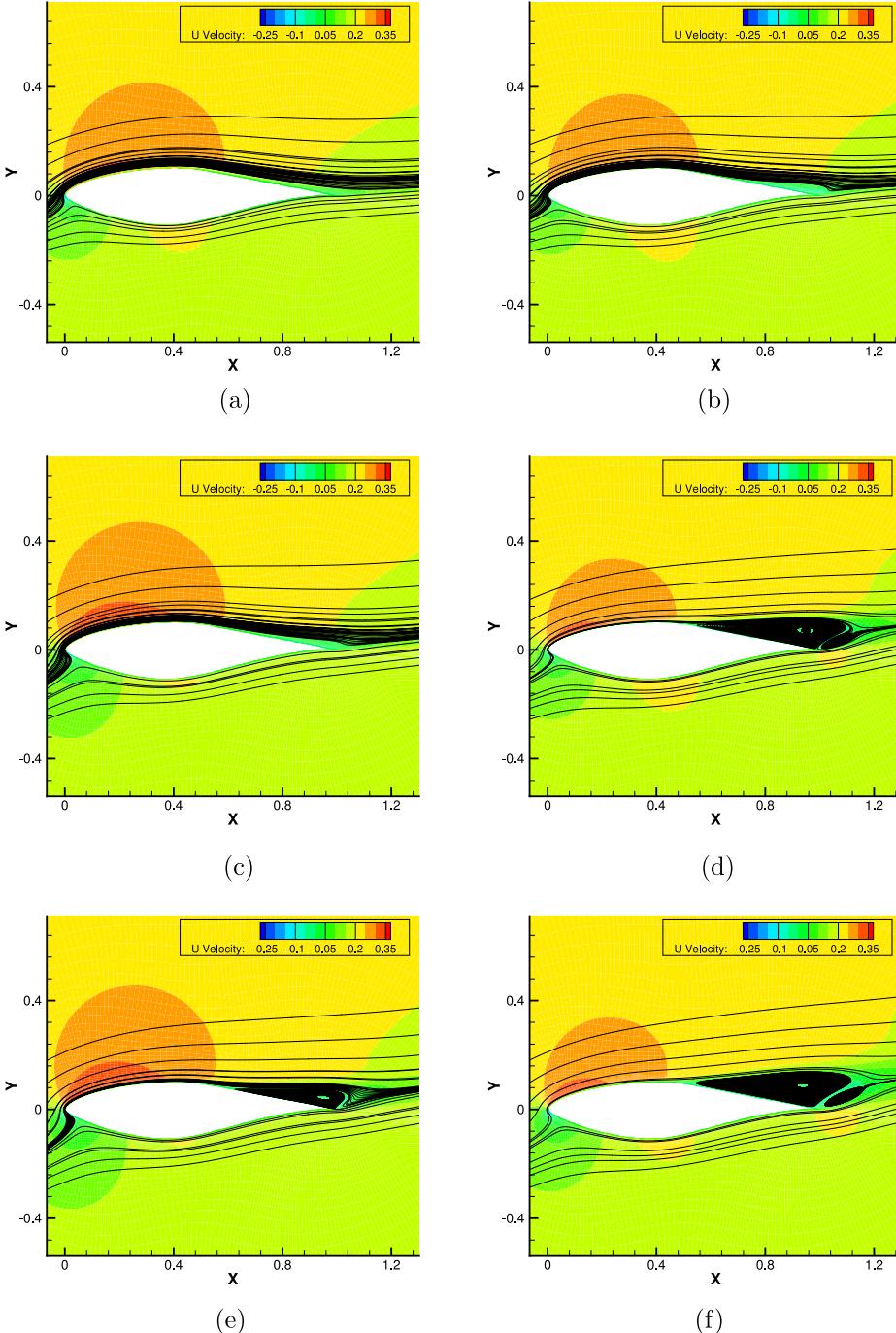


FIG. 15. X-velocity contour along with the streamlines near the surface of the airfoil. (a) Base SA, $\alpha = 9.22^\circ$. (b) MAP SA, $\alpha = 9.22^\circ$. (c) Base SA, $\alpha = 12.20^\circ$. (d) MAP SA, $\alpha = 12.20^\circ$. (e) Base SA, $\alpha = 14.24^\circ$. (f) MAP SA, $\alpha = 14.24^\circ$.

