Convolutional neural network for transition modeling based on linear stability theory

Muhammad I. Zafar and Heng Xiao[*]
Kevin T. Crofton Department of Aerospace and Ocean Engineering, Virginia Tech, Blacksburg, Virginia 24061, USA

Meelan M. Choudhari, Fei Li, and Chau-Lyan Chang
NASA Langley Research Center, Hampton, Virginia 23681, USA

Pedro Paredes and Balaji Venkatachari
National Institute of Aerospace, Hampton, Virginia 23666, USA

Transition prediction is an important aspect of aerodynamic design because of its impact on skin friction and potential coupling with flow separation characteristics. Traditionally, the modeling of transition has relied on correlation-based empirical formulas based on integral quantities such as the shape factor of the boundary layer. However, in many applications of computational fluid dynamics, the shape factor is not straightforwardly available or not well-defined. We propose using the complete velocity profile along with other quantities (e.g., frequency, Reynolds number) to predict the perturbation amplification factor. While this can be achieved with regression models based on a classical fully connected neural network, such a model can be computationally more demanding. We propose a convolutional neural network inspired by the underlying physics as described by the stability equations. Specifically, convolutional layers are first used to extract integral quantities from the velocity profiles, and then fully connected layers are used to map the extracted integral quantities, along with frequency and Reynolds number, to the output (amplification ratio). Numerical tests on classical boundary layers clearly demonstrate the merits of the proposed method. More importantly, we demonstrate that, for Tollmien-Schlichting instabilities in two-dimensional, low-speed boundary layers, the proposed network encodes information in the boundary-layer profiles into an integral quantity that is strongly correlated to a well-known, physically defined parameter—the shape factor.

I. INTRODUCTION

Laminar-turbulent transition of boundary-layer flows can have a strong impact on the performance of flight vehicles because of their influence on surface skin friction and aerodynamic heating. Therefore, transition prediction is a key issue for the design of next-generation aerospace configurations. Indeed, according to the CFD Vision 2030 Study [1], the most critical area in computational fluid dynamics (CFD) simulation capability that will remain a pacing item for the foreseeable future is the ability to adequately predict viscous flows involving transition-to-turbulence and/or flow separation.

---

[*]Author to whom all correspondence should be addressed: hengxiao@vt.edu

Under the benign disturbance environment in flight, boundary-layer transition is often initiated by the amplification of linearly unstable eigenmodes of the laminar boundary layer. For two-dimensional (2D) and weakly 3D boundary layers developing over a nominally smooth surface, the dominant instability mechanisms correspond to streamwise propagating Tollmien-Schlichting (TS) waves at subsonic speeds, oblique first-mode disturbances at supersonic edge Mach numbers, and again planar waves of the second mode, i.e., Mach mode type in hypersonic flows. Additionally, centrifugal instabilities in the form of Görtler vortices are known to influence the transition process over surfaces with significant regions of concave longitudinal curvature. Finally, attachment line and crossflow instabilities come into play when the flow becomes three-dimensional.

The transition process begins with the generation of instability waves via the interaction (i.e., receptivity) of the laminar boundary layer to its disturbance environment. However, the onset of turbulence ensues only after the instability waves have gained sufficiently large amplitudes to undergo a sequence of nonlinear processes that culminates with the breakdown to turbulence. Because the nonlinear breakdown tends to be relatively rapid, the slow amplification of the linear instability waves accounts for a majority of the laminar flow region preceding the onset of transition. As a result, the linear amplification ratio, e^N, of the most amplified instability mode can often be used to predict the experimental trends in the transition location. The linear amplification ratio is usually computed by using the classical stability theory based on the assumption of a quasiparallel boundary layer flow. Reviews of the linear stability theory for fluid flows may be found in Refs. [2–5]. The connection between the stability theory and laminar-turbulent transition in boundary-layer flows has been discussed in Refs. [6–8]. Prior work [9,10] has shown that the N-factor values between 9 and 11 correlate with the transition locations measured in a broad class of boundary-layer flows.

Direct computations of boundary-layer stability place rather stringent demands on the accuracy of mean flow calculations, much more so in comparison with that required for the prediction of aerodynamic metrics such as the skin friction drag. In addition, the solution to the eigenvalue problem associated with the discretized version of the linear stability equations incurs a significant computational cost. Furthermore, due to the complex nature of the eigenvalue spectra and their sensitivity to both input parameters and numerical discretization, stability computations are difficult to automate, and they also require significant user expertise in the details of the hydrodynamic stability theory. Consequently, the task of transition prediction based on the N-factor methods has been a specialist's domain, and is often performed as a post-processing step that follows the computation of the laminar boundary layer over the flow configuration of interest. Implicit in this post-processing approach is the assumption of a weak coupling between the transition location and the basic state computation.

The weak-coupling assumption may be justified for fully attached boundary-layer flows such as aircraft wings at the cruise condition. However, a number of technological applications, such as high-lift systems [11], rotorcraft [12], and other configurations involving flow separation, entail a strong viscous-inviscid interaction, requiring an iterative prediction approach that reflects the stronger coupling between transition and the overall flow field. To that end, the CFD Vision 2030 Study has called for transition prediction methods that can be fully integrated into the overall process of aerodynamic prediction. In the past, several attempts have been made to simplify the application of the N-factor methods in the engineering environment, ranging from analytical but potentially complex data fits [9,13,14] to numerical, table-lookup procedures based on a prior database of stability results [15–22]. In recent work, such empirical fits have also been incorporated into CFD integrated transition models based on auxiliary transport equations, such as the amplification factor transport model [23].

In the above-mentioned database-query techniques, a response surface model is developed in terms of a small number of scalar input parameters representing the combination of the global flow parameters, selected measures of boundary-layer profiles, and the relevant disturbance characteristics such as frequency and wave-number parameters. Almost universally, one or more shape factors of the boundary-layer profiles have been used to encapsulate the complex dependence of the disturbance amplification rates on the underlying mean flow. This tends to limit the expressive

power of the model as discussed by Crouch *et al.* [24]. Secondarily, while such shape factors can be easily evaluated when the mean flow computation is based on the classical boundary-layer theory, it is not easy to compute them in a consistent and accurate manner when Navier-Stokes codes are used for the basic state computation, especially for unstructured grid solutions. Such situations arise rather commonly in high-speed applications, such as the flow past blunt nosed bodies, where the inviscid flow beyond the edge of the boundary layer includes nonzero vorticity as a result of vorticity generation at the curved shock. The database methods can perform rather poorly in these cases, as demonstrated by Paredes *et al.* [25].

Stability predictions based on artificial neural networks [24,26,27] allow additional features of the boundary-layer profiles to be taken into account without sacrificing the computational efficiency and robustness of the conventional methods based on a previously computed database of amplification characteristics. The neural network methods can also be easily generalized to higher-dimensional input features. This allows the multiparameter dependence of the stability characteristics to be accounted for, whereas the conventional methods for database query do not scale very well as the number of independent parameters becomes significantly large. Neural-network-based stability predictions for free-shear layer flows were first presented by Fuller *et al.* [26]. However, a significant advance related to transition prediction was made by Crouch *et al.* [24], who found that the expressivity of the model could be improved by augmenting the set of scalar variables used in conventional database methods via the slopes of the appropriately normalized velocity profiles at six equidistant points across the boundary layer. The details of the neural network architecture used in that paper are somewhat limited; however, a feed-forward network based on fully connected hidden layers was used to approximate the maximum amplification rate among all unstable modes at any given station. This maximum amplification rate was integrated along the airfoil surface to evaluate the *N*-factors.

The selection of a smaller number of input features by Crouch *et al.* [24] was somewhat arbitrary and is unlikely to lend itself to other instability mechanisms without further modifications. In the present work, we present an alternate approach based on convolutional neural networks that can automatically learn a reduced-order representation of the boundary-layer profiles in terms of a specified number of most significant features that can optimally predict the targeted linear stability characteristics across the training space. As such, the proposed architecture can be easily adapted to predict the amplification characteristics of a broad range of very different instability mechanisms.

The objective of this paper is to present a preliminary proof of concept to establish the potential of convolutional neural networks (CNNs) to distill the latent features of the boundary layer profiles. To that end, it is sufficient to consider the simplest case of TS instability waves in two-dimensional, incompressible boundary layers. Because the architecture of the CNN is not related to the specific physics of the TS instability mechanism or to the flow geometry, the proposed neural network can be easily generalized to other classes of instability waves.

The rest of the paper is organized as follows. An overview of the methodology is presented in Sec. II, which includes a summary of modal stability analysis for two-dimensional, incompressible flows, followed by a description of the proposed architecture of a hybrid convolutional/fully connected neural network. Section III presents the results based on the training and validation of the proposed network architecture, including an assessment of its generalization capability for real-world applications. In particular, we highlight the capability of the proposed network for encoding boundary-layer profiles into integral parameters in an automatic, data-driven manner. Section IV concludes the paper.

II. METHODOLOGY

Using the e^N method, the onset of laminar-turbulent transition is predicted to occur where the logarithmic amplification ratio N of the most amplified instability mode reaches an empirically defined critical value, denoted herein as $N = N_{tr}$. As described in Sec. II A, the logarithmic amplification ratio N can be computed by solving the governing equations based on linear stability theory.

