

Nonlinearly most dangerous disturbance for high-speed boundary-layer transition

Reza Jahanbakhshi¹ and Tamer A. Zaki^{1,†}

¹Department of Mechanical Engineering, Johns Hopkins University, Baltimore, MD 21218-2682, USA

(Received 8 November 2018; revised 24 June 2019; accepted 24 June 2019;
first published online 31 July 2019)

Laminar-to-turbulent transition in a zero-pressure-gradient boundary layer at Mach 4.5 is studied using direct numerical simulations. For a given level of total disturbance energy, the inflow spectrum was designed to correspond to the nonlinearly most dangerous condition that leads to the earliest possible transition Reynolds number. The synthesis of the inlet disturbance is formulated as a constrained optimization, where the control vector is comprised of the amplitudes and relative phases of the inlet modes; the constraints are the prescribed total energy and that the flow evolution satisfies the full nonlinear compressible Navier–Stokes equations; the cost function is defined in terms of the mean skin-friction coefficient and, once maximized, ensures the earliest possible transition location. An ensemble-variational (EnVar) technique is developed to solve the optimization problem. Starting from an initial guess, here a broadband disturbance, EnVar updates the estimate of the control vector at the end of each iteration using the gradient of the cost function, which is evaluated from the outcomes of an ensemble of possible solutions. Two inflow conditions are computed, each corresponding to a different level of energy, and their spectra are contrasted: the lower-energy case includes two normal acoustic waves and one oblique vorticity perturbation, whereas the higher-energy condition consists of oblique acoustic and vorticity waves. The focus is placed on the former case because it cannot be categorized as any of the classical breakdown scenarios (fundamental, subharmonic or oblique), while the higher-energy condition undergoes a second-mode oblique transition. At the lower energy level, the instability wave that initiates the rapid breakdown to turbulence is not present at the inlet plane. Instead, it appears at a downstream location after a series of nonlinear interactions that spur the fastest onset of turbulence. The results from the nonlinearly most potent inflow condition are also compared to other inlet disturbances that are selected solely based on linear theory, and which all yield relatively delayed transition onset.

Key words: instability, transition to turbulence, compressible boundary layers

1. Introduction

The location and mechanism of laminar-to-turbulence transition in hypersonic boundary layers are sensitive to the flow environment. Linear theory provides criteria for the onset of exponential instability, and these primary waves subsequently undergo

† Email address for correspondence: t.zaki@jhu.edu

secondary instability and breakdown to turbulence. However, similar criteria for transition onset cannot be established theoretically because breakdown to turbulence is nonlinear. Nonlinear simulations can predict the evolution of any inflow perturbation, although starting with the linearly most unstable mode does not guarantee the earliest possible transition location. And when the inflow perturbation is composed of multiple waves, the choice of their relative amplitudes and phases can appreciably influence transition onset. These parameters are herein optimized using an ensemble-variational (EnVar) algorithm in order to determine the earliest possible transition Reynolds number in a zero-pressure-gradient (ZPG) high-speed boundary layer, for a specified level of energy of the inlet disturbance.

1.1. Transition in high-speed boundary layers

The precursors that ultimately lead to transition to turbulence in high-speed boundary layers can be traced upstream to the early amplification of small-amplitude acoustic, entropic or vortical fluctuations (Leyva 2017). When their amplitudes are infinitesimal, these waves and their growth rates are accurately modelled by linear stability theory (LST) (Mack 1984). According to the LST, oblique first-mode instabilities, which are similar to the Tollmien–Schlichting waves at lower speeds, are most amplified in supersonic conditions. With increasing Mach number, however, two-dimensional acoustic instability waves, which are commonly referred to as Mack’s (second) modes (Mack 1984), play an increasingly important role. Secondary instability theory was also developed (Herbert 1988) in order to examine how the state comprised of the mean-flow profile and the primary instability will itself become unstable to new perturbations. In order to account for non-parallel effects on the amplification of instability waves, the linear parabolized stability equations were introduced by Bertolotti, Herbert & Spalart (1992). In addition, when the disturbance amplitudes become appreciable, the nonlinear parabolized stability equations (Bertolotti *et al.* 1992; Chang & Malik 1994; Herbert 1997) can account for nonlinear modal interactions and the base-state distortion. The final stage of transition is marked by the appearance of localized turbulence patches, or spots, which spread as they are advected downstream. Intermittency, which is defined as the fraction of time that the flow at a given streamwise location is turbulent (Narasimha 1985), thus rises from its initial value of zero in the laminar regime to unity where the spots merge to form a fully turbulent boundary layer. Recent numerical and experimental studies have focused on advancing our fundamental understanding of transition (see the reviews by Fedorov 2011; Zhong & Wang 2012; Schneider 2015). A summary of select efforts is provided here, with the objective of highlighting transition mechanisms that are relevant to boundary layers at free-stream Mach numbers larger than four.

A key determining factor of the transition mechanism, be that in simulations or in experiments, is the spectral makeup of the upstream disturbance. For example, in numerical simulations, this inflow disturbance must be specified, and changes in the relative amplitudes and phases of inflow waves can lead to quantitative and qualitative changes in the transition process. Franko & Lele (2013, 2014) performed direct numerical simulations of three transition mechanisms at Mach six, in ZPG and adverse-pressure-gradient (APG) boundary layers: first-mode oblique breakdown, second-mode oblique breakdown and second-mode fundamental resonance. In all cases, the disturbance frequencies and spanwise wavenumbers were chosen based on LST and the e^N method. They concluded that, for all three mechanisms, APG increases the linear growth rates and promotes earlier breakdown to turbulence.

Nonetheless, the nonlinear breakdown of each of the three mechanisms was qualitatively similar for ZPG and APG conditions. Sivasubramanian & Fasel (2015, 2016) also performed direct numerical simulations of fundamental resonance and of oblique breakdown in a Mach 6 boundary layer on a sharp cone, with a half-vertex angle equal to 7° . A series of low-resolution simulations, informed by LST, were used to identify the instability waves that promote early transition. These waves are then used to perform fully resolved simulations. They reported that both second-mode fundamental and oblique breakdown lead to strong nonlinear interactions, and are viable candidates to affect transition on the cone geometry. Novikov & Egorov (2016) performed a numerical study of laminar–turbulent transition over a flat plate at Mach 5.37. In that effort, perturbations were introduced into the boundary layer through a small round hole on the surface of the plate by forced pulsations of the vertical velocity component. The authors reported that the linear stages of transition are dominated by first-mode oblique waves, while second-mode plain waves become dominant in nonlinear regions. Most recently, Zhao *et al.* (2018) studied the effect of local wall heating and cooling on the stability of a flat-plate boundary layer at Mach 6. Their results revealed that the location of thermal treatment relative to the synchronization point of the slow and the fast modes (Fedorov 2011) significantly affects the stability of a second-mode instability. These studies highlight the variety of transition mechanisms that can be explored using nonlinear computations, and by varying the inflow condition.