B. Bayesian inversion setup

As in the case of curved channel flow, a correction function $\beta(x, y)$ is multiplied to the production term. Experimental coefficient of pressure (C_p) is used as the benchmark for inversion. The experimental data are extracted from the report by Somers.⁴⁶ The data extraction is prone to error, especially near the leading edge, where the pressure changes rapidly. The effect of this error is minimized by utilizing *only the suction-side pressure taps, starting from $0.05c$ to $0.8c$* , for the inversion.

TABLE II. Comparison of prior, posterior, and experimental lift coefficients for $\alpha = 9.22^\circ, 12.2^\circ, 14.24^\circ$.

	$\alpha = 9.22^\circ$	$\alpha = 12.2^\circ$	$\alpha = 14.24^\circ$
Experiment	1.010	1.007	1.083
Prior SA (% error in mean)	$1.166 \pm 0.005(15\%)$	$1.411 \pm 0.017(40\%)$	$1.453 \pm 0.039(34\%)$
Posterior SA (% error in mean)	$1.051 \pm 0.005(4\%)$	$1.028 \pm 0.004(2\%)$	$1.140 \pm 0.004(5\%)$

A diagonal prior with a standard deviation, $\sigma_{prior} = 1.0$ and a diagonal observational covariance with standard deviation, $\sigma_{obs} = 10^{-2}$ are used. This corresponds to a maximum error of 0.2% in the C_p . Note that σ_{obs} should ideally be derived from the experimental data. However the experimental error is unavailable for this case, and the posterior results were confirmed to be insensitive to small changes in the prior covariance. The optimization process is identical to the previous cases. The prior and the posterior model are constructed using 1000 realizations of β .

C. Results

Figure 13 shows the inferred field β_{MAP} for all the cases. Turbulent production is reduced on the suction side. The tendency of the inferred field is to significantly reduce the production of eddy viscosity starting in an extremely thin region close to the leading edge and continuing into the edge of the boundary layer and into the separated shear layer. This reduction is consistent with the transitioning and separating boundary layer and may be used to construct a transition model and a model for adverse pressure gradients. Figure 14 shows that the experimental C_p is well-matched by the inversion, and the posterior variance in C_p is reduced significantly. Though the results are not shown here, it was confirmed that solely by using the pressure data at the three most upstream locations on the suction side, the posterior pressure distribution could be improved to the level in Figure 14 over the complete airfoil surface.

Figure 15 compares the near-wall streamlines for the base SA model and the MAP, showing the large differences in the prior and the posterior separation bubble. For $\alpha = 12.20^\circ$ and 14.24° , the posterior model predicts the separation at $0.50c \pm 0.01c$ (which represents a 68% confidence interval) and $0.48c$, respectively, compared to the prior predictions of $0.73c \pm 0.04c$ and $0.56c \pm 0.02c$, respectively. The predictions are consistent with the experiments, where the separation is observed near the mid-chord for these angles. For $\alpha = 9.22^\circ$, turbulent separation is observed in the experiment but not in the prior SA model. The posterior, however, predicts the separation at $0.71c \pm 0.01c$. Table II compares the prior and the posterior lift coefficient (C_l) with the experiment. For $\alpha = 14.24^\circ$, the error in the C_l is reduced from 34% to 5% and the trend is preserved for other angles of attack.

VII. SUMMARY AND PERSPECTIVES

A data-based approach was presented with the objective of quantifying functional errors and uncertainties in turbulence closure models. Bayesian inversion techniques were adapted to extract field variables (source terms) that are consistent with their use in closure equations. The problem is made finite-dimensional by re-defining it such that the distribution of the unknown functions is sought at every point in the computational domain of the RANS grid. With the additional assumption of a Gaussian distribution, the inversion is reduced to the solution of a deterministic optimization problem to compute the maximum a posteriori estimate and sampling to obtain the posterior variabilities.