Thus, N is a direct function of the laminar boundary-layer profiles (e.g., U, dU/dy, and d^2U/dy^2), the flow Reynolds number Re_θ based on local momentum thickness of the boundary layer, and disturbance parameters such as the frequency of the instability wave ω and, for 3D disturbances, the spanwise wave number. The objective herein is to develop a surrogate transition model based on a neural network that would incorporate the physics of the transition phenomenon and predict the transition onset location without requiring the direct computations using linear stability theory. Since this data-driven surrogate transition model is being developed based on linear stability theory, any limitations of the quasiparallel theory are inherited by the present implementation of the neural-network-based surrogate model. However, since the methodology developed here has no intrinsic dependence on the quasiparallel theory, the basic framework can be extended to use any advanced simulation method, such as parabolized stability equations (PSEs), to generate the data and thus to overcome of the limitations of the classical theory. This would be a fruitful avenue for future work.

We propose a hybrid neural network (consisting of convolutional and fully connected layers) that allows the relevant flow information (boundary-layer profiles and scalar disturbance characteristics) to be processed in a physically informed manner. The resulting regression models predict the local instability amplification rates, σ, for the relevant frequencies of the instability waves as they propagate through the laminar boundary layer. The N-factor curves can be computed by integrating the predicted σ values to allow the transition onset location to be predicted on the basis of the empirically defined critical value, N_{tr}. Moreover, as a baseline to compare the performance of the proposed neural network, we also consider the fully connected neural network architecture based on the idea proposed earlier by Crouch *et al.* [24]. In the first part of this section, we discuss the basis of the e^N method and the linear stability theory. The transition models based on the proposed neural network and the fully connected neural network are outlined in the second part. The database used to train these neural-network-based transition models is discussed in the last part of this section.

A. The e^N method

For simplicity, we outline the transition prediction procedure in the context of an incompressible, fully attached flow over a two-dimensional airfoil. An orthogonal, body-fitted, curvilinear coordinate system (s, n) is introduced such that s denotes the distance from the stagnation point, measured along the airfoil contour on either the suction or the pressure side of the airfoil, and n represents the outward surface normal. Consistent with the boundary-layer character of the basic state, both coordinates are normalized by a length scale comparable to the local thickness of the boundary layer, which is taken to be the local momentum thickness $\theta(s)$ without any loss of generality. The two-dimensional boundary-layer flow over the airfoil is represented by the velocity field (U, V), where the velocities are normalized by a local velocity scale $U_e(s)$, taken to be the flow speed at the edge of the mean boundary layer. The Reynolds-number parameter based on the free-stream speed and the momentum thickness is denoted by Re_θ.

We consider small-amplitude, time-harmonic, spatially evolving perturbations to the mean boundary-layer flow of the form

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} U(n) \\ V(n) \end{bmatrix} \exp\left(i[\varphi(s) - \omega t]\right), \tag{1}$$

where $i = \sqrt{-1}$ is the imaginary unit, $\omega \equiv 2\pi f$ denotes the real valued disturbance frequency, f is the frequency parameter in Hz, t is the appropriately normalized time, and $d\varphi/ds = \alpha$ denotes the complex streamwise wave number. Substituting the above normal mode ansatz into the linearized Navier-Stokes equations and neglecting the weak nonparallel effects associated with the $\mathrm{O}(Rde^{-1})$ velocity component $\mathbf{V}$ and the slow streamwise evolution $\partial U/\partial s$ of the tangential velocity field, one obtains the quasiparallel form of the disturbance equations that must be solved with homogeneous boundary conditions for U and V at the surface ($n = 0$) as well as in the free stream ($n \to \infty$). For an incompressible flow, the quasiparallel disturbance equations can be combined into a single

equation for the wall-normal velocity perturbation, yielding an eigenvalue problem based on the well known Orr-Sommerfeld (OS) equation [28]:

$$(\alpha U - \omega)(V'' + \alpha^2 V) - \alpha U'' V = (V'''' - 2\alpha^2 V'' + \alpha^4)/(i \operatorname{Re}_\theta), \tag{2}$$

along with homogeneous Dirichlet boundary conditions:

$$V(0) = V'(0) = 0 \quad \text{and} \quad V(\infty) = V'(\infty) = 0, \tag{3}$$

where a prime denotes the derivative with respect to the wall-normal coordinate n. The solution to the eigenvalue problem [Eqs. (2) and (3)] at a given station s determines the local value of the complex streamwise wave number α as a function of the frequency parameter ω. The local, streamwise amplification rate of a disturbance at frequency ω corresponds to $\sigma = -\operatorname{Im}(\alpha(\omega, s))$, where $\operatorname{Im}(\cdot)$ denotes the imaginary part of a complex quantity. Hence, the logarithmic amplification of the disturbance amplitude with respect to the neutral station, where the disturbance first begins to amplify, is given by

$$N(\omega, s) = \int_{s_0}^{s} \sigma(\omega, \tilde{s}) d\tilde{s}, \tag{4}$$

where the subscript "0" denotes the neutral station. The e^N method postulates that transition is likely to occur when the envelope N-factor, $N_e(s) = \sup(N(\omega, s))$, reaches the critical value of $N_e(s) = N_{\text{tr}}$. Here "sup" denotes the maximum over the frequency range of all unstable disturbances. Values of $N_{\text{tr}} = 9$–11 have been found to correlate with the onset of transition in a number of subsonic and supersonic flows [9,10].

B. Convolutional neural network

A neural network is a sequence of composite functions representing the mapping from an input vector $\mathbf{q}$ to output vector $\mathbf{y}$. Each member of the sequence is parametrized by the weight matrix $\mathbf{W}$ and bias vector $\mathbf{b}$, which can both be learned iteratively by using the available training data. This sequence of composite functions is arranged in the form of layers that consist of several neurons in general. For example, a neural network with one intermediate (hidden) layer $\mathbf{h}$ between the input layer ($\mathbf{q}$) and output layer ($\mathbf{y}$) may be represented by the following composite functional mapping:

$$\mathbf{y} = \mathbf{W}^{(2)}\mathbf{h} + \mathbf{b}^{(2)} \quad \text{with} \quad \mathbf{h} = f[\mathbf{W}^{(1)}\mathbf{q} + \mathbf{b}^{(1)}], \text{ or equivalently}$$
$$\mathbf{y} = \mathbf{W}^{(2)}(f[\mathbf{W}^{(1)}\mathbf{q} + \mathbf{b}^{(1)}]) + \mathbf{b}^{(2)}, \tag{5}$$

where f is an activation function, and $\mathbf{W}^{(i)}$ and $\mathbf{b}^{(i)}$ represent the weight matrix and biases vector for the ith layer, respectively. Activation functions introduce the nonlinearity in the composite functions that enables them to represent arbitrarily complex functional relationships. Several different activation functions have been proposed for this purpose, such as the sigmoid function $f(x) = 1/(1 + e^{-x})$ or the rectified linear unit (ReLU) $f(x) = \max(0, x)$. The training of the neural network consists of successive adjustments of the weights and the biases in order to minimize the squared error between the predicted and truth values of the output feature, i.e., the local amplification rate σ in the present application. Neural networks with at least one hidden layer are *universal approximators*, i.e., they can represent any continuous function on a compact domain to arbitrary accuracy, given a sufficient number of neurons in the hidden layer [29].

In a fully connected neural network, each neuron in a given layer is connected to every neuron in the adjacent layer, yielding a generic connection pattern that makes no assumptions about the input features in the data. The schematic of a fully connected neural network with multiple input features is displayed in Fig. 1. As mentioned in the Introduction, the existing models starting with Ref. [13] have used analytical curve fits or other database query methods to predict the local amplification rate of the instability wave as a function of the disturbance frequency ω, local Reynolds number Re_θ, and a scalar parameter such as the shape factor H. A fully connected neural network such
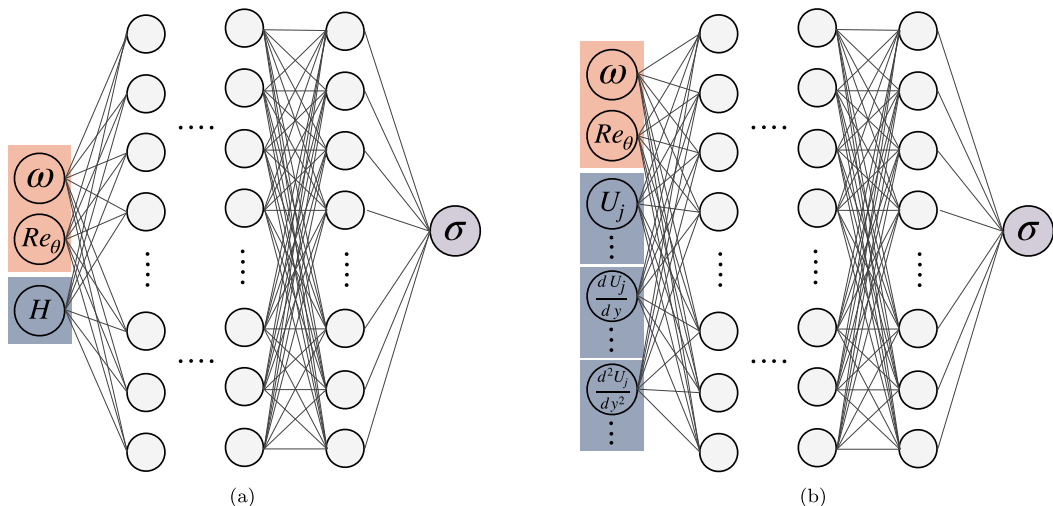
FIG. 1. Fully connected neural networks, (a) with input features including a scalar boundary-layer parameter, shape factor H, and (b) with boundary-layer profiles as direct input features. Other scalar parameters in both architectures represent disturbance characteristics of instability waves, i.e., frequency (ω) and Reynolds number (Re_θ).