In a Mach 6 wind tunnel, Zhang, Tang & Lee (2013), Zhang *et al.* (2015) and Zhu *et al.* (2016, 2018) investigated transition on a flared cone geometry using Rayleigh scattering visualization, fast-response pressure measurements and particle image velocimetry. They found that the second-mode instability is a key modulator of the transition process, through the formation of high-frequency vortical waves and triggering a fast breakdown to turbulence. They also reported that the second-mode waves, accompanied by high-frequency alternating fluid compression and expansion (dilatation), produce intense aerodynamic heating in a small region. Casper *et al.* (2016) studied boundary-layer transition on a sharp seven-degree cone in hypersonic wind tunnels at Mach 5, 6, 8 and 14. At the lowest Mach number, transition was initiated by a combination of first- and second-mode instabilities, while at higher values the boundary layer was dominated by second-mode instabilities. Laurence, Wagner & Hannemann (2016) and Kennedy *et al.* (2018) also investigated transition on a slender cone at Mach 6–7 and Mach 14, respectively, using high-speed schlieren visualizations. They specifically focused on formation and evolution of second-mode instabilities, and reported that the wave packets form rope-like structures that become progressively more interwoven as transition develops. The experiments were performed for both low- and high-enthalpy conditions, and highlighted the difference in the wall-normal distribution of the disturbances in each condition. Despite laboratory experiments accounting for all stages of transition fully, starting from receptivity through the development of the turbulent boundary layer, they do not necessarily reproduce the environmental disturbances of high-altitude flight. As a result, flight experiments have been performed, and more than 20 such tests were surveyed by Schneider (1999). The author noted the high cost of each experiment which often precludes repeating the tests. More recently, the Hypersonic International Flight Research Experimentation (HIFiRE) program performed a series of tests to study transition scenarios on a straight cone (HIFiRE-1; see e.g. Stanfield *et al.* 2015) and on an elliptic cone (HIFiRE-5; see e.g. Juliano, Adamczak & Kimmel 2015; Kimmel *et al.* 2018) in hypersonic flights. The HIFiRE data show that, depending on the

flow condition and angle of attack, second-mode-, cross-flow- and separation-induced transitions are all potential causes of turbulence on different parts of the vehicle. It is important to note that only a limited amount of flight data on transition are available and that in-flight environmental conditions are variable and uncertain.

1.2. *The nonlinearly most dangerous disturbance*

With a wealth of instability waves at hypersonic speeds, transition mechanisms in numerical simulations can vary quantitatively and qualitatively depending on the inflow disturbance spectrum. Existing approaches synthesize the inflow as a superposition of instability waves that are often selected based on their growth rates or appearance in experiments. The former approach does not guarantee that these waves are the most dangerous with respect to the onset of nonlinear breakdown, and relative modal amplitudes and phases are not prescribed in a manner that exposes the most dangerous transition scenario. When motivated by experimental observations, the choice of the computational inflow condition attempts to reproduce the instability waves that were experimentally measurable; precursor interactions that may have led to the generation of these modes may not be included. In addition, inflow conditions that are relevant to in-flight environments are often uncertain or unknown. These challenges motivate the present work and the question: how can we guarantee robust flow design in high-speed flows? The adopted approach is to determine, for a given level of disturbance energy, the nonlinearly most dangerous inflow condition that will cause transition at the lowest possible Reynolds number. No other inflow disturbance with the same energy can cause earlier breakdown to turbulence.

Compared to theory and nonlinear simulations of boundary-layer stability, studies of the nonlinearly most dangerous disturbances in transitional flows are relatively recent (Pringle & Kerswell 2010; Cherubini *et al.* 2011; Rabin, Caulfield & Kerswell 2012). The limited number of previous efforts all focused on incompressible flows; the present study is the first to consider high-speed flow. Another common feature in earlier efforts is the use of adjoint-variational techniques, which are commonly adopted in flow control (Bewley, Moin & Temam 2001; Cherubini, Robinet & De Palma 2013; Luchini & Bottaro 2014; Xiao & Papadakis 2017), in order to optimize a control vector that is either the initial or inflow disturbance – the optimal value is the nonlinearly most dangerous disturbance. The two main ingredients of the algorithm are a cost function whose optimality corresponds to the earliest transition location, and iterative numerical solution of the forward and adjoint equations to evaluate the local gradient of this cost function with respect to the control vector. One key advantage of adjoint methods is that a forward-adjoint loop directly yields the local sensitivity of the cost function to the control vector. However, the adjoint approach has a number of limitations that are relevant for the present purposes. Firstly, an accurate adjoint model is required and is not always available, for example in the case of nonlinear parabolized stability equations. Secondly, when the forward equations are nonlinear, the adjoint model depends on the time history of the forward state variables which must therefore be stored with full or high spatial and temporal resolution. Lastly, in chaotic systems it is very difficult to accurately satisfy the forward-adjoint duality relation over long time horizons and, as a result, forward-adjoint loops are not guaranteed to provide accurate or convergent gradients of the cost function. As a result, previous efforts have often been restricted to short optimization horizons (e.g. Xiao & Papadakis 2017).

In order to avoid some of the limitations of adjoint-variational methods, we developed an EnVar algorithm to evaluate the nonlinearly most dangerous inflow

disturbance. This technique does not require an adjoint solver and is therefore applicable to any forward model, e.g. direct and large-eddy simulations, nonlinear parabolized stability equations, etc. In addition, without the requirement of an adjoint, the difficulties associated with storage requirements and the stability of the adjoint integration are no longer relevant. Applications of EnVar in fluid mechanics are recent, and are generally in the field of data assimilation (Colburn, Cessna & Bewley 2011; Suzuki 2012; Kato *et al.* 2015; Yang *et al.* 2015; Mons *et al.* 2016; Gao *et al.* 2017). The current effort is the first to develop EnVar methods for computing the nonlinearly most dangerous disturbance, and the first to examine these disturbances in hypersonic boundary layers.

In §2, the governing equations and a detailed description of the EnVar algorithm are presented. Section 3 provides a summary of the parameters of the numerical simulations at the two levels of inflow energy. The nonlinearly most dangerous routes to turbulence are discussed in detail in §4, and conclusions are provided in §5.

2. Theoretical formulation

The high-speed flow of a compressible fluid satisfies the Navier–Stokes equations

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{2.1a}$$

$$\frac{\partial \rho \mathbf{u}}{\partial t} + \nabla \cdot (\rho \mathbf{u} \mathbf{u} + p \mathbf{I} - \boldsymbol{\tau}) = 0, \tag{2.1b}$$

and

$$\frac{\partial E}{\partial t} + \nabla \cdot (\mathbf{u}[E + p] + \boldsymbol{\theta} - \mathbf{u} \cdot \boldsymbol{\tau}) = 0, \tag{2.1c}$$

where ρ is the density, \mathbf{u} is the velocity vector, p is the pressure, \mathbf{I} is the unit tensor, $E = \rho e + 0.5 \rho \mathbf{u} \cdot \mathbf{u}$ is the total energy, e is the specific internal energy, $\boldsymbol{\tau}$ is the viscous stress tensor and $\boldsymbol{\theta}$ is the heat flux vector. To close the system of equations, the calorically perfect gas relations are used

$$p = (\gamma - 1) \rho e, \tag{2.2a}$$

and

$$T = \frac{\gamma - 1}{R} e, \tag{2.2b}$$

where T is the temperature, γ is the ratio of specific heats and R is the gas constant. Furthermore, the viscous stress tensor and the heat flux are defined as

$$\boldsymbol{\tau} = \mu_s (\nabla \mathbf{u} + (\nabla \mathbf{u})^T) + \left(\mu_b - \frac{2}{3} \mu_s \right) (\nabla \cdot \mathbf{u}) \mathbf{I}, \tag{2.3}$$

and

$$\boldsymbol{\theta} = -\kappa \nabla T, \tag{2.4}$$

respectively, where μ_s is the dynamic shear viscosity, which is computed from the local temperature via Sutherland’s law (Sutherland 1893), μ_b is the bulk viscosity and κ is the thermal conductivity. The viscosity variables are related via the Stokes

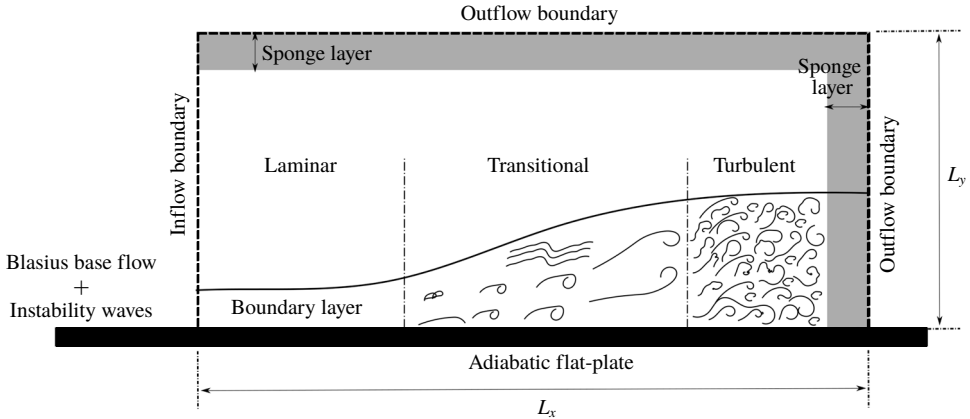


FIGURE 1. Schematic of the computational domain for a ZPG transitional boundary layer over a flat plate.

hypothesis and the thermal conductivity is computed by assuming a constant Prandtl number and specific heat. We will refer to our Navier–Stokes solution algorithm, which is discussed later in § 3, in operator form as \mathcal{N} .