A number of problems are studied, including one-dimensional channel flows, boundary layers with curvature, airfoil flows, and oblique shock-turbulent boundary layer interactions. In all of these cases, given a few data points (typically wall-pressure or skin-friction), complete realizations of the closure model are obtained. For the one-dimensional channel flow, the posterior

model suggests a universal time scale correction in the near wall region. For flows with curvature, velocity profiles are improved in the complete flow domain by virtue of assimilating just the skin-friction at the wall. The required corrections in the turbulence production term were compared to an existing curvature-correction procedure. The entire flow field is well-predicted in the oblique shock-turbulent boundary layer interaction problem by matching the surface pressure. The point of separation and lift is more accurately predicted for the S809 airfoil by matching the experimental surface pressure at a few points on the suction side.

While the examples presented in this work infer the form of source terms in closure equations, a similar procedure can be used to extract, for instance, a field distribution of a Prandtl or a Schmidt number. Nevertheless, it is pertinent to recognize that inferring a field variable changes the balance of the entire model in a non-local sense.

The posterior information is, of course, sensitive to the prior assumptions to a certain degree. For all the examples in this work, the mean of the prior distribution was chosen to be same as the baseline model. This is a reasonable choice to make, but within the setting of model development, the knowledge of the prior will improve with increasing datasets. Similarly, once a number of inverse problems have been computed, the variability in the posterior distributions can provide some additional information about the prior covariance for future computations.

Having extracted pertinent data, it remains to use it to improve models. As a first step, a larger, and more diverse set of problems will have to be considered. In this process, care has to be devoted to the choice of the objective function. For instance, the coefficient of pressure is found to be a more *global* benchmark when compared to the friction and leads to a model that is closer to reality. Friction, a highly localized quantity, may lead to a degree of over-fitting by adjusting the model in a small near-wall region in many problems. *The inferred functions can be used to assess the structural form of the model and make improvements.* As an alternate step, machine learning algorithms^{14,15,50} may also be directly used to reconstruct model terms.

Based on the posterior results in this paper, it could also be argued that some of the criticism of eddy viscosity models that is usually attributed to the isotropic stress-strain relationship is partially a result of poor functional forms within the model rather than the limiting nature of the Boussinesq assumption itself. Models should be assessed by evaluating the type and variability of corrections that will be required in a diverse set of problems. Another indicator of the robustness of a model structure and its improvement via field inversion is to inspect the posterior distributions of variables other than those which were used in the inversion.

While inverse models have a tendency of over-fitting a model to a given dataset, the specification of an appropriate prior distribution can mitigate this effect. As the diversity of data is increased, the choice of the prior will become less arbitrary.

We note that additional work may be required to assess the role of numerical errors,⁵⁵ sensitivity to error in the data, impact of the Gaussian assumption, etc. Efficiencies can be gained by decomposing the inferred functions in terms of a priori bases and estimating coefficients of basis functions. Current work is being directed towards applying these techniques to second moment closures.

ACKNOWLEDGMENTS

This work was supported by NASA LEARN Project No. NNX15AN98A (Technical Monitor: Gary Coleman) and by the NSF via Grant No. 1507928 (Program Officer: Dimitrios Papavassiliou).

APPENDIX A: MARKOV CHAIN MONTE CARLO SAMPLING FOR THE PLANAR CHANNEL FLOW

As mentioned in Section II, the maximum a posteriori (MAP) estimate and an approximate covariance around the MAP solution is used to characterize the properties of the posterior distribution. This assumption is used to promote tractability of the inversion process. It is to be noted that even in the simplest case presented in this paper, 201 values of β are estimated and hence