as that shown in Fig. 1(a) provides a suitable architecture to achieve a similar functionality by using a neural network in place of the database interpolation or curve fitting. However, we note that the scalar boundary-layer parameter H does not appear directly in the governing equations for linear stability theory, and hence it is only indirectly related to the amplification characteristics of the instability modes. Moreover, the shape factor cannot be determined in a straightforward and/or accurate manner for several boundary layers, such as high-speed flows over blunt nose configurations [25]. Therefore, it is desirable to introduce the boundary-layer profiles directly into the predictive model, which is *physically more consistent* with the underlying linear stability equation. Crouch *et al.* [24] presented a model of this type by including a coarse representation of the mean velocity profiles as part of the input for the fully connected neural network. Whereas they used the velocity and its first derivative at just six equidistant points, Fig. 1(b) presents a model architecture based on a well-resolved representation of the boundary-layer profiles (i.e., velocity profile U and its derivatives dU/dy and d^2U/dy^2) as input features for the fully connected neural network. However, by using a large number of parameters to characterize the boundary-layer profiles, a model architecture of this type risks a potential misrepresentation of the dependence of instability growth rates on the two remaining physical parameters, namely ω and Re_θ. The balance involving the number of input parameters becomes more lopsided for high-speed flows, since the boundary-layer profiles for temperature or density are also required for a reliable prediction of the amplification rate σ.

A convolutional neural network (CNN) [30] is composed of a number of convolutional and pooling layers, which enable it to automatically extract the latent features of the input data, which are considered to be an ordered data structure. Specifically, the boundary-layer profiles can now be considered to be an ordered array, in contrast to the fully connected neural network that makes no assumption about the ordering of basic state quantities across the input profile. Furthermore, the CNN exploits two special attributes to learn efficiently with a smaller number of model parameters. First, each neuron in the convolutional layers has only local connections to the neurons in the previous layer, allowing it to develop a correlation with the neighboring neurons. Second, the convolutional kernel has translational invariance in terms of model parameters, which leads to a drastic reduction in the number of network parameters. These attributes allow the CNN to achieve
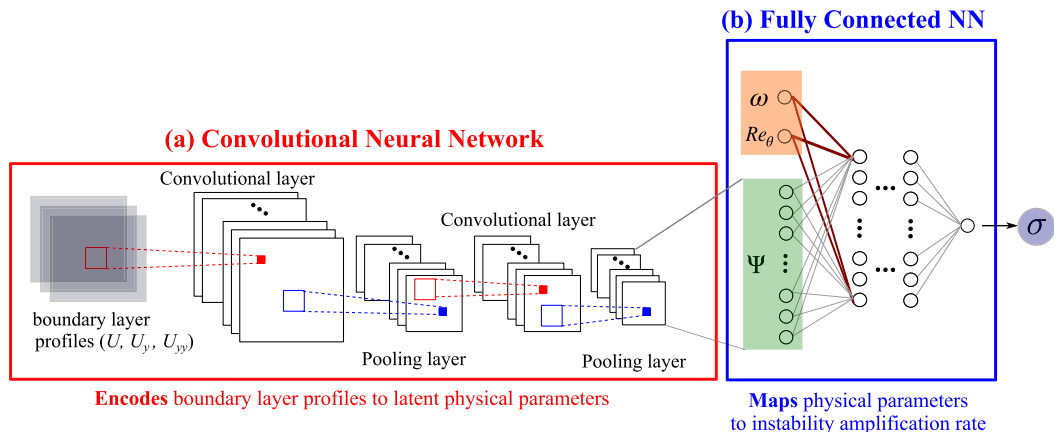
FIG. 2. Proposed hybrid convolutional neural network architecture including (a) a regular CNN, which encodes the boundary-layer profiles to a set of latent physical parameters Ψ, and (b) a fully connected neural network, which maps the CNN-extracted features Ψ along with other physical parameters (frequency of the instability wave ω and Reynolds number Re_θ) to the output (instability amplification rate σ).

a comparable predictive accuracy much more efficiently in terms of the training cost and/or the amount of data required for training [31].

A schematic of the proposed hybrid neural network is presented in Fig. 2. In this network architecture, the CNN first [Fig. 2(a)] maps the boundary-layer profiles to a specified number of latent features in a physically consistent manner while accounting for the spatial proximity of neighboring points across the boundary-layer profiles. The encoding of the boundary-layer profiles in the form of these latent features is denoted by the vector Ψ. Following a preliminary assessment, the number of latent features in vector Ψ was empirically chosen to be 8. However, the results are not significantly sensitive to this parameter. Next, the revised set of input parameters comprising the vector Ψ and the remaining physical scalar parameters (ω and Re_θ) is nonlinearly mapped through a fully connected neural network [Fig. 2(b)] to yield the local instability amplification rate σ as the final output. Observe that the dependence of σ on the physical parameters ω and Re_θ is introduced into the network architecture in an explicit yet flexible manner. In particular, the relationship between the local instability amplification rate and the parameters ω and Re_θ, along with the boundary-layer profiles, is known from the linear stability theory [Eqs. (2) and (3)], and that quantitative relation can be inferred via the training process.

The hyperparameters of the proposed neural network (Fig. 2) and those of fully connected neural networks (Fig. 1) have been empirically selected to yield an adequate complexity of the neural network model for learning all of the required information, without causing an overfitting of the training data. The list of primary hyperparameters includes the number of convolutional layers, the number of channels in each convolutional layer, the number of fully connected layers, the number of neurons in each fully connected layer, and the learning rate for the training of the model. A summary of the relevant model architectures is given in Table I, wherein each category of architecture has been labeled as A, B, or $C_{1,2,3}$ for future reference. The number of neurons in each fully connected layer and the number of channels in each convolutional layer have also been listed for the respective architectures. The number of input channels for networks C_1, C_2, and C_3 can be varied to accommodate the desired number of boundary-layer profiles that are to be used as input to the convolutional layers. Convolutional kernels of size 3×1 have been used to extract the latent features from the boundary-layer profiles defined by 41 equidistant points along the wall-normal direction. In all networks, the ReLU is used as the activation function, and the Adam optimization algorithm [32] has been chosen to minimize the sum squared error during the training process. All of

TABLE I. Details of network architectures along with respective input features. Acronyms used: FC, fully connected; NN, neural network; NL, neurons in FC network layers; CH, number of channels in CNN layers. The first fully connected layer of the C_i ($i = 1, 2, 3$) networks has (2+8) neurons, where the first two neurons correspond to the physical parameters (ω, Re_θ), and the eight additional neurons correspond to a vector of parameters, Ψ, that encodes the information from boundary-layer profiles. The C_i^* ($i = 1, 2, 3$) networks use only one scalar value for representing parameter Ψ, hence the first fully connected layer has (2+1) neurons.

Network	Input features	Architecture type	Architecture	Number of parameters
A	ω, Re_θ, H	Fig. 1(a): Fully connected NN	NL:[3,96,96,96,96,96,96,1]	56 353
B	ω, Re_θ, U_j, $\frac{dU}{dy}\big|_j$, $\frac{d^2U}{dy^2}\big|_j$	Fig. 1(b): Fully connected NN	NL:[125,96,96,96,96,96,96,1]	57 261
C_1	ω, Re_θ, U_j, $\frac{dU}{dy}\big|_j$, $\frac{d^2U}{dy^2}\big|_j$	Fig. 2: Convolutional NN + fully connected NN	CH: [3,6,8,4] + NL: [2+8,96,96,96,96,96,96,1]	57 337
C_2	ω, Re_θ, U_j, $\frac{dU}{dy}\big|_j$	Fig. 2: Convolutional NN + fully connected NN	CH: [2,4,8,4] + NL: [2+8,96,96,96,96,96,96,1]	57 257
C_3	ω, Re_θ, U_j	Fig. 2: Convolutional NN + fully connected NN	CH: [1,4,8,4] + NL: [2+8,96,96,96,96,96,96,1]	57 245
C_i^* $i = 1, 2, 3$	Same as C_i above	Same as C_i above	Same as C_i above, but with only 2+1 input features for the fully connected network	57 245

the neural network architectures considered herein have been implemented in the machine learning framework PYTORCH.

C. Generation of training data

The training database is obtained by solving the Orr-Sommerfeld (OS) eigenvalue problem [Eq. (3)] for the Falkner-Skan family of self-similar boundary-layer profiles over a wide range of pressure gradient parameter β_H and local Reynolds number $Re_\theta(s)$. It was generated by using the stability analysis software, LASTRAC, developed at the NASA Langley Research Center [33]. LASTRAC is a well-known software suite that has been extensively validated against existing benchmark data including direct numerical simulations.

The convolutional neural network model maps the complex dependence of the local instability amplification rate on the relevant disturbance characteristics and the mean flow parameters. The training database consists of a tabular listing of stability characteristics for the Falkner-Skan boundary layers, which includes a comprehensive, discrete sampling of the complex-valued local wave number α of the TS instability wave as a function of the real-valued frequency of the wave ω and the mean flow parameters, which include the Hartree pressure gradient parameter β_H, or equivalently the shape factor H of the Falkner-Skan velocity profile, the local Reynolds number Re_θ based on the momentum thickness, and the velocity profile and its first- and second-order derivatives.