The flow configuration is a ZPG boundary layer over a flat plate, and is shown schematically in figure 1. The computational domain starts downstream of the leading edge, and the inflow condition is a superposition of the compressible Blasius solution and perturbations:

$$\Phi = \Phi_B + \varphi', \tag{2.5}$$

where $\Phi = [\rho \ u \ v \ w \ T]^tr$ are the inflow variables, the subscript B denotes the Blasius base state, $\varphi' = [\rho' \ u' \ v' \ w' \ T']^tr$ are the perturbations with respect to this base state and the superscript tr denotes the transpose. In the present work, the unknown is the most unstable disturbance vector, φ' , for a given level of initial energy.

2.1. Problem definition

The nonlinearly most dangerous disturbance is defined as the inflow perturbation (i) with a prescribed amount of total energy; (ii) that satisfies the Navier–Stokes equations; and (iii) that undergoes the fastest breakdown to turbulence (i.e. shortest streamwise distance to establishing a fully turbulent boundary layer). The first two elements of the definition are the constraints while the third is the objective. This problem can be cast as a constrained optimization: find the control vector φ' that optimizes the cost function

$$\mathcal{J}(\varphi') = \frac{1}{2} \|\mathcal{G}(\mathbf{q})\|^2, \tag{2.6a}$$

while satisfying the constraints

$$\mathbf{q} = \mathcal{N}(\varphi'), \tag{2.6b}$$

and

$$\mathcal{E}(\varphi') = E_0, \tag{2.6c}$$

where $\mathbf{q} = [\rho \ \rho \mathbf{u} \ E]^tr$ is the state vector, $\mathcal{G}(\mathbf{q})$ is the observation vector that quantifies the breakdown to turbulence, \mathcal{N} represents the Navier–Stokes solver, \mathcal{E} is the energy operator and E_0 is the constrained energy. Choices of $\mathcal{G}(\mathbf{q})$ can be formulated based on the skin friction, wall temperature, heat flux or perturbation energy. In the present work, the observation vector is proportional to the skin-friction coefficient, $\mathcal{G}(\mathbf{q}) = C_f(dx/L_x)^{1/2}$. In addition, the total energy of the perturbations, used in constraint (2.6c), is defined as

$$\mathcal{E}(\boldsymbol{\varphi}') = \frac{1}{2} \int_0^{L_y} \left(\bar{\rho}[u^2 + v^2 + w^2] + \frac{R}{\gamma - 1} \left[\frac{\bar{T}}{\bar{\rho}} \rho'^2 + \frac{\bar{\rho}}{\bar{T}} T'^2 \right] \right) dy. \tag{2.7}$$

In the above equation and throughout, the overbar denotes the mean which is evaluated from averaging in time and in the homogeneous spanwise direction.

The control vector is the inflow disturbance, which is expressed as a superposition of instability waves

$$\boldsymbol{\varphi}' = \sum_{n,m} \text{Real} \left\{ c_{n,m} \hat{\boldsymbol{\varphi}}_{n,m}(x_0, y) e^{i(\beta_m z - \omega_n t - \theta_{n,m})} \right\}, \tag{2.8}$$

where ω_n and β_m are the frequency and spanwise wavenumber of each instability mode and $\hat{\boldsymbol{\varphi}}_{n,m} = [\hat{\rho}_{n,m} \ \hat{u}_{n,m} \ \hat{v}_{n,m} \ \hat{w}_{n,m} \ \hat{T}_{n,m}]^tr$ are the complex wall-normal mode shapes which are obtained from solution of the Orr–Sommerfeld–Squire eigenvalue problem. At each (n, m) pair, only the most unstable mode will be included at the inflow, which in the present study will be the slow mode (Fedorov 2011). Since eigenmodes of the Orr–Sommerfeld–Squire problem can have arbitrary amplitudes, they must be appropriately normalized; here we enforce that each mode $\hat{\boldsymbol{\varphi}}_{n,m}$ has unit energy

$$\frac{1}{2} \int_0^{L_y} \left(\bar{\rho}[\hat{u}^* \hat{u} + \hat{v}^* \hat{v} + \hat{w}^* \hat{w}] + \frac{R}{\gamma - 1} \left[\frac{\bar{T}}{\bar{\rho}} \hat{\rho}^* \hat{\rho} + \frac{\bar{\rho}}{\bar{T}} \hat{T}^* \hat{T} \right] \right)_{n,m} dy = 1, \tag{2.9}$$

where star denotes the complex conjugate. It should be emphasized that (2.9) is simply a normalization of the eigenmodes, whose amplitudes in (2.8) are prescribed using the positive real-valued $c_{n,m}$ and their relative phases are $\theta_{n,m}$. Therefore, assuming knowledge of the relevant range of frequencies and spanwise wavenumbers, the only unknowns are $c_{n,m}$ and $\theta_{n,m}$ that lead to the most upstream transition location. The control vector can then be redefined in terms of those two parameters

$$\mathbf{c} = [c_{n,m} \ \dots \ \theta_{n,m} \ \dots]^tr, \tag{2.10}$$

and its size is $2M$, or twice the total number of instability modes. In terms of the new control vector, the perturbation energy (2.7) can be expressed as

$$\mathcal{E}(\boldsymbol{\varphi}') = \frac{1}{2} \mathbf{c}^tr \mathbf{A}_1 \mathbf{c}, \tag{2.11}$$

where

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{bmatrix}_{2M \times 2M}, \tag{2.12}$$

and \mathbf{O} and \mathbf{I} are zero and identity matrices of size $M \times M$, respectively.

Now the constrained optimization problem (2.6) can be rewritten as identifying \mathbf{c} that optimizes the cost function

$$\mathcal{J}(\mathbf{c}) = \frac{1}{2} \|\mathcal{G}(\mathbf{q})\|^2, \quad (2.13a)$$

while satisfying the constraints

$$\mathbf{q} = \mathcal{N}(\mathbf{c}), \quad (2.13b)$$

$$\frac{1}{2} \mathbf{c}^{tr} \mathbf{A}_1 \mathbf{c} = E_0, \quad (2.13c)$$

and

$$\mathbf{A}_1 \mathbf{c} \geq 0. \quad (2.13d)$$

The constraint (2.13d) is introduced because $c_{n,m}$ in (2.8) are real positive values. An EnVar approach is adopted to solve the optimization problem (2.13).

2.2. Ensemble-variational optimization

The optimization problem (2.13) starts with an estimate, or guess, of the solution $\mathbf{c}^{(e)}$. An ensemble of N_{en} variants, $\mathbf{c}^{(r)}$, where $r = 1, \dots, N_{en}$, is also constructed whose mean is equal to $\mathbf{c}^{(e)}$. The optimal control vector is then expressed as the weighted sum

$$\mathbf{c} = \mathbf{c}^{(e)} + \mathbf{E}' \mathbf{a}, \quad (2.14)$$

where \mathbf{E}' is the deviation of the ensemble members from their mean,

$$\mathbf{E}' = [\mathbf{c}^{(1)} - \mathbf{c}^{(e)} \quad \dots \quad \mathbf{c}^{(r)} - \mathbf{c}^{(e)} \quad \dots \quad \mathbf{c}^{(N_{en})} - \mathbf{c}^{(e)}]_{2M \times N_{en}}, \quad (2.15)$$

and \mathbf{a} is the weight vector,

$$\mathbf{a} = [a^{(1)} \quad \dots \quad a^{(r)} \quad \dots \quad a^{(N_{en})}]^{tr}. \quad (2.16)$$

Since the mean $\mathbf{c}^{(e)}$ and members $\mathbf{c}^{(r)}$ of the ensemble are known, the control vector becomes the optimal weights \mathbf{a} .