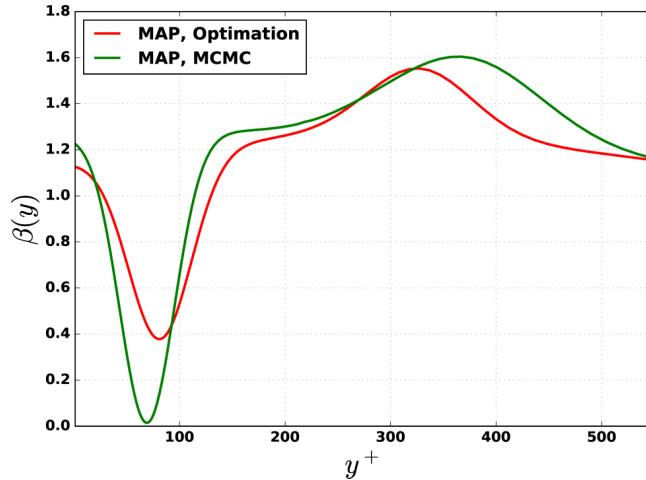


FIG. 16. Comparison of the MAP $\beta(y)$ field obtained by solving the deterministic optimization problem (and approximate Gaussian posterior) with a generic MCMC method.

sampling methods become prohibitively expensive. This section assesses the implications of pursuing sampling-based Markov-chain Monte Carlo-based (MCMC²⁰) methods by considering a further simplified example.

The channel flow example from Section III is considered with $Re_\tau = 550$. The prior and the observational distributions are still assumed to be Gaussian with diagonal covariances. To facilitate the application of sampling methods, the number of parameters to be estimated is reduced to 15. This is accomplished by reconstructing $\beta(y)$ using a set of Radial basis functions (RBFs) in contrast to assigning β to individual cells in the finite volume framework. For the results presented here,

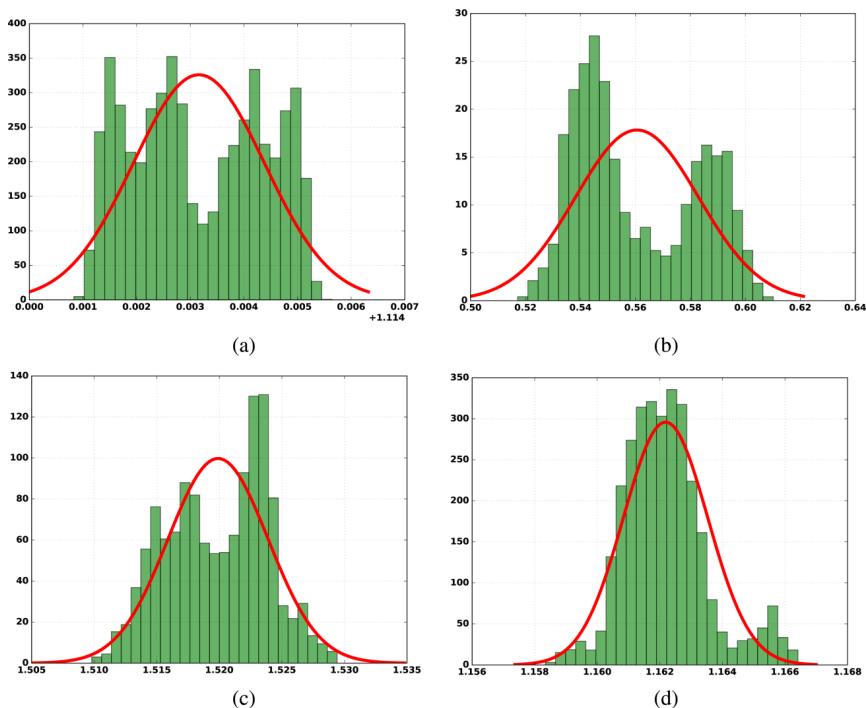


FIG. 17. Probability distribution function of the inferred function $\beta(y)$ at different locations. Red line represents an approximate Gaussian distribution. (a) $y^+ = 11$. (b) $y^+ = 99$. (c) $y^+ = 308$. (d) $y^+ = 550$.