The Falkner-Skan group of boundary layers supports a single instability mode that corresponds to the viscous-inviscid interactive TS waves by themselves ($\beta_H \geqslant 0$) or a combination of instability mechanisms involving the TS waves and the predominantly inviscid Rayleigh instabilities ($\beta_H < 0$). All stability calculations were carried out for a compressible boundary-layer flow with a vanishingly

TABLE II. Input features for neural network models.

Feature	Definition	Expression
q_1	Nondimensional frequency of the instability wave	ω
q_2	Reynolds number based on edge velocity and momentum thickness	Re_θ
q_3	Local value of velocity profile shape factor (derived parameter)	H
q_4	Velocity profile as a function of wall normal coordinate y	$U_j, j = 1, 2, \ldots, 41$
q_5	First-order derivative of velocity profile	$\frac{dU}{dy}\big|_j, j = 1, 2, \ldots, 41$
q_6	Second-order derivative of velocity profile	$\frac{d^2U}{dy^2}\big|_j, j = 1, 2, \ldots, 41$

small Mach number of 10^{-5}, and a stagnation temperature of 311.11 K, along with an adiabatic thermal wall boundary condition and zero transpiration velocity at the surface. All parameters included in the database are nondimensional. Lengths are scaled with respect to local momentum thickness, velocities with respect to the flow speed at the edge of the boundary layer, and temperature with respect to the local edge temperature. The database includes Hartree pressure gradient parameters in the range of $\beta_H \in [-0.1988, 1]$, corresponding to the discrete values given by

$$\beta_H = [-0.1988, -0.19, -0.18, -0.16, -0.14, -0.12, -0.10, -0.075, -0.05, -0.025,$$
$$0, 0.025, 0.05, 0.075, 0.1, 0.15, 0.2, 0.3, 0.4, 0.5, 0.6, 0.8, 1.0].$$

Because of a rapid change in the instability characteristics for $\beta_H < 0$, especially as $\beta_H \to -0.1988$ (i.e., when the boundary layer is on the verge of separation), the sampling in β_H is chosen to be denser at the negative values of β_H. This bias in sampling may result in a bias in the learning of the neural network models toward the lower limit of the β_H values. Stability computations were carried out for Reynolds numbers extending from just below the minimum critical Reynolds number (below which all disturbances are predicted to decay) up to $\text{Re}_s \equiv U_e s / \nu = 10^{10}$, where ν denotes the kinematic viscosity of the fluid. The frequency range at each Reynolds number included the entire range of unstable disturbances as well as a modest range of stable disturbances in the vicinity of the neutral stability curve. Because the parameter range covered multiple orders of magnitudes, a logarithmic increment was used along both axes. Due to computational considerations, a quarter of the data points in the database were used, which amounts to approximately 400 000. The results are not influenced by the down-sampling of training data. The input parameters used to train and predict the local instability amplification rate correspond to a suitable subset of the various features listed in Table II. Boundary-layer profiles include the velocity profile along with the first- and second-order derivatives, sampled at 41 equidistant points in order to resolve each profile. The scalar input features ($q_1 : \omega$, $q_2 : \text{Re}_\theta$, and $q_3 : H$) have been scaled and shifted to the range of [0, 1]. We note in passing that modified input features based on a logarithmic scale along the Re_θ and ω axes were also considered on the basis of the high Reynolds number asymptotic theory of Tollmien-Schlichting waves [34], but no significant improvement in the testing error was noted.

The present database has been generated as part of the NASA Langley Research Center's effort to use machine learning methods to enable robust, CFD-solver-friendly models for boundary-layer transition. This database will be made available in an electronic form to encourage the development of physics-based transition models that can be integrated with CFD solvers.

III. RESULTS

In this section, we demonstrate the predictive performance of the proposed convolutional neural network (Fig. 2) and compare it with the performance of the fully connected neural networks [Figs. 1(a) and 1(b)]. In the first part of this section, the proposed neural network model is validated using the Falkner-Skan database (presented in Sec. II C) by randomly splitting the data, with 90%

TABLE III. Comparison of the validation error corresponding to neural network models from Table I. Training and validation datasets correspond to a random 90–10 % split of the available database over the entire range of the pressure gradient parameter β_H.

Network	Input features	Validation error
A	$\omega, \mathrm{Re}_\theta, H$	0.58%
C_1	$\omega, \mathrm{Re}_\theta, U_j, \frac{dU}{dy}\big|_j, \frac{d^2U}{dy^2}\big|_j$	0.41%
C_2	$\omega, \mathrm{Re}_\theta, U_j, \frac{dU}{dy}\big|_j$	0.44%
C_3	$\omega, \mathrm{Re}_\theta, U_j$	0.46%

of the data points used for training and the remaining 10% for validation. Both the training and validation datasets contain data from the entire range of the pressure gradient parameter β_H. We then assess the proposed model for interpolation and extrapolation cases by splitting the Falkner-Skan database into training and testing datasets based on the data corresponding to each pressure gradient parameter β_H.

Finally, the capability of the network trained on the Falkner-Skan database (with self-similar boundary layers) to generalize its predictions to realistic flow configurations is evaluated by comparing the predictions of the neural network model with actual stability computations for those configurations. Specifically, we consider two different airfoils with non-self-similar boundary layers for this purpose, namely a symmetric 2D HSK airfoil [35] and an asymmetric NLF-0416 airfoil [36,37]. The Reynolds number parameter based on the free-stream speed and the chord length of the airfoil is chosen to be $\mathrm{Re}_c = 1.23 \times 10^6$ for the HSK airfoil and $\mathrm{Re}_c = 9 \times 10^6$ for the NLF-0416 airfoil. At the selected flow conditions, the boundary-layer instability is dominated by the TS instabilities of interest. In all cases, the following metric corresponding to the percent error based on the Frobenius norm is used for the evaluation of the model throughout this paper:

$$\epsilon_\sigma = 100 \times \frac{\|\sigma_{\mathrm{truth}} - \sigma_{\mathrm{predicted}}\|_F}{\|\sigma_{\mathrm{truth}}\|_F}, \tag{6}$$

where the Frobenius norm is defined as $\|X\|_F = \sqrt{\sum_i |X_i|^2}$. In the second part of this section, we analyze the feature learning and encoding capability of the convolutional neural network and how it makes the proposed model more robust and generalizable to other flow regimes. In the third part of this section, the potential advantages of the proposed model over the previously proposed model architecture [24] are analyzed.

A. Demonstration of predictive performance

The predictive performance of the convolutional neural network (Fig. 2) is first validated by using the Falkner-Skan database and is also compared with the performance of the fully connected network [Fig. 1(a)] with scalar input features. In these cases, the training process utilized a randomly sampled subset of the available dataset, amounting to 90% of the total data points. The validation is conducted by using the remaining 10% of the data that were never seen by the neural network model during the training process. The results presented in Table III show that the proposed model demonstrates slightly improved predictive performance (Network C_1, 0.41%) as compared to that of the fully connected neural network with scalar input features (Network A, 0.58%). Furthermore, the proposed neural network provides qualitatively similar results when the number of velocity profiles used as input features is varied from the velocity and its two derivatives (network C_1) to the velocity profile alone (network C_3). The data presented in Table III indicate that the inclusion of the velocity derivatives decreases the validation error, but the improvement is rather small. The small reduction in validation error is consistent with the fact that the derivative information is contained within the

TABLE IV. Results for interpolation and extrapolation in the Hartree pressure gradient parameter. The testing dataset is comprised of data at a specified value of β_H, while the remaining data are used as the training dataset.

Cases	Testing dataset	Testing error
Interpolation	$\beta_H = -0.1$	3.97%
	$\beta_H = 0.1$	3.02%
	$\beta_H = 0.5$	4.68%
Extrapolation	$\beta_H = -0.1988$	15.34%
	$\beta_H = 1.0$	28.52%

velocity profile itself, and therefore we believe that the observed improvement is attributed to the well-resolved yet finite sampling of the velocity profile. Since the validation error for the neural network based on the velocity profiles alone (i.e., network C_3) is already small, the margin for improvement is rather limited, and the findings in Table III confirm this expected behavior. Given the small differences in validation error corresponding to the networks C_1 through C_3, all of the remaining assessments reported in this paper are based on a single set of input features, which is chosen to include the velocity profile U and its derivatives dU/dy and d^2U/dy^2. Thus, the mapping sought by the neural network may be represented as

$$\left(\omega, \mathrm{Re}_\theta, U_j, \left. \frac{dU}{dy} \right|_j, \left. \frac{d^2U}{dy^2} \right|_j \right) \quad \longmapsto \quad \sigma, \quad \text{where} \quad j = 1, 2, \ldots, 41.$$

We now evaluate the performance of the proposed neural network for more challenging interpolation and extrapolation cases, where the Falkner-Skan database has been split for testing and training based on the Hartree pressure gradient parameter β_H. In these cases, data corresponding to a selected value of β_H are reserved for testing while the data corresponding to the remaining values of β_H are used for the training process. Interpolation cases were considered by isolating a single value of β_H for evaluating the testing error, and the corresponding results for $\beta_H = -0.1$, 0.1, and 0.5 are given in Table IV. The proposed neural network model is able to interpolate the remaining database at each of these three values of β_H with a testing error of between 3% and 5%. We also considered two additional cases based on testing corresponding to the two extremes of the β_H range, namely $\beta_H = -0.1988$ and 1.0, which amount to an extrapolation from the training database. As expected, the corresponding results in Table IV indicate significantly higher testing errors with the extrapolation in comparison with the testing errors for the three interpolation cases discussed above. Since most of the neural networks are intrinsically interpolators and do not perform well for something beyond the distribution of training data, the proposed neural network is not able to extrapolate well. Some of the neural network architectures that impose physical knowledge such that the model satisfies the governing equations (i.e., cost functions based on ordinary differential equations) are capable of extrapolating significantly [38]. In current work, while sound physical knowledge has been used to guide learning of underlying functions, the model is still an interpolator just like most of the other neural networks, random forest models, or other regression models [39–41] in physical modeling. Further, the testing error percentage toward the lower limit ($\beta_H = -0.1988$) is significantly lower (15.3%) than that toward the higher limit of the range of pressure gradient parameters (namely, 28.5% error for $\beta_H = 1.0$). This disparity in testing errors may be due to the significantly denser sampling in β_H toward the lower end of the range, indicating that the overall accuracy of the neural network could perhaps be improved by increasing the database size by including additional data at higher values of β_H.