The optimization problem can be rewritten in terms of the new control vector. Using (2.14), the energy constraint (2.13c) becomes

$$\frac{1}{2} \mathbf{c}^{(e)tr} \mathbf{A}_1 \mathbf{c}^{(e)} + \mathbf{c}^{(e)tr} \mathbf{A}_1 \mathbf{E}' \mathbf{a} + \frac{1}{2} \mathbf{a}^{tr} \mathbf{E}'^{tr} \mathbf{A}_1 \mathbf{E}' \mathbf{a} = E_0. \quad (2.17)$$

Similarly, the inequality constraint (2.13d) is recast as

$$\mathbf{A}_1 \mathbf{c}^{(e)} + \mathbf{A}_1 \mathbf{E}' \mathbf{a} \geq 0. \quad (2.18)$$

The details of constructing the ensemble members $\mathbf{c}^{(r)}$ are provided in appendix A. The procedure ensures the following three properties: (i) the mean of the ensemble members is $\mathbf{c}^{(e)}$; (ii) each member and the mean satisfy the energy constraint, $\frac{1}{2} \mathbf{c}^{(r)tr} \mathbf{A}_1 \mathbf{c}^{(r)} = E_0$ for $r = 1, \dots, N_{en}$ and $\frac{1}{2} \mathbf{c}^{(e)tr} \mathbf{A}_1 \mathbf{c}^{(e)} = E_0$; and (iii) the deviations of the members from the mean are controlled using a covariance matrix. It is important to note that the arithmetic mean cannot satisfy (i) and (ii) simultaneously.

Instead, the modal amplitudes and phases of the ensemble members and the mean control vector are related by, respectively,

$$\mathbf{A}_1 \mathbf{c}^{(e)} \circ \mathbf{A}_1 \mathbf{c}^{(e)} = \frac{1}{N_{en}} \sum_{r=1}^{N_{en}} \mathbf{A}_1 \mathbf{c}^{(r)} \circ \mathbf{A}_1 \mathbf{c}^{(r)}, \tag{2.19a}$$

and

$$\mathbf{A}_2 \mathbf{c}^{(e)} = \frac{1}{N_{en}} \sum_{r=1}^{N_{en}} \mathbf{A}_2 \mathbf{c}^{(r)}, \tag{2.19b}$$

where \circ is element-wise product of two vectors and

$$\mathbf{A}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}_{2M \times 2M}. \tag{2.20}$$

The optimization procedure requires computation of the gradient of the cost function with respect to the control variable, and therefore \mathcal{J} must be expressed as a differentiable function of \mathbf{a} . First, the governing Navier–Stokes equations (2.13b) are written in terms of the new control vector

$$\mathbf{q} = \mathcal{N}(\mathbf{c}^{(e)} + \mathbf{E}'\mathbf{a}). \tag{2.21}$$

The equation is expanded around the mean vector using Taylor series

$$\mathbf{q} = \mathbf{q}^{(e)} + \left. \frac{\partial \mathcal{N}}{\partial \mathbf{c}} \right|_{\mathbf{c}^{(e)}} \mathbf{E}'\mathbf{a} + \frac{1}{2!} \left. \frac{\partial^2 \mathcal{N}}{\partial \mathbf{c}^2} \right|_{\mathbf{c}^{(e)}} (\mathbf{E}'\mathbf{a} \circ \mathbf{E}'\mathbf{a}) + \dots, \tag{2.22}$$

where $\mathbf{q}^{(e)} = \mathcal{N}(\mathbf{c}^{(e)})$. By assuming the members of the ensemble are close to the mean, or $\mathbf{E}'\mathbf{a}$ is small, the high-order terms in (2.22) can be ignored as follows:

$$\mathbf{q} \approx \mathbf{q}^{(e)} + \left. \frac{\partial \mathcal{N}}{\partial \mathbf{c}} \right|_{\mathbf{c}^{(e)}} \mathbf{E}'\mathbf{a}. \tag{2.23}$$

Substituting (2.23) into (2.13a) and using Taylor series expansion, the cost function is approximated as

$$\mathcal{J}(\mathbf{a}) \approx \frac{1}{2} \left\| \mathcal{G} \left(\mathbf{q}^{(e)} + \left. \frac{\partial \mathcal{N}}{\partial \mathbf{c}} \right|_{\mathbf{c}^{(e)}} \mathbf{E}'\mathbf{a} \right) \right\|^2 \approx \frac{1}{2} \left\| \mathcal{G}(\mathbf{q}^{(e)}) + \left. \frac{\partial \mathcal{G}}{\partial \mathbf{q}} \right|_{\mathbf{q}^{(e)}} \left. \frac{\partial \mathcal{N}}{\partial \mathbf{c}} \right|_{\mathbf{c}^{(e)}} \mathbf{E}'\mathbf{a} \right\|^2. \tag{2.24}$$

As long as the linear approximations in (2.23) and (2.24) are valid, constraint (2.13b) is satisfied and the observation matrix $\mathbf{H} \equiv (\partial \mathcal{G} / \partial \mathbf{q})|_{\mathbf{q}^{(e)}} (\partial \mathcal{N} / \partial \mathbf{c})|_{\mathbf{c}^{(e)}} \mathbf{E}'$ can be evaluated as

$$\mathbf{H} \approx [\mathcal{G}(\mathbf{q}^{(1)}) - \mathcal{G}(\mathbf{q}^{(e)}) \quad \dots \quad \mathcal{G}(\mathbf{q}^{(N_{en})}) - \mathcal{G}(\mathbf{q}^{(e)})]_{N_G \times N_{en}}, \tag{2.25}$$

where N_G is the size of observation vector. Note that the validity of (2.23) to (2.25) is predicated on $\mathbf{E}'\mathbf{a}$ being small – a condition that is ensured during the generation of the ensemble members (appendix A). Assembling matrix \mathbf{H} is the most expensive

part of the algorithm since it requires $N_{en} + 1$ solutions of the governing equations. While in the current study direct numerical simulations are used to evaluate \mathbf{H} , the derivation of (2.25) did not invoke any assumptions regarding the choice of the numerical approach to computing $\mathcal{G}(\mathbf{q}^{(r)})$. As a result, other approaches such as large-eddy simulations and the nonlinear parabolized stability equations can potentially be adopted.

In summary, the constrained optimization (2.13) can be recast as seeking the weight vector \mathbf{a} that optimizes the cost function

$$\mathcal{J}(\mathbf{a}) \approx \frac{1}{2} \|\mathcal{G}(\mathbf{q}^{(e)}) + \mathbf{H}\mathbf{a}\|^2, \tag{2.26a}$$

while satisfying the constraints

$$\frac{1}{2} \mathbf{c}^{(e)tr} \mathbf{A}_1 \mathbf{c}^{(e)} + \mathbf{c}^{(e)tr} \mathbf{A}_1 \mathbf{E}' \mathbf{a} + \frac{1}{2} \mathbf{a}^{tr} \mathbf{E}^{tr} \mathbf{A}_1 \mathbf{E}' \mathbf{a} - E_0 = 0, \tag{2.26b}$$

and

$$-\mathbf{A}_2 \mathbf{E}' \mathbf{a} \leq \mathbf{A}_2 \mathbf{c}^{(e)}. \tag{2.26c}$$

Interior-point method (Nocedal & Wright 2006; Wächter & Biegler 2006) is used to solve (2.26) by generating iterates that satisfy the inequality bounds, strictly. The full procedure to solve the optimization problem, and hence identify the nonlinearly most dangerous disturbance, is shown in algorithm 1.