five radial basis functions are used, each containing three parameters: the weight(c_j), location of the RBF center(y_j) and the radius of influence(r_j). Given the RBF parameters, the $\beta(y)$ field can be reconstructed as follows:

$$\beta(y) = \sum_{j=1,5} c_j \exp \left[-\frac{(y - y_j)^2}{r_j^2} \right]. \quad (\text{A1})$$

The initial values of these parameters were estimated by utilizing the full field $\beta_{MAP}(y)$ solution from Section III. The prior distribution for the RBF parameters is taken to be Gaussian with standard deviations $\sigma_{c_j} = 0.1$, $\sigma_{y_j} = 0.02$, and $\sigma_{r_j} = 0.03$. The observational standard deviation, σ_{obs} , is taken to be 10^{-3} . MCMC sampling is performed using a Python package PyMC3⁵¹ which utilizes a Metropolis-Hastings algorithm to determine the posterior distribution. A total of 5 million samples were required for convergence. Realizations of $\beta(y)$ are constructed a posteriori using the samples of RBF parameters. Figure 16 compares the $\beta(y)_{MAP}$ from MCMC with $\beta(y)_{MAP}$ using the Gaussian assumption as described in Section II. Further, the posterior probability distribution at four different locations are shown in Figure 17 along with an approximate Gaussian probability density function. It is evident that in the near-wall region, the posterior PDF has a bi-modal character and transitions to a near-Gaussian PDF in the outer region.

It is to be emphasized that even after the reduction in parameters from 201 to 15, over 5 million forward solutions were required to obtain a sufficiently converged MCMC solution.

APPENDIX B: GRADIENT CALCULATION

The derivative of the objective function with respect to β is required for the optimization process. In the discretized framework, the size of β is the same as the number of finite volume cells, which can easily run into thousands, rendering finite difference extremely expensive. Therefore, the adjoint formulation is used to evaluate the derivatives efficiently. The derivative of the prior part of the objection function is trivial, hence we just focus on the derivative of the observation part of the objective function, which is given by

$$\mathfrak{J}_{obs}(\beta) = \frac{1}{2\sigma_{obs}^2} \mathbf{F}^T \mathbf{F} = \frac{1}{2\sigma_{obs}^2} \left(\sum_i f_i^2 \right), \quad (\text{B1})$$

where

$$f_i = d_{i,RANS} - d_{i,benchmark}. \quad (\text{B2})$$

In the adjoint framework, the total derivative with respect to β is given by

$$\frac{d\mathfrak{J}_{obs}}{d\beta} = \frac{\partial \mathfrak{J}_{obs}}{\partial \beta} + \psi^T \frac{\partial \mathbf{R}}{\partial \beta}, \quad (\text{B3})$$

where the adjoint variable, ψ , is calculated using

$$\left[\frac{\partial \mathbf{R}}{\partial \mathbf{U}} \right]^T \psi = - \left[\frac{\partial \mathfrak{J}_{obs}}{\partial \mathbf{U}} \right]^T. \quad (\text{B4})$$

The partial derivatives in Eq. (B4) are calculated using automatic differentiation and the system is solved using pseudo-time stepping. In Eq. (B3), the first term is zero and the second term requires knowledge of the partial derivative of the residual with respect to β . Since β is explicitly present only as a multiplier to the production term,

$$\frac{\partial \mathbf{R}_{\tilde{v}_t}}{\partial \beta} = -\mathbf{P}(\tilde{v}_t, \mathbf{U}), \quad (\text{B5})$$

where $\mathbf{R}_{\tilde{v}_t}$ refers to the residual of the SA equation. Then the expression for the derivative is given by

$$\frac{d\mathfrak{J}_{obs}}{d\beta} = \frac{\partial \mathfrak{J}_{obs}}{\partial \beta} - \psi_{\tilde{v}}^T \mathbf{P}(\tilde{v}, \mathbf{U}). \quad (\text{B6})$$