As the underlying purpose of the proposed neural network model is to predict the transition onset location, we next evaluate the predictive performance of the neural network model C_1 for a symmetric HSK airfoil section and an asymmetric NLF-0416 airfoil section. These two examples
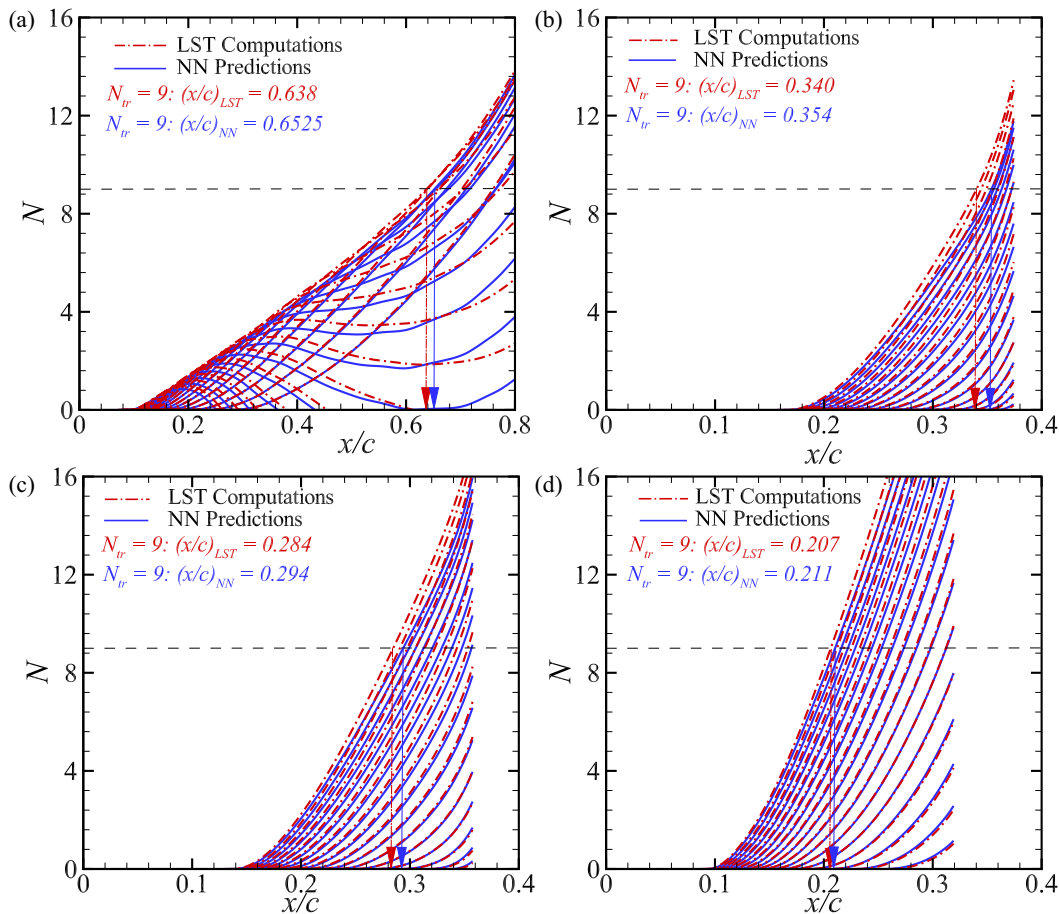
FIG. 3. *N*-factor curves for non-self-similar boundary-layer profiles over the upper surface of a symmetric HSK airfoil section and an asymmetrical NLF-0416 airfoil at different angles of attack (α). Transition location corresponds to the critical value of $N = 9$ (marked by a dashed line). Corresponding transition onset locations are mentioned on the upper left corner and marked on the horizontal axis as predicted by the proposed neural network C_1 model (blue arrow) and computed by linear stability theory (LST) (red arrow). The neural network model was trained using the Falkner-Skan database (with self-similar boundary layers). Red and blue lines correspond to *N*-factor curves. (a) HSK airfoil, $\alpha = 0°$, (b) NLF-0416, $\alpha = 0°$, (c) NLF-0416, $\alpha = 2°$, (d) NLF-0416, $\alpha = 5°$.

collectively cover both favorable and adverse pressure gradients. For this assessment, the proposed neural network model has been trained on the complete Falkner-Skan database while the testing dataset corresponds to the upper surface of the airfoil sections. Unlike the Falkner-Skan database, the boundary-layer flows on these airfoils evolve in a non-self-similar manner. Therefore, the evaluation of the neural network trained on the database of self-similar profiles allows us to gauge the practical utility of this model, i.e., the capacity to generalize the predictions to arbitrary, but still attached, boundary-layer profiles. Figure 3 shows the corresponding results where the predicted *N*-factor curves for instability waves with a selected set of disturbance frequencies have been superposed on those based on the linear stability theory (LST). The predicted *N*-factor curves are computed by integrating the local amplification rates σ predicted by the proposed neural network model. Figure 3(a) shows the comparison plots for the symmetric HSK airfoil section wherein the abscissa corresponds to the surface location along the upper surface of the airfoil section (scaled with respect

to the airfoil chord length) and the N-factor values are plotted along the ordinate. As discussed in Sec. II A, transition is predicted to occur when the value of N reaches a critical value, which is chosen to be $N_{tr} = 9$ for the purpose of this comparison [9,10]. This critical value of N is marked by a dashed line in the plot, whereas the corresponding transition locations predicted by the neural network and the LST are indicated by blue and red arrows, respectively, and are also listed in the upper left portion of the figure. The error in the neural network prediction for the transition location based on $N_{tr} = 9$ is approximately 2% for the HSK airfoil. Similar results are shown in Figs. 3(b)–3(d) for the predicted transition location along the suction surface of the asymmetric NLF-0416 airfoil section at selected angles of attacks. For these cases, the transition onset location has been predicted to within an error of approximately 2–5 % by the neural network model.

B. Automatic, data-driven feature extraction of boundary layer profiles

The primary advantage of the proposed neural network is the feature extraction capability of the CNN (Fig. 2). The CNN is able to distill information from the boundary-layer profiles in a physically consistent manner, i.e., by considering the boundary-layer profiles as continuous functions and by encoding the information into a set of parameters indicated by the (green) shaded neurons Ψ in Fig. 2(b). The CNN provides a mapping from the space of boundary-layer profiles to the physical parameter space in an automated, data-driven fashion, i.e., without requiring the user to specify an explicitly defined learning target for the CNN. To demonstrate this capability of the proposed neural network, we consider the case in which the CNN maps the distilled information from the boundary-layer profile $U_j(j = 1, 2, \ldots, 41)$ to a single parameter Ψ at the interface between the CNN and the fully connected network in Fig. 2. In essence, this process mimics the behavior of the fully connected network A from Table II by choosing the CNN parameters to encode a single feature from the boundary-layer profiles. Because the CNN does not make any prior assumptions about what this single feature should be, one might expect that the predictive performance of the CNN architecture C_1^* would be better than that of the fully connected network A. Somewhat surprisingly, however, comparison of the respective testing errors for the asymmetrical NLF-0416 airfoil indicates that the network A performs slightly better than C_1^* (testing error of 21.2% versus 24.5%). The explanation of this relative performance is left as a topic for future studies. However, it does seem to provide independent evidence that supports the practice of using the analytically defined shape factor H as a nearly optimal scalar representation of the boundary-layer profile for the purpose of predicting the amplification rates. Next, we evaluated the parameter Ψ for each of the 111 boundary-layer profiles along the upper surface of the asymmetrical NLF-0416 airfoil by using the convolutional neural network model that was trained on the full Falkner-Skan database. For convenience of interpretation, we normalize (i.e., shift and scale) the learned parameter Ψ to $\tilde{\Psi}$ such that the latter falls within the range of [0, 1]. Figure 4 indicates the variation in $\tilde{\Psi}$ with a similarly normalized shape factor $\tilde{H}$ of the velocity profiles that is defined in the inset of the figure. The plot shows a nearly linear relationship between the CNN-extracted feature $\tilde{\Psi}$ and the physically defined counterpart $\tilde{H}$. We point out that, because the activation function for the last convolutional layer is linear, the mapping to the parameter Ψ remains unchanged if we scale all of the weights leading into this layer by an arbitrary factor a and the value of all neurons from the layer by $1/a$. As such, the normalization process is well justified.