Algorithm 1: Finding the nonlinearly most dangerous perturbation using an ensemble-variational optimization technique.

- *Iteration* = 0.
- Establish a first guess for the control vector, $\mathbf{c}^{(e)}$, that satisfies the constraints.
- Simulate the evolution of the mean control vector, $\mathbf{q}^{(e)} = \mathcal{N}(\mathbf{c}^{(e)})$.
- Compute the mean observation vector $\mathcal{G}(\mathbf{q}^{(e)})$.

while *The convergence condition is not satisfied* **do**

- construct an ensemble of $\mathbf{c}^{(r)}$ that satisfy the constraints (see appendix A);
- compute the deviation matrix \mathbf{E}' (equation 2.15);
- simulate the evolution of each ensemble member, $\mathbf{q}^{(r)} = \mathcal{N}(\mathbf{c}^{(r)})$;
- compute the observation vectors of the members of the ensemble $\mathcal{G}(\mathbf{q}^{(r)})$;
- compute the observation matrix \mathbf{H} (equation 2.25);
- solve the constrained optimization (2.26) using interior-point method;
- update the mean control vector as $\mathbf{c}^{(e)} \rightarrow \mathbf{c}^{(e)} + \mathbf{E}' \mathbf{a}$;
- simulate the evolution of the mean control vector, $\mathbf{q}^{(e)} = \mathcal{N}(\mathbf{c}^{(e)})$;
- compute the mean observation vector $\mathcal{G}(\mathbf{q}^{(e)})$;
- check for convergence;
- *Iteration* = *Iteration* + 1.

end

Case	Ma_∞	$\sqrt{Re_{x_0}}$	L_x	L_y	L_z	N_x	N_y	N_z	$E_0 \times 10^5$	F_l	F_u	$k_{z,l}$	$k_{z,u}$
E1	4.5	1800	2983.5	200	150	1989	130	108	2	10	250	0	15
E2	4.5	1800	2983.5	200	150	1989	130	108	100	10	250	0	15

TABLE 1. Physical and computational parameters for the cases of the present study.

3. Computational aspects

During the search for the nonlinearly most dangerous inflow disturbance, the optimization algorithm involves the solution of the Navier–Stokes equations for every ensemble member and observation of the wall friction. The simulations are performed using code Hybrid (Johnsen *et al.* 2010), which solves the compressible Navier–Stokes equations (2.1) on a structured grid and is designed for accurate simulations of high-speed transitional and fully turbulent flows. It is nominally free of numerical dissipation, which is important for accurate prediction of the evolution of instability waves and transition onset. The time advancement is fourth-order accurate using the Runge–Kutta scheme, and the spatial discretization in the absence of shocks is sixth-order central with a split form which improves nonlinear stability, and a fifth-order weighted essentially non-oscillatory scheme with Roe flux splitting is used near shocks. Code Hybrid was described and adopted in a number of previous studies (see e.g. Larsson & Lele 2009; Johnsen *et al.* 2010; Kawai & Larsson 2012).

A schematic of the simulated domain is shown in figure 1. Two cases are considered to highlight the effect of the energy level on the nonlinearly most dangerous inflow perturbations. Table 1 summarizes the computational and physical parameters of each case, which are identical except for the inflow disturbance energy. The free-stream values are adopted as reference scales for normalization of velocity, temperature, density, viscosity, specific heat and conductivity. The reference length is the Blasius length scale, $(\tilde{\mu}_\infty \tilde{x}_0 / \tilde{\rho}_\infty \tilde{U}_\infty)^{1/2}$, at the inlet location, where the $\tilde{\cdot}$ represents dimensional quantities and \tilde{x}_0 is the start location of the simulation relative to the virtual boundary-layer origin.

The boundary conditions are periodic in the spanwise (z) direction, while characteristic conditions are adopted on the remaining four boundaries in the streamwise (x) and wall-normal (y) directions. The marked sponge regions in figure 1 along the outflow boundaries minimize reflections of outgoing disturbances. Within these layers, a relaxation term, $\sigma d^2(\mathbf{q}_{ref} - \mathbf{q})$, is introduced in the governing equations to force the solution towards the spanwise-averaged \mathbf{q}_{ref} at each time step. In this expression, $d \in [0, 1]$ is the normalized distance within the sponge region. In the streamwise direction the width of the sponge is 250 units and $\sigma = 0.7$, while in the wall-normal direction the sponge height is 30 units and $\sigma = 0.85$. These values were verified to effectively damp outgoing vortical, acoustic and entropic waves, without causing reflections.

The free-stream Mach number in all simulations is $Ma_\infty = 4.5$, the operating gas is air ($Pr = 0.72$ and $\gamma = 1.4$) and the free-stream temperature used in Sutherland’s law is 65.15 K. The Reynolds number based on inlet is Re_{x_0} and the streamwise position of the inflow plane is $x_0 = \sqrt{Re_{x_0}} = 1800$, which was selected based on the transition Reynolds numbers in high-altitude flight tests being $\sqrt{Re_{x,tr}} > 2000$ for $Ma_\infty > 4$ (Harvey 1978; Schneider 1999). Starting the simulations from smaller Reynolds numbers requires a much longer domain size, which poses prohibitively high computational requirements. The lengths of the computational domain in the x ,

y and z directions are L_x , L_y and L_z , which are discretized with N_x , N_y and N_z grid points, respectively (see table 1). The computational grid is uniform in the x and z directions, and a hyperbolic tangent stretching is used in the y direction. Our choice of computational parameters was informed by preliminary simulations with nonlinear instability amplitudes that cause transition at different streamwise locations. We have also validated our algorithm by simulating the evolution of very small-amplitude instability waves, which are very sensitive to numerical accuracy, and comparing to the parabolized stability equations.

The two cases, E1 and E2, are distinguished by the total disturbance energy at the inflow plane, E_0 , which is prescribed as a constraint in the optimization problem (constraint (2.26b)). Comparison of E1 and E2 highlights the effect of the energy constraint on the spectral makeup of the nonlinearly most dangerous inflow perturbation, and the associated transition mechanism. The amount of energy for case E1 is chosen to be of the same order as the turbulence kinetic energy (TKE) in stratospheric layers, measured during a recent experimental campaign by Haack, Gerding & Lübken (2014) with the balloon-borne instrument Leibniz Institute Turbulence Observations in the Stratosphere (LITOS).

The normalized frequencies of the disturbance (2.8) are assumed to range from F_l to F_u with an increment $\Delta F = F_l = 10$, where

$$F = \frac{\omega_n}{\sqrt{Re_{x_0}}} \times 10^6 = nF_l \quad \text{for } n = 1, 2, \dots, \frac{F_u}{F_l}. \quad (3.1)$$

Similarly, $k_{z,l}$ and $k_{z,u}$ in table 1 are the lower and upper bounds of the integer spanwise wavenumbers of the inflow disturbance, where

$$k_z = \frac{\beta_m L_z}{2\pi} = m \quad \text{for } m = k_{z,l}, \dots, k_{z,u}. \quad (3.2)$$

For the three-dimensional modes at the inlet, $k_z > 0$, the modal energy is divided equally between the positive and negative spanwise wavenumbers, $\pm\beta_m$, with the same phase $\theta_{n,m}$ in (2.8). Therefore, hereafter only $k_z > 0$ will be reported since they are representative of the pairs of oblique modes. The total number of frequencies (3.1) and spanwise wavenumbers (3.2) at the inlet is $M = 400$. And since both the amplitude and phase of each instability wave must be optimized, the size of the control vector is equal to $2M$. The number of unknowns is directly proportional to the computational cost that is required to solve the optimization problem, which places a practical constraint on M . Based on the present size of the control vector, and guided by some preliminary tests, the size of the ensemble that is adopted in the EnVar algorithm 1 was $N_{en} = 20$.