APPENDIX C: HESSIAN APPROXIMATION

For an objective function in Equation (B1), the Hessian matrix is given by

$$\frac{d^2\mathfrak{J}_{\text{obs}}}{d\boldsymbol{\beta}_j d\boldsymbol{\beta}_k} = \frac{1}{\sigma_{\text{obs}}^2} \sum_i \left(\frac{df_i}{d\boldsymbol{\beta}_j} \frac{df_i}{d\boldsymbol{\beta}_k} + f_i \frac{d^2f_i}{d\boldsymbol{\beta}_j d\boldsymbol{\beta}_k} \right) \quad (\text{C1})$$

$$= \frac{1}{\sigma_{\text{obs}}^2} \left(\nabla \mathbf{F}^T \nabla \mathbf{F} + \sum_i f_i \frac{d^2f_i}{d\boldsymbol{\beta}_j d\boldsymbol{\beta}_k} \right). \quad (\text{C2})$$

In the Gauss–Newton approximation, the second term in Equation (C2) is omitted. The approximation seems reasonable when either all the f_i 's are zero or the gradient of f_i 's (extrema) are zero, which is ideally the case at the MAP solution. The matrix $\nabla \mathbf{F}$ is constructed by running the adjoint solver for every data point used for the inversion and with an objective function of f_i .