A similar evaluation of the data-driven feature extraction capability of the CNN was performed by varying the number of boundary-layer profiles used as input to the neural network. Specifically, the correlation analysis from Fig. 4 was repeated by including the first and second derivatives of the velocity profile in addition to the velocity profile itself. The results of this analysis for all three networks (namely, network C_1^* with U, dU/dy, and d^2U/dy^2 profiles as input, C_2^* with U and dU/dy as input, and C_3^* with U only) are presented and compared in Fig. 5. It can be observed that the correlation between $\tilde{\Psi}$ and $\tilde{H}$ is linear for C_3^*, where only the velocity profile U is used as input. In comparison, a mild nonlinearity may be observed in the $\tilde{\Psi}$-$\tilde{H}$ relations for the C_2^* and C_1^* networks, where the first and second derivatives dU/dy and d^2U/dy^2 of the velocity are
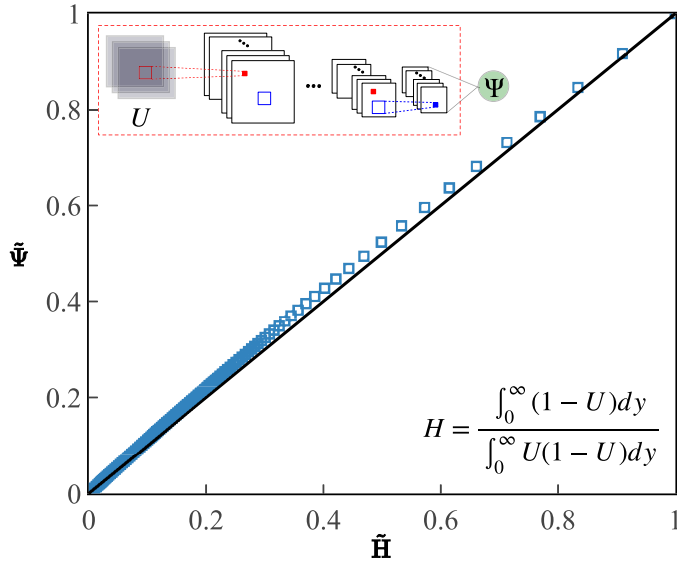
FIG. 4. Correlation between normalized CNN learned parameter ($\tilde{\Psi}$) from non-self-similar boundary-layer profile U and normalized shape factor ($\tilde{H}$), at 111 locations along the upper surface of an asymmetric NLF-0416 airfoil section at 0 degrees angle of attack. The learned parameter Ψ and the shape parameter H are normalized to within the range [0, 1] to facilitate comparison. The Falkner-Skan database (with self-similar boundary layers) has been used for training.
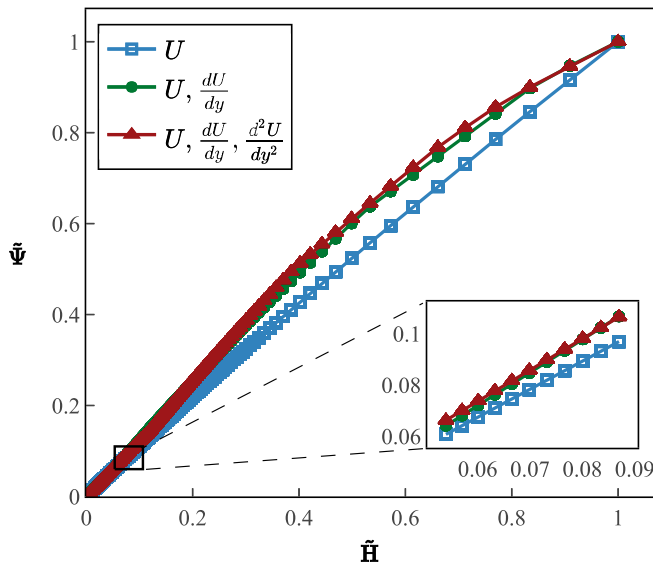

FIG. 5. Comparison of correlation between normalized CNN learned parameter ($\tilde{\Psi}$) for varying sets of non-self-similar boundary layer profiles (corresponding to networks C_1^*, C_2^*, and C_3^*) and normalized shape factor ($\tilde{H}$), at 111 locations along the upper surface of an asymmetric NLF-0416 airfoil section at 0 degrees angle of attack. Falkner-Skan database (with self-similar boundary layers) has been used for training.
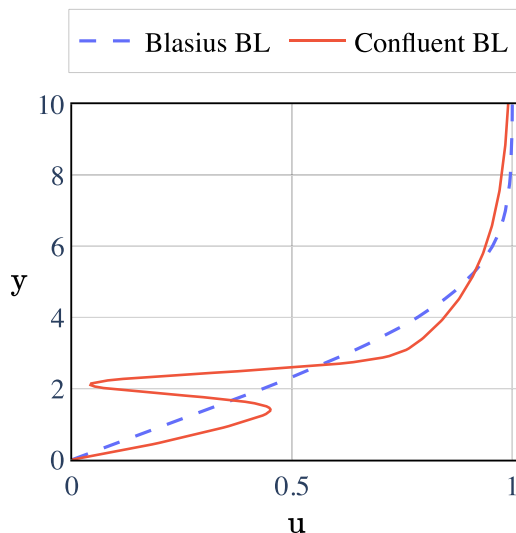
FIG. 6. Comparison of Blasius profile and a confluent boundary layer involving the wake deficit due to an upstream element. The shape factor value of both profiles is 2.59, but the instability characteristics are expected to be very different from each other.

also introduced as additional input features to the CNN. However, whether the relation is linear or nonlinear does not have any major consequence for our purpose. What is more important is that all three networks exhibit a one-to-one correspondence between the CNN-extracted quantity Ψ from the boundary-layer profiles and the physically defined quantity $\tilde{H}$. Thus, we may conclude that the proposed CNN architecture has a robust performance in goal-oriented feature extraction. In particular, without any explicit instruction from the user, the CNN has been able to encode the boundary-layer profiles (i.e., velocity and its derivatives) into a quantity Ψ that is predictive of the amplification rate σ when used in conjunction with the other physical parameters ω and Re_{θ}. The quantity Ψ is strongly correlated to the shape factor H, which is known to be correlated with the stability characteristics [13]. This observation points to the physically consistent nature of the proposed neural network architecture. Furthermore, profiles with substantially different stability characteristics can have the same integral shape factor, and therefore the shape factor alone cannot be an adequate predictor of the stability characteristics. As an example, Fig. 6 presents a comparison between the self-similar Blasius boundary-layer profile and a confluent boundary-layer profile with a wake deficit, which may be encountered in the case of boundary layers over multielement airfoil configurations. The integral shape factor value for both profiles is 2.59; however, the stability characteristics of both profiles are expected to be substantially different. The Blasius profile would have the viscous-inviscid interactive instability concentrated closer to the wall, whereas the confluent boundary-layer profile would have primarily inviscid instability within the wake deficit region. However, both profiles would appear the same in the shape factor space. In comparison, the proposed CNN model would use the well-resolved boundary-layer profile to capture the differences between the two profiles toward the prediction of their instability characteristics. In addition to the theoretical importance, the results also have practical importance. Specifically, the results above highlight the intrinsic potential of the proposed network to allow additional features of the boundary-layer profiles to be taken into account with manageable computational costs and without compromising the robustness of the network performance.

Admittedly, the shape factor H of the Falkner-Skan boundary layers can be easily evaluated and then used to map the complex dependence of the disturbance amplification rate on the underlying mean flow through a fully connected neural network [Fig. 1(a)]. However, in several other cases,

such as high-speed flows over blunt nose configurations or flows where the edge of the boundary layer cannot be easily determined, the shape factor H cannot be defined in a consistent and accurate manner and/or computed straightforwardly. References [42,43] highlight some of the difficulties in determining the boundary-layer edge, and hence in computing the shape factor of both transonic [42] and hypersonic [43] laminar boundary-layer flows encountered in practical applications. In such cases, a numerical solution to the full Navier-Stokes equations must be used to compute the boundary-layer flow. For these flows, the proposed convolutional neural network provides a more general and effective architecture for modeling the local instability amplification rates based on the boundary-layer profiles and other relevant physical parameters. Such capability has been demonstrated in a companion paper [44] for the case of a blunt-nosed body in hypersonic flows where, irrespective of the existence of a shape factor (H), the CNN model has been able to predict local instability amplification rates accurately. Although a reduced representation of the input features can also be achieved via dimensionality reduction techniques, such as principal component analysis [20], the proposed neural network architecture provides an easier technique to encode the targeted information from boundary-layer profiles into a smaller set of parameters. This feature extraction capability of the CNN is likely to assume an even greater significance for (i) three-dimensional boundary-layer profiles involving the additional crossflow velocity component, (ii) high-speed flows that involve the profiles of thermodynamic quantities such as the density and/or temperature, and (iii) boundary-layer flows that are inhomogeneous in two spatial coordinates instead of just the wall-normal coordinate, e.g., planar boundary-layer profiles that vary along both the wall-normal and the spanwise directions. One expects that more complex flows such as three-dimensional boundary layers are likely to require increasingly long feature vectors as compared to the Falkner-Skan boundary layers where even a scalar representation of Ψ provided satisfactory results.