Figure 2 shows the neutral curves from linear spatial stability analysis of a ZPG boundary layer at $Ma_\infty = 4.5$. The lower unstable regions in the figure are commonly referred to as (Mack's) first-mode instability regions, whereas the upper unstable regions are called (Mack's) second-mode disturbances (Mack 1984). The ranges of F and k_z in table 1 are appropriately selected such that they include all the relevant frequencies and spanwise wavenumbers at the inlet location. For every pair of frequency and spanwise wavenumber, the most unstable discrete mode in the Orr–Sommerfeld and Squire spectra is included; these instabilities correspond to slow modes (Fedorov 2011). For the 400 inlet instability waves, the spatial resolutions reported in table 1 correspond to a minimum of nine and a maximum of 225 grid points within one streamwise wavelength, and a minimum of seven and a maximum of 108 grid points in one spanwise wavelength. In the wall-normal direction, 40 grid points are within the boundary layer at the inlet plane.

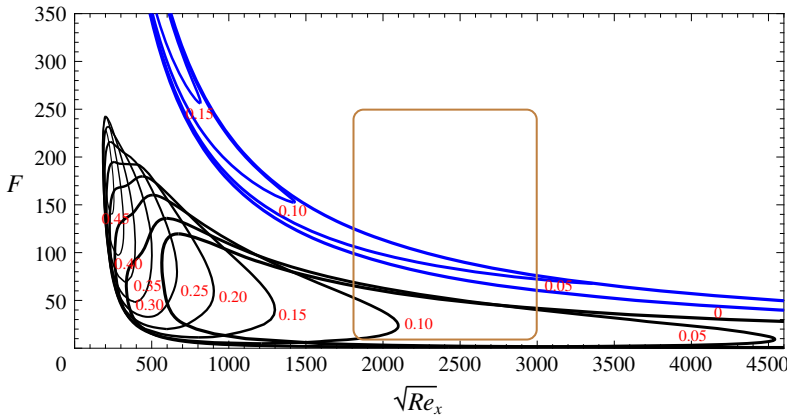


FIGURE 2. (Colour online) Neutral curves of a ZPG boundary layer at $Ma_\infty = 4.5$. The values marked on the curves correspond to the normalized spanwise wavenumber, $b = (\beta/\sqrt{Re_x}) \times 10^3$. The marked rectangular region corresponds to the streamwise extent of the simulation domain and the range of inflow frequencies.

4. Results and discussion

4.1. Nonlinearly most dangerous inflow spectra

The optimization algorithm 1 seeks the nonlinearly most dangerous disturbance by maximizing the cost function $\mathcal{J} = \frac{1}{2} \|\mathcal{G}(\mathbf{q})\|^2 = \frac{1}{2L_x} \int_{x_0}^{x_f} C_f^2 dx$, where

$$C_f = \frac{\tau_{wall}}{\frac{1}{2} \rho_\infty U_\infty^2}. \tag{4.1}$$

The initial guess of the disturbance spectra was equipartition of the energy among all instability waves and randomly assigned phases. In addition to the mean, 20 ensemble members were used in each iteration. The stopping criterion of the optimization was $(\mathcal{J}_i - \mathcal{J}_{i-1})/\mathcal{J}_i < 10^{-3}$, where i is the iteration number. It should be noted that the optimization was not repeated for any other initial guesses. As such, similar to other gradient-based methods, the reported results are only guaranteed to be local optima for a given choice of the initial guess.

For each iteration, C_f , $\mathcal{J}/\mathcal{J}_0$ and $\sqrt{Re_{x,tr}}$ associated with the mean control vectors $\mathbf{c}^{(e)}$ are plotted in figure 3; the transition Reynolds number $\sqrt{Re_{x,tr}}$ is defined as the location of minimum C_f . For case E1, the flow remains laminar throughout the computational domain for the initial guess and first iteration. As a result, the C_f curves of these two inflow perturbations nearly coincide and, while $(\mathcal{J}_1 - \mathcal{J}_0)/\mathcal{J}_1 \approx 9.99 \times 10^{-4} < 10^{-3}$, further iterations are performed. The convergence rate of the optimization algorithm is highly dependent on the observation matrix, defined in (2.25). Figure 3(a,c) shows that the convergence rate of the algorithm for case E1 is slow during the first two iterations. This behaviour is the outcome of the observation matrix, formed by the skin-friction profiles, being nearly singular for these two iterations. Transition to turbulence generally advances upstream (i.e. smaller $\sqrt{Re_{x,tr}}$) with successive iterations, with convergence observed for both E1 and E2. For E1, transition location is nearly unchanged from the tenth to the eleventh iteration, and for case E2 a similar behaviour is observed from the seventh to the eighth iteration. Quantitatively, the change in transition Reynolds number during the last two

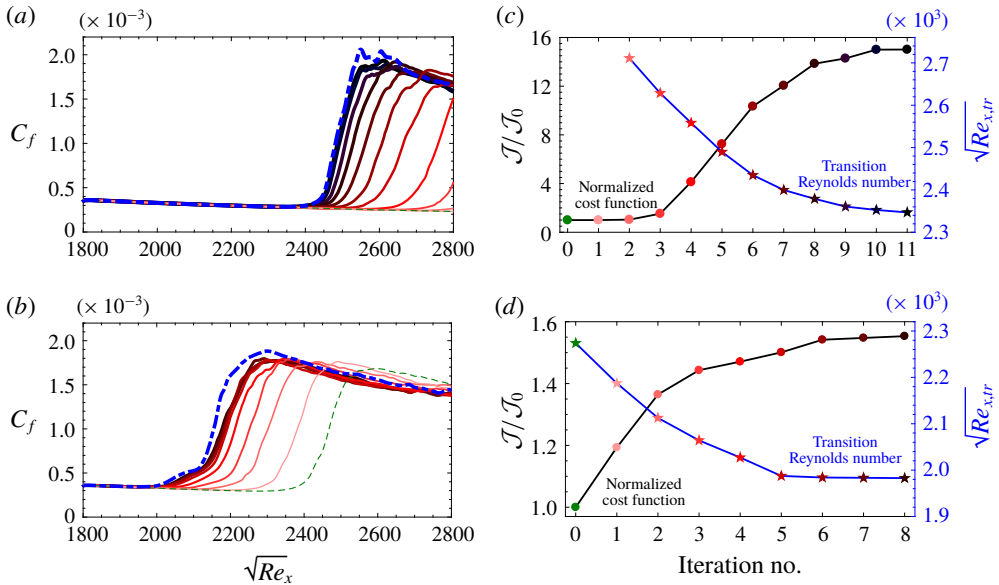


FIGURE 3. (Colour online) (a,b) Skin friction and (c,d) normalized cost function and transition Reynolds number for cases (a,c) E1 and (b,d) E2. In the C_f plots, the dashed green profiles correspond to the initial guess, and thin to thick (also light to dark coloured) solid lines represent the result for the mean control vector after consecutive iterations. The dashed-dotted blue curves correspond to the nonlinearly most dangerous inflow spectra for each case (see table 2).

iterations is approximately 0.2% and 0.03% for cases E1 and E2, respectively. Each iteration of the optimization procedure required approximately 50 000 CPU hours, and therefore the total computational cost was half a million CPU hours per inflow energy level.

The choice of the cost function in algorithm 1 is not unique. Depending on the application of interest, measures based on the total energy of the perturbations within the domain or the wall temperature could have been adopted. Every new cost function, however, requires repeating the entire optimization procedure, which is not performed here. Instead, for each iterate of the optimization where the cost function was proportional to skin friction, the average integrated energy and the average wall temperature were recorded and are plotted in figures 4 and 5. Note that in figure 4, the energy \mathcal{E} is computed from (2.7) with fluctuations evaluated relative to a laminar solution over the entire plate. As a result, a base-state distortion term with frequency and spanwise wavenumber $(0, 0)$ contributes to the perturbation energy.