- ¹ J. Slotnick, A. Khodadoust, J. Alonso, D. Darmofal, W. Gropp, E. Lurie, and D. Mavriplis, “CFD vision 2030 study: A path to revolutionary computational aerosciences,” NASA Technical Report, NASA/CR-2014-218178, NF1676L-18332, 2014.
- ² G. A. Gerolymos, E. Sauret, and I. Vallet, “Contribution to single-point closure Reynolds-stress modelling of inhomogeneous flow,” *Theor. Comput. Fluid Dyn.* **17**, 407–431 (2004).
- ³ M. E. Olsen, “Prediction of separation with a third-order-moment model,” AIAA 2015-1968, 2015.
- ⁴ S. Poroseva and S. M. Murman, “Velocity/pressure-gradient correlations in a forans approach to turbulence modeling,” AIAA 2014-2207, 2014.
- ⁵ T. A. Oliver, G. Terejanu, C. S. Simmons, and R. D. Moser, “Validating predictions of unobserved quantities,” *Comput. Methods Appl. Mech. Eng.* **283**, 1310–1335 (2015).
- ⁶ J. Brynjarsdóttir and A. O’Hagan, “Learning about physical parameters: The importance of model discrepancy,” *Inverse Probl.* **30**, 114007 (2014).
- ⁷ S. V. Poroseva and S. M. Murman, “Reynolds-stress simulations of a fully developed channel flow using a new velocity/pressure-gradient model,” in *International Symposium on Turbulence and Shear Flow Phenomena (TSFP-9)* (Proceeding of the symposium, No. P-38, 2015).
- ⁸ S. Yarlanki, B. Rajendran, and H. Hamann, “Estimation of turbulence closure coefficients for data centers using machine learning algorithms,” in *13th IEEE Intersociety Conference on Thermal and Thermomechanical Phenomena in Electronic Systems (ITHERM)* (IEEE, 2012), pp. 38–42.
- ⁹ H. Kato and S. Obayashi, “Data assimilation for turbulent flows,” AIAA 2014-1177, 2014.
- ¹⁰ W. Edeling, P. Cinnella, R. P. Dwight, and H. Bijl, “Bayesian estimates of parameter variability in the $k-\epsilon$ turbulence model,” *J. Comput. Phys.* **258**, 73–94 (2014).
- ¹¹ E. Dow and Q. Wang, “Quantification of structural uncertainties in the $k-\omega$ turbulence model,” in *Structures, Structural Dynamics, and Materials and Co-located Conferences* (AIAA, 2011).
- ¹² S. H. Cheung, T. A. Oliver, E. E. Prudencio, S. Prudhomme, and R. D. Moser, “Bayesian uncertainty analysis with applications to turbulence modeling,” *Reliab. Eng. Syst. Saf.* **96**, 1137–1149 (2011).
- ¹³ M. Emory, J. Larsson, and G. Iaccarino, “Modeling of structural uncertainties in Reynolds-averaged Navier-Stokes closures,” *Phys. Fluids* **25**, 110822 (2013).
- ¹⁴ J. Ling and J. Templeton, “Evaluation of machine learning algorithms for prediction of regions of high Reynolds averaged Navier Stokes uncertainty,” *Phys. Fluids* **27**, 085103 (2015).
- ¹⁵ E. Parish and K. Duraisamy, “A paradigm for data-driven modeling using functional inversion and machine learning,” *J. Comput. Phys.* **305**, 758–774 (2016).
- ¹⁶ K. Duraisamy, Z. J. Zhang, and A. P. Singh, “New approaches in turbulence and transition modeling using data-driven techniques,” AIAA 2015-1284, 2015.
- ¹⁷ Z. J. Zhang and K. Duraisamy, “Machine learning methods for data-driven turbulence modeling,” AIAA 2015-2460, 2015.
- ¹⁸ D. D. Gray and A. Giorgini, “The validity of the Boussinesq approximation for liquids and gases,” *Int. J. Heat Mass Transfer* **19**, 545–551 (1976).
- ¹⁹ A. M. Stuart, “Inverse problems: A Bayesian perspective,” *Acta Numer.* **19**, 451–559 (2010).
- ²⁰ W. R. Gilks, *Markov Chain Monte Carlo* (Wiley Online Library, 2005).
- ²¹ R. C. Aster, B. Borchers, and C. H. Thurber, *Parameter Estimation and Inverse Problems* (Academic Press, 2013).
- ²² K. Duraisamy and J. Alonso, “Adjoint based techniques for uncertainty quantification in turbulent flows with combustion,” in *42nd AIAA Fluid Dynamics Conference and Exhibit* (AIAA, 2012), p. 2711; available at <http://arc.aiaa.org/doi/abs/10.2514/6.2012-2711>.
- ²³ R. Pecnik, V. E. Terrapon, F. Ham, G. Iaccarino, and H. Pitsch, “Reynolds-averaged Navier-Stokes simulations of the HyShot II scramjet,” *AIAA J.* **50**, 1717–1732 (2012).
- ²⁴ Q. Wang, K. Duraisamy, J. J. Alonso, and G. Iaccarino, “Risk assessment of scramjet unstart using adjoint-based sampling methods,” *AIAA J.* **50**, 581–592 (2012).
- ²⁵ M. B. Giles, M. C. Duta, J.-D. M-acute, and N. A. Pierce, “Algorithm developments for discrete adjoint methods,” *AIAA J.* **41**, 198–205 (2003).
- ²⁶ A. Walther and A. Griewank, “Getting started with ADOL-C,” *Comb. Sci. Comput.* 181–202 (2012).
- ²⁷ S. Hoyas and J. Jiménez, “Scaling of the velocity fluctuations in turbulent channels up to $Re_\tau = 2003$,” *Phys. Fluids* **18**, 011702 (2006).
- ²⁸ M. Lee and R. D. Moser, “Direct numerical simulation of turbulent channel flow up to $Re_\tau = 5200$,” preprint [arXiv:1410.7809](https://arxiv.org/abs/1410.7809) (2014).