Finally, we emphasize that while the shape factor is known to work well for predicting the growth characteristics of TS waves, we emphasize that the proposal to use the shape factor was completely ad hoc and based on domain expertise. The shape factor does not appear anywhere in the mathematical statement of the stability theory, or for that matter in the analytic solutions based on high Reynolds number asymptotic theories of TS waves. The present work presents completely independent evidence extracted from data that the shape factor is a nearly optimal scalar predictor of TS growth rates and how one may develop additional predictors of instability growth rates. Feature extraction capabilities of translation invariant images have been demonstrated in recent work in computer vision and even geosciences [31]; however, the principle of translation invariance does not apply to the boundary-layer velocity profiles, and the success of the CNN as demonstrated here is by no means a foregone conclusion.

C. Comparison to a fully connected network with profile inputs

The proposed convolutional neural network has also been assessed against the straightforward method of directly introducing the full boundary-layer profiles as input features to a fully connected neural network, as shown in Fig. 1(b). The latter architecture may be viewed as a generalization of the architecture in Ref. [24]. Figure 7 shows a comparison between the predictive performances of the proposed convolutional neural network (Fig. 2) and the fully connected network [Fig. 1(b)]. Both neural network models have been trained using the stability database for Falkner-Skan profiles, and the predictive performance is evaluated for instability amplification over the upper surface of the HSK airfoil section. To help ensure a fair comparison, the number of model parameters for both models is kept approximately equal as given in Table I for networks C_1 and B, respectively. The predicted transition onset location based on each model is marked by an arrow on the x-axis and mentioned below the legends in Figs. 7(b) and 7(d). For reference, the predictions based on direct stability computations are also shown. Although the validation plots for instability amplification rates in Figs. 7(a) and 7(c) show better predictive performance for the convolutional neural network, the prediction of transition onset location is better for the fully connected network. The transition
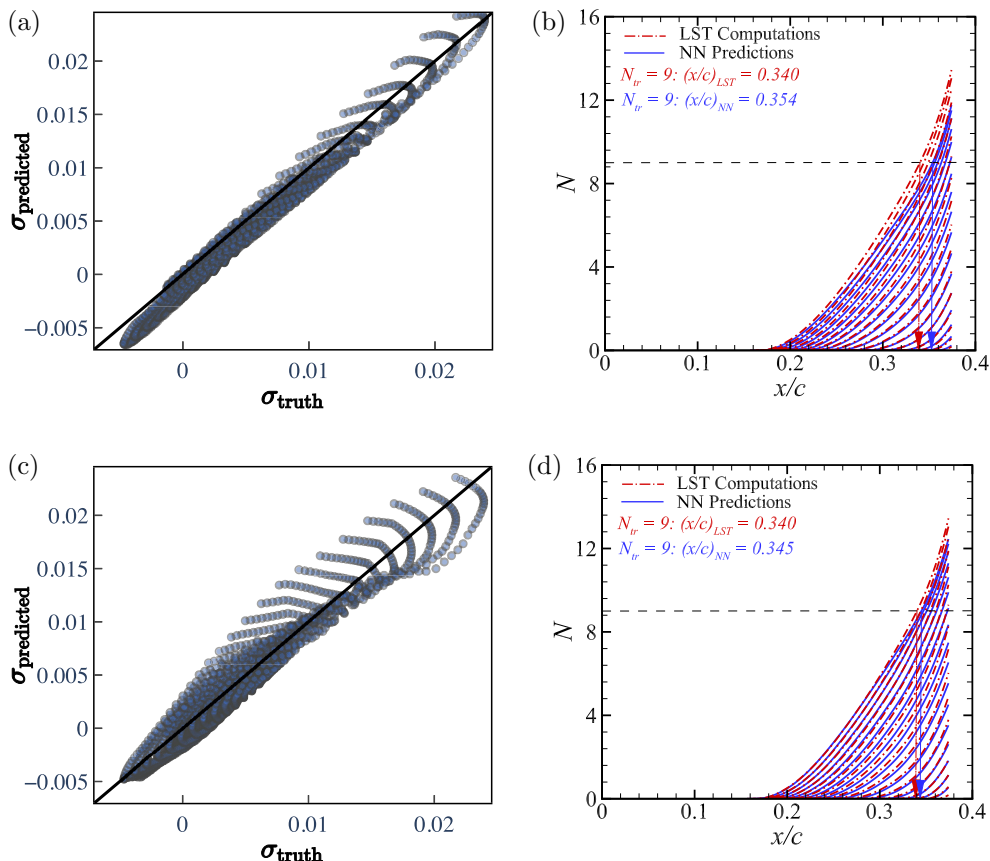
FIG. 7. Comparison between fully connected network B and proposed convolutional neural network C_1. Validation plots of local instability amplification rates σ and N-factor curves for an asymmetric NLF-0416 airfoil section at 0 degrees angle of attack are given for both networks. Transition location corresponds to the critical value of $N = 9$ (marked by a dashed line). Corresponding transition onset locations are mentioned below the legends in (b) and (d) and marked on the horizontal axis as predicted by the neural network model (NN) (blue arrow) and computed by linear stability theory (LST) (red arrow). (a) Validation plot, growth rate predictions for CNN C_1, (b) N-factor curves, CNN C_1, (c) Validation plot, growth rate predictions for fully connected network B and (d) N-factor curves, fully connected network B.

location predicted by the fully connected network is within 1.1% of the transition onset location based on the linear stability theory, whereas the convolutional neural network predicts the same with a 2.3% error. Even though both measures of error are rather small, the qualitative trend is somewhat unexpected and requires further investigation in future studies.

One significant advantage of the proposed convolutional neural network (Fig. 2) over the fully connected network [Fig. 1(b)] pertains to the number of trainable model parameters required to achieve a comparable performance. With full information available as input features, a fully connected network would require a significantly higher number of model parameters, and consequently a higher training cost to provide a comparable performance, whereas the convolutional neural network is likely to provide more robust predictive performance with a smaller number of model parameters. Figure 8 presents the results of the analysis designed to verify this behavior. Here, the test error percentage has been plotted against the size of the neural network model as measured by the total number of trainable model parameters. We observe that the performance of
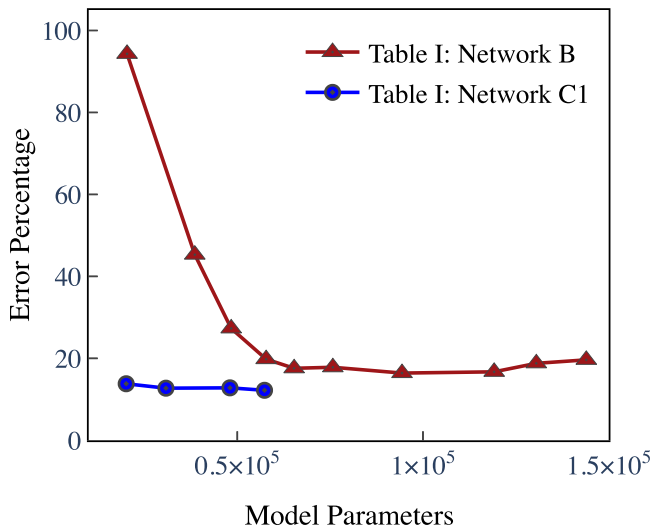
FIG. 8. Comparison of testing error for the fully connected neural network (FC NN) [Fig. 1(b), Network B from Table I] and the convolutional neural network (Fig. 2, Network C_1 from Table I) as a function of the total number of learnable model parameters. Input features for both models include boundary-layer profiles. The Falkner-Skan database (with self-similar boundary layers) was used for training while the testing error was evaluated for an asymmetric NLF-0416 airfoil section with non-self-similar boundary layers at 0 degrees angle of attack.

the fully connected neural network deteriorates significantly as the number of model parameters is reduced, while the proposed convolution-based neural network is able to maintain its predictive performance. The convolutional neural network in the proposed model encodes the information from the boundary-layer profiles into a significantly smaller number of scalar parameters (eight parameters for network C_3) before folding them into the ensuing fully connected portion of the overall network, along with the other physical parameters (namely, frequency of the instability wave ω and the local Reynolds number Re_θ). Such feature engineering for the boundary-layer profiles via the convolutional neural network enables one to achieve a comparable performance with a significantly reduced number of model parameters and training cost.

IV. CONCLUSION

A neural-network-based transition model has been presented that is capable of accurately predicting the transition onset location for incompressible, two-dimensional attached flows in a physically informed manner without requiring the direct computations using the linear stability theory. The proposed model has the ability to encode information, using convolutional layers, from boundary-layer profiles (velocity and its derivatives) into a set of integral quantities. More importantly, the encoded feature $\hat{\Psi}$ shows strong correlation with, or even one-to-one correspondence to the physically defined shape parameter H, which clearly demonstrates the physically consistent nature of the proposed neural network. These encoded integral quantities are then nonlinearly mapped to local instability amplification rate σ along with other scalar disturbance characteristics (ω and Re_θ). The proposed model is shown to have robust predictive performance, clear physical interpretation, and superior computational efficiency.

The CNN architecture presented herein can be easily generalized to other instability mechanisms and, in follow-on (and as yet unpublished work), we have demonstrated the application of this architecture to second mode instabilities in high-speed boundary layers that cannot be predicted well on the basis of local shape factors of the boundary-layer profiles. Thus, it could become the means

for physics-based transition prediction in practical applications of computational fluid dynamics codes.