The streamwise integrals of the energy, $\mathcal{M}_{\mathcal{E}} = (1/L_x) \int_{x_0}^{x_f} \mathcal{E} dx$, and of the wall temperature, $\mathcal{M}_{\bar{T}_{wall}} = (1/L_x) \int_{x_0}^{x_f} \bar{T}_{wall} dx$, are also plotted in figures 4 and 5, normalized by their values from the initial guess. Note, however, that these norms were not used in an optimization procedure. The ratio $\mathcal{M}_{\mathcal{E}}/\mathcal{M}_{\mathcal{E},0}$ increases with every iteration for both inflow perturbation levels.

Comparison of figures 4 and 3 highlights the rationale for choosing C_f for the definition of the cost function. The increase in the perturbation energy takes place throughout the domain, initially through the amplification of primary instabilities, subsequent secondary instabilities and ultimately breakdown to turbulence if it

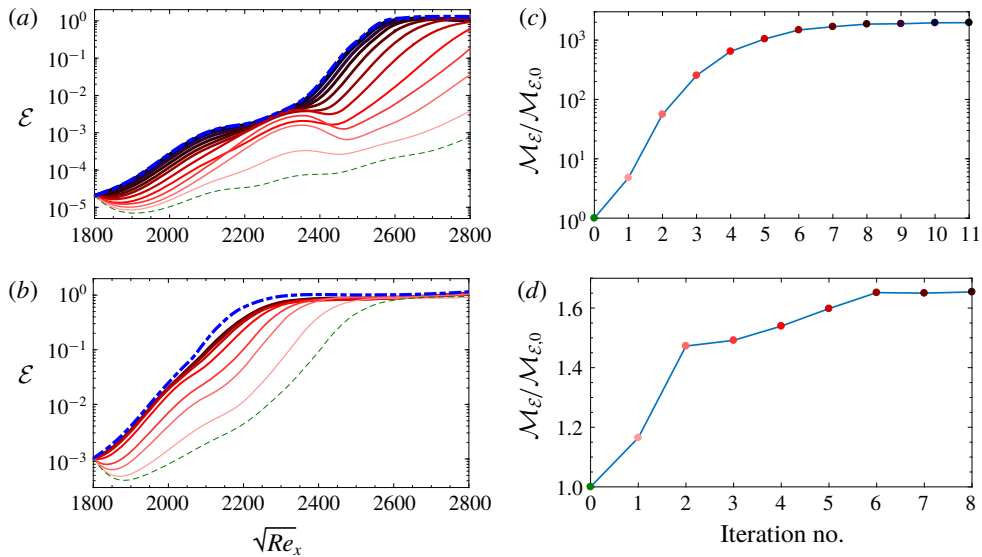


FIGURE 4. (Colour online) (a,b) Curves of \mathcal{E} versus downstream Reynolds number and (c,d) the ratio $\mathcal{M}_{\mathcal{E}}/\mathcal{M}_{\mathcal{E},0}$, for cases (a,c) E1 and (b,d) E2. In the \mathcal{E} plots, the dashed green profiles correspond to the initial guess (iteration no. 0 in algorithm 1), thin to thick (also light to dark coloured) solid lines represent the optimal solutions after consecutive iterations and the dashed-dotted blue profiles correspond to the nonlinearly most dangerous inflow spectra for each case (the spectra in table 2).

takes place within the domain. For case E1, the initial guess yields an increasing perturbation energy even though the flow is laminar throughout the computational domain; the second iterate too yields a laminar solution, even though its energy integral is higher. Therefore, using \mathcal{E} as the observation vector may not in theory guarantee the fastest breakdown to turbulence, although in practice it is effective (Pringle & Kerswell 2010). In contrast, the skin-friction curve reproduces the laminar behaviour throughout the pre-transitional region and only increases towards the turbulent correlation when intermittency is finite, where intermittency is the fraction of time that the flow is turbulent. As a result, a cost function based on C_f only shows appreciable increase in case E1 when the flow undergoes transition to turbulence. In addition, increasing the cost function based on C_f is directly tied to generation of turbulence early upstream, unlike a cost function based on the perturbation energy that includes contributions from the amplification of instability waves in the laminar region of the flow.

The mean wall temperature is plotted versus downstream Reynolds number in figure 5, along with $\mathcal{M}_{\bar{T}_{wall}}/\mathcal{M}_{\bar{T}_{wall},0}$ at each iteration and for both levels of the inflow perturbation energy. The figure shows that, while a cost function based on skin friction promotes early breakdown to turbulence, it does not guarantee that $\mathcal{M}_{\bar{T}_{wall}}/\mathcal{M}_{\bar{T}_{wall},0}$ is highest. In addition, based on the temperature profiles, while a cost function based on skin friction promotes early breakdown to turbulence, it does not guarantee that the peak temperature in the transition zone is largest. These results demonstrate that optimizing for minimum viscous losses or thermal loading in hypersonic applications would not necessarily lead to the same outcome; an appropriate cost function must be adopted for each choice and the results must be compared.

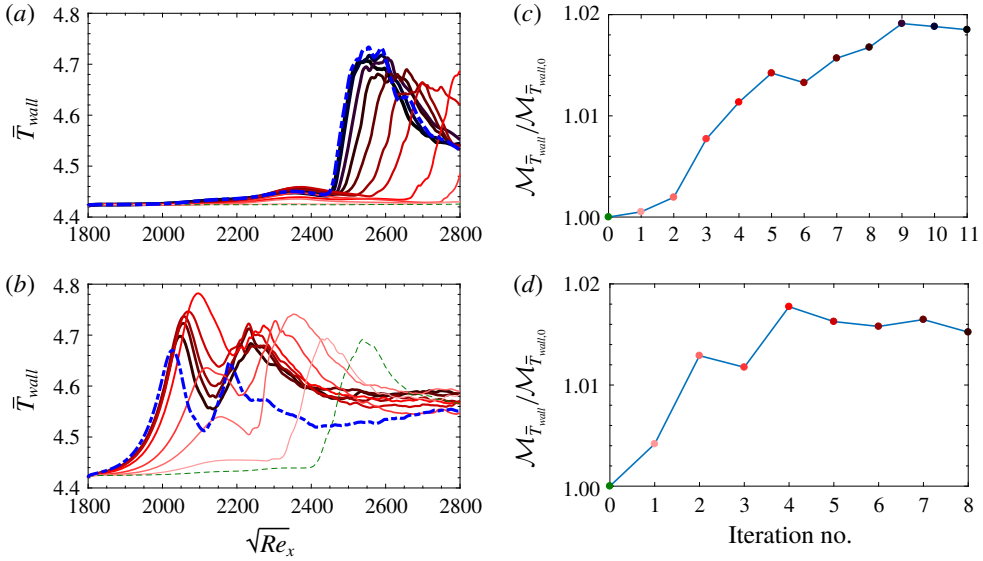


FIGURE 5. (Colour online) (a,b) Curves of \bar{T}_{wall} versus downstream Reynolds number and (c,d) the ratio $\mathcal{M}_{\bar{T}_{wall}}/\mathcal{M}_{\bar{T}_{wall},0}$, for cases (a,c) E1 and (b,d) E2. In the \bar{T}_{wall} plots, the dashed green profiles correspond to the initial guess (iteration no. 0 in algorithm 1), thin to thick (also light to dark coloured) solid lines represent the optimal solutions after consecutive iterations and the dashed-dotted blue profiles correspond to the nonlinearly most dangerous inflow spectra for each case (the spectra in table 2).

The focus is hereafter placed on the results from the optimization algorithm, where the cost function is defined using the wall friction. Figure 6 shows the energy spectra of the optimal inflow disturbances

$$\mathcal{E}_{(F,k_z)} = \frac{1}{2} \int_0^{L_y} \hat{\psi}_{(F,k_z)}^* \text{diag}(\xi) \hat{\psi}_{(F,k_z)} dy, \tag{4.2}$$

where $\hat{\psi}_{(F,k_z)}$ are the two-dimensional Fourier coefficients of the perturbation vector $\psi' \equiv [\rho' \ u' \ v' \ w' \ T']^T$, the subscripts correspond to the frequency $F \equiv 10^6 \omega / \sqrt{Re_{x_0}}$ and integer spanwise wavenumber $k_z \equiv \beta / (2\pi / L_z)$ and $\hat{\psi}^*$ is the complex-conjugate transpose. In (4.2), the term $\xi = [(R/\gamma - 1)(\bar{T}/\bar{\rho}) \ \bar{\rho} \ \bar{\rho} \ \bar{\rho} \ (R/\gamma - 1)(\bar{\rho}/\bar{T})]^T$ ensures that the kinetic and internal energies are appropriately weighted; note that overline in ξ denotes the laminar state.

For case E1 (figure 6a), most of the available energy is allocated to modes $\langle 40, 3 \rangle$, $\langle 100, 0 \rangle$ and $\langle 110, 0 \rangle$. The energy ratio among these modes is (1.66 : 1.18 : 1). The downstream development of the spectra for all 400 instability waves that comprise the optimal perturbation was evaluated from the direct numerical simulation data (figure 1a in the supplementary material is available at <https://doi.org/10.1017/jfm.2019.527>), and confirmed that the three identified modes are indeed the most important instability waves at the inlet of the computational domain. Collectively they are responsible for the earliest transition to turbulence for case E1. In order to verify this assertion, a few test cases were performed with variations to the inflow spectrum, which were motivated by figure 6(a) and also by analysis of the downstream development of the

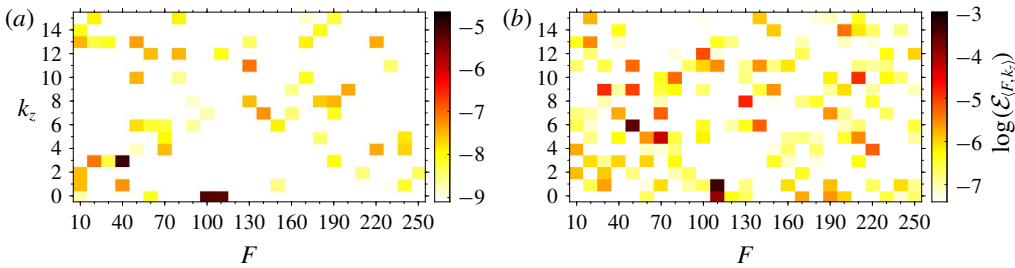


FIGURE 6. (Colour online) The energy distribution amongst the instability modes at the inflow corresponding to the optimal solutions after the final iteration for cases (a) E1 and (b) E2.

Case	$E_0 \times 10^5$	Inlet mode $\langle F, k_z \rangle$	$\mathcal{E}_{\langle F, k_z \rangle} / E_0$	$\theta_{\langle F, k_z \rangle} / \pi$
E1N	2	$\langle 40, 3 \rangle$	0.43	0.09
		$\langle 100, 0 \rangle$	0.31	0.00
		$\langle 110, 0 \rangle$	0.26	0.12
E2N	100	$\langle 50, 6 \rangle$	0.32	0.00
		$\langle 110, 1 \rangle$	0.68	0.02
E1L1	2	$\langle 100, 0 \rangle$ broadband	0.99 0.01	Randomly assigned
E1L2	2	$\langle 90, 0 \rangle$ broadband	0.99 0.01	Randomly assigned
E1L3	2	$\langle 20, 2 \rangle$ broadband	0.99 0.01	Randomly assigned
E2L	100	$\langle 110, 0 \rangle$ broadband	0.99 0.01	Randomly assigned

TABLE 2. Cases E1N and E2N are the nonlinearly most dangerous inflow spectra for the low and high energy levels. Cases E1L1, E1L2, E1L3 and E2L are the linearly most unstable inflow spectra.

instability waves. All tests confirmed that a disturbance comprised of mode $\langle 40, 3 \rangle$ with 43% of total energy, mode $\langle 100, 0 \rangle$ with 31% of total energy and mode $\langle 110, 0 \rangle$ with 26% of total energy is the most potent inflow spectra for case E1. This reduced form of the disturbance is reported in table 2, and will be referred to as E1N, and is examined in detail in § 4.2. The associated skin-friction coefficient is reported in figure 3(a) (dashed blue line).

The spectrum of case E2 after convergence of the optimization algorithm is reported in figure 6(b): most of the available energy is in modes $\langle 50, 6 \rangle$ and $\langle 110, 1 \rangle$. The ratio of their energy content is (1 : 2.17). Similar to E1, the downstream development of the spectra for E2 was computed from direct numerical simulation data, for all the 400 wavenumber pairs at the inflow and which can potentially be excited downstream (figure 1b in the supplementary material). Careful assessment of the

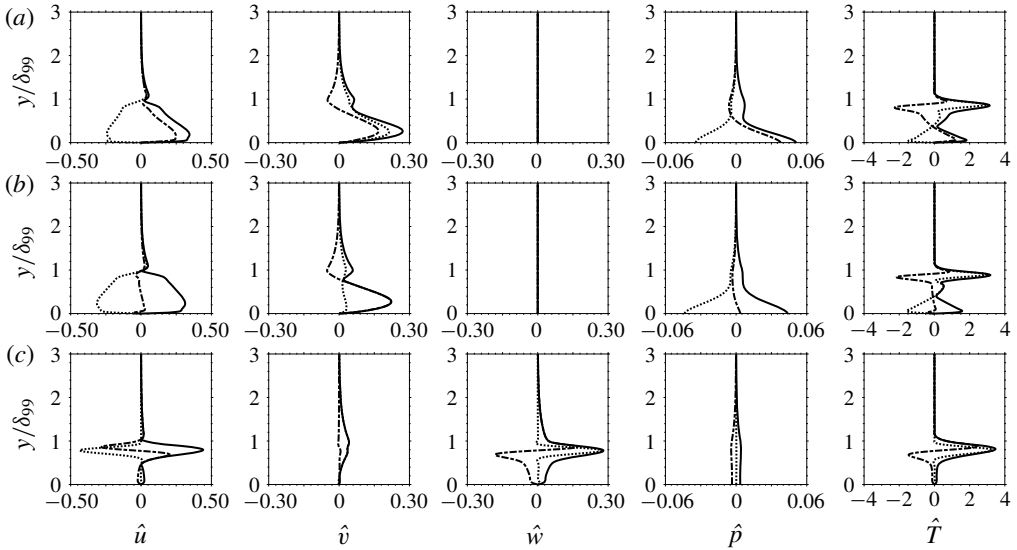


FIGURE 7. The mode shapes of the instability waves of case E1N at the inlet, $\sqrt{Re_x} = 1800$. (a) Mode $\langle 110, 0 \rangle$, (b) mode $\langle 100, 0 \rangle$ and (c) mode $\langle 40, 3 \rangle$. The solid, dashed-dotted and dotted lines are the magnitude, real part and imaginary part of the mode shapes, respectively.

spectra indicated that modes $\langle 50, 6 \rangle$ and $\langle 110, 1 \rangle$ are the most important elements of the inflow spectra, and are responsible for the earliest transition to turbulence for case E2. This conclusion was verified using several additional simulations where the inflow spectrum was modified and the transition location was compared to the optimal. In one of the tests, the spanwise size of the domain was doubled and the energy of mode $\langle 110, 1 \rangle$ was reassigned to mode $\langle 110, \frac{1}{2} \rangle$ in order to ensure that the width of the domain did not influence the outcome of the optimization procedure. All tests confirmed that an inflow disturbance spectrum composed of modes $\langle 50, 6 \rangle$ and $\langle 110, 1 \rangle$ with 32% and 68% of total energy, respectively, is the most potent inflow condition for case E2. The skin-friction coefficient associated with this disturbance is reported in figure 3(b) as the dashed blue line. It is designated E2N in table 2, and a detailed description of its transition mechanism is discussed in § 4.2.

The mode shapes that play an important role in cases E1N and E2N are plotted in figures 7 and 8, respectively. As the