- ²⁹ J. C. del Alamo and J. Jimenez, "Direct numerical simulation of the very large anisotropic scales in a turbulent channel," preprint [arXiv:1309.2322](https://arxiv.org/abs/1309.2322) (2013).
- ³⁰ D. C. Wilcox, *Turbulence Modeling for CFD* (D C W Industries, 2006), Vol. 3.
- ³¹ P. Bradshaw, "The analogy between streamline curvature and buoyancy in turbulent shear flow," *J. Fluid Mech.* **36**, 177–191 (1969).
- ³² J. Gillis and J. Johnston, "Turbulent boundary-layer flow and structure on a convex wall and its redevelopment on a flat wall," *J. Fluid Mech.* **135**, 123–153 (1983).
- ³³ V. Patel and F. Sotiropoulos, "Longitudinal curvature effects in turbulent boundary layers," *Prog. Aerosp. Sci.* **33**, 1–70 (1997).
- ³⁴ W. S. Saric, "Gortler vortices," *Annu. Rev. Fluid Mech.* **26**, 379–409 (1994).
- ³⁵ P. Spalart and M. Shur, "On the sensitization of turbulence models to rotation and curvature," *Aerosp. Sci. Technol.* **1**, 297–302 (1997).
- ³⁶ A. Hellsten, "Some improvements in menter's $k-\omega$ SST turbulence model," in *Fluid Dynamics and Co-located Conferences* (AIAA-98-2554, 1998).
- ³⁷ K. Duraisamy and G. Iaccarino, "Curvature correction and application of the v2-f turbulence model to tip vortex flows," in *Center for Turbulence Research Annual Research Briefs* (CTR, Stanford, 2015), pp. 157–168; <https://web.stanford.edu/group/ctr/ResBriefs05/duraisamy.pdf>.
- ³⁸ M. L. Shur, M. K. Strelets, A. K. Travin, and P. R. Spalart, "Turbulence modeling in rotating and curved channels: Assessing the spalart-shur correction," *AIAA J.* **38**, 784–792 (2000).
- ³⁹ S. K. Arolla and P. A. Durbin, "LES of spatially developing turbulent boundary layer over a concave surface," *J. Turbul.* **16**, 81–99 (2015).
- ⁴⁰ P. Durbin, personal communication, Iowa State University (2015).
- ⁴¹ A. Ferri, *Experimental Results with Airfoils Tested in the High-Speed Tunnel at Guidonia* (National Advisory Committee for Aeronautics, 1940).
- ⁴² D. S. Dolling and M. T. Murphy, "Unsteadiness of the separation shock wave structure in a supersonic compression ramp flowfield," *AIAA J.* **21**, 1628–1634 (1983).
- ⁴³ D. S. Dolling, "Fifty years of shock-wave/boundary-layer interaction research: What next?," *AIAA J.* **39**, 1517–1531 (2001).
- ⁴⁴ D. D. Knight and G. Degrez, "Shock wave boundary layer interactions in high mach number flows: A critical survey of current CFD prediction capabilities," in *AGARD AR-319-Hypersonic Experimental and Computational Capability, Improvement and Validation* (NATO Science and Technology Organization, 1998).
- ⁴⁵ B. Morgan, K. Duraisamy, N. Nguyen, S. Kawai, and S. Lele, "Flow physics and RANS modelling of oblique shock/turbulent boundary layer interaction," *J. Fluid Mech.* **729**, 231–284 (2013).
- ⁴⁶ D. Somers, "Design and experimental results for the S809 airfoil," NREL/SR-440-6918, 1997.
- ⁴⁷ Y. Chang, S. Yang, and O. Arici, "Flow field computation of the NREL S809 airfoil using various turbulence models," Technical Report(s) CONF-960154—ISBN 0-9648731-8-4; TRN: IM9641%210, American Society of Mechanical Engineers, New York, NY, USA, 1996.
- ⁴⁸ S. Yang, Y. Chang, and O. Arici, "Incompressible Navier-Stokes computation of the NREL airfoils using a symmetric total variational diminishing scheme," *J. Sol. Energy Eng.* **116**, 174–182 (1994).
- ⁴⁹ W. Wolfe and S. Ochs, "CFD calculations of S809 aerodynamic characteristics," in *Aerospace Sciences Meetings* (AIAA, 1997).
- ⁵⁰ B. D. Tracey, K. Duraisamy, and J. J. Alonso, "A machine learning strategy to assist turbulence model development," AIAA 2015-1287, 2015.
- ⁵¹ C. F. John Salvatier and T. V. Wiecki, Probabilistic programming in python using pymc, (2015), <https://github.com/pymc-devs/pymc3>.
- ⁵² Note that is still a very high-dimensional problem.
- ⁵³ This dimensionality could be reduced to a degree by expressing β as a weighted sum of appropriately defined basis functions.
- ⁵⁴ R is an upper triangular matrix.
- ⁵⁵ In this work, all the RANS results (prior to inversion) were grid-converged.