ACKNOWLEDGMENTS

[1] J. E. Green, Air Travel–Greener by Design Mitigating the environmental impact of aviation: Opportunities and priorities, Aeronaut. J. **109**, 361 (2005).

[2] L. M. Mack, Review of linear compressible stability theory, in *Stability of Time Dependent and Spatially Varying Flows*, edited by D. L. Dwoyer and M. Y. Hussaini, ICASE NASA LaRC Series (Springer, New York, 1987).

[3] H. L. Reed, W. S. Saric, and D. Arnal, Linear stability theory applied to boundary layers, Annu. Rev. Fluid Mech. **28**, 389 (1996).

[4] M. P. Juniper, A. Hanifi, and V. Theofilis, Modal stability theory: Lecture notes from the flow-nordita summer school on advanced instability methods for complex flows, Appl. Mech. Rev. **66**, 024804 (2014).

[5] K. Taira, S. L. Brunton, S. T. M. Dawson, C. W. Rowley, T. Colonius, B. J. McKeon, O. T. Schmidt, S. Gordeyev, V. Theofilis, and L. S. Ukeiley, Modal analysis of fluid flows: An overview, AIAA J. **55**, 4013 (2017).

[6] D. Arnal, Boundary layer transition: Prediction based on linear stability theory, in *Special Course on Progress in Transition Modeling, AGARD-R-793* (AGARD, Neuilly-sur-Seine, France, 1994), pp. 1–63.

[7] L. M. Mack, Transition prediction and linear stability theory, in *Special Course on Stability and Transition of Laminar Flows, AGARD R-709* (AGARD, Neuilly-sur-Seine, France, 1984), pp. 1–232.

[8] E. Reshotko, Boundary-layer stability and transition, Annu. Rev. Fluid Mech. **8**, 311 (1976).

[9] J. van Ingen, The e^n method for transition prediction, Historical review of work at TU Delft, in *38th Fluid Dynamics Conference and Exhibit, AIAA 2008-3830* (AIAA, Seattle, 2008), Vol. 1, pp. 1–49.

[10] D. M. Bushnell, M. R. Malik, and W. D. Harvey, Transition prediction in external flows via linear stability theory, in *Symposium Transsonicum III*, edited by J. Zierep and H. Oertel (Springer, Berlin, 1989), pp. 225–242.

[11] K. Kusunose, L. Wigton, and P. Meredith, A rapidly converging viscous/inviscid coupling code for multi-element airfoil configurations, in *29th Aerospace Sciences Meeting, AIAA-91-0177* (AIAA, Seattle, 1991), Vol. 1, pp. 1–11.

[12] C. Sheng, Role of transition modeling in rotor hover predictions, J. Aircraft **55**, 23 (2018).

[13] M. Drela and M. B. Giles, Viscous-inviscid analysis of transonic and low Reynolds number airfoils, AIAA J. **25**, 1347 (1987).

[14] J. P. Perraud and A. Durant, Stability-based Mach zero to four longitudinal transition prediction criterion, J. Spacecr. Rockets **53**, 730 (2016).

[15] J. R. Dagenhart, Amplified cross flow disturbances in the Laminar layer on swept wings with suction, Technical Report No. NASA-TP-1902, NASA Langley Research Center, 1981.

[16] H. W. Stock and E. Degenhart, A simplified e^n method for transition prediction in two-dimensional, incompressible boundary layers, Z. Flugwissen. Weltraumforsch. **13**, 16 (1989).

[17] M. Gaster and F. Jiang, Rapid scheme for estimating transition on wings by linear stability theory, in *Proceedings ICAS International Council of the Aeronautical Sciences* (ICAS, Anaheim, CA, 1995), Vol. 3, pp. 1104–1113.

[18] M. Langlois, C. Masson, F. Kafyeke, and I. Paraschivoiu, Automated method for transition prediction on wings in transonic flows, J. Aircraft **39**, 460 (2002).

[19] A. Krumbein, e^n transition prediction for 3D wing configurations using database methods and a local, linear stability code, Aerosp. Sci. Technol. **12**, 592 (2008).

[20] D. G. Rajnarayan and P. Sturdza, Extensible rapid transition prediction for aircraft conceptual design using modal decomposition, in *51st AIAA Aerospace Sciences Meeting including the New Horizons Forum and Aerospace Exposition, AIAA 2013-0231* (AIAA, Grapevine, TX, 2013), Vol. 1, pp. 1–21.

[21] G. Begou, H. Deniau, O. Vermeersch, and G. Casalis, Database approach for laminar-turbulent transition prediction: Navier-Stokes compatible reformulation, AIAA J. **55**, 3648 (2017).

[22] F. Pinna, L. Zanus, S. Demange, and M. Olazabal-Loume, Reduced model for transition prediction in hypersonic flows, in *2018 Fluid Dynamics Conference, AIAA 2018-3697* (AIAA, Atlanta, GA, 2018), Vol. 1, pp. 1–10.

[23] J. G. Coder and M. D. Maughmer, Computational fluid dynamics compatible transition modeling using an amplification factor transport equation, AIAA J. **52**, 2506 (2014).

[24] J. D. Crouch, I. W. M. Crouch, and L. L. Ng, Transition prediction for three-dimensional boundary layers in computational fluid dynamics applications, AIAA J. **40**, 1536 (2002).

[25] P. Paredes, B. Venkatachari, M. Choudhari, F. Li, C.-L. Chang, M. I. Zafar, and H. Xiao, Extensible rapid transition prediction for aircraft conceptual design using modal decomposition, in *AIAA SciTech Forum* (AIAA, Orlando, FL, 2020), Vol. 1, pp. 1–21.

[26] R. M. Fuller, W. R. Saunders, and U. Vandsburger, Neural network estimation of disturbance growth using a linear stability numerical model, in *35th Aerospace Sciences Meeting and Exhibit*, AIAA-97-0559 (AIAA, Reno, 1997), Vol. 1, pp. 1–10.

[27] F. Danvin, F. Pinna, and M. Olazabal-Loume, Laminar to turbulent transition prediction in hypersonic flows with meta-models, in *2018 Fluid Dynamics Conference, AIAA-2018-3701* (AIAA, Atlanta, 2018), Vol. 1, pp. 1–18.

[28] P. G. Drazin and W. H. Reid, *Hydrodynamic Stability* (Cambridge University Press, Cambridge, 1981).

[29] K. Hornik, M. Stinchcombe, and H. White, Multilayer feedforward networks are universal approximators, Neural Netw. **2**, 359 (1989).

[30] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).

[31] J. Wu, X. Yin, and H. Xiao, Seeing permeability from images: fast prediction with convolutional neural networks, Sci. Bull. **63**, 1215 (2018).

[32] D. P. Kingma and L. J. Ba, Adam: A method for stochastic optimization, in *3rd International Conference for Learning Representations* (Ithaca, NY, 2015).

[33] C.-L. Chang, Langley Stability and Transition Analysis Code (LASTRAC) Version 1.2 User Manual, NASA/TM-2004-213233 (2004).

[34] F. T. Smith, On the non-parallel flow stability of the Blasius boundary layer, Proc. R. Soc. London A **366**, 91 (1979).

[35] H. Kanner, J. Schetz, and R. Wlezien, The evolution of an acoustic disturbance up to transition in the boundary layer on an airfoil, in *30th Fluid Dynamics Conference, AIAA-99-3791* (AIAA, Norfolk, VA, 1999), Vol. 1, pp. 1–11.

[36] D. M. Somers, Design and experimental results for a natural–laminar-flow airfoil for general aviation applications, Technical Report No. NASA-TP-1861, NASA Langley Research Center, 1981.

[37] A. Gopalarathnam and M. S. Selig, Low-speed natural-laminar-flow airfoils: Case study in inverse airfoil design, J. Aircraft **38**, 57 (2001).

[38] M. Raissi, P. Perdikaris, and G. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, J. Comput. Phys. **378**, 686 (2019).

[39] J. Ling, A. Kurzawski, and J. Templeton, Reynolds averaged turbulence modeling using deep neural networks with embedded invariance, J. Fluid Mech. **807**, 155 (2016).

[40] J. Wang, J. L. Wu, and H. Xiao, Physics informed machine learning approach for reconstructing Reynolds stress modeling discrepancies based on DNS data, Phys. Rev. Fluids **2**, 034603 (2017).

[41] J. L. Wu, H. Xiao, and E. G. Paterson, Physics-informed machine learning approach for augmenting turbulence models: A comprehensive frame-work, Phys. Rev. Fluids **3**, 074602 (2018).

[42] J. Schneider and B. Ewald, Calculation of transonic laminar flow airfoils using a Navier-Stokes method and linear stability theory, in *Proceedings of 1994 International Council of the Aeronautical Sciences*, Anaheim ICAS-94-4.7.2, 2456 (AIAA, Anaheim, 1994).

[43] D. A. Saunders and D. K. Prabhu, Blayer user guide, NASA/TM|—2018-219749, NASA Ames Research Center, 2018.

[44] P. Paredes, B. Venkatachari, M. Choudhari, F. Li, C.-L. Chang, M. I. Zafar, and H. Xiao, Toward a practical method for hypersonic transition prediction based on stability correlations, AIAA J. **58**, 4475 (2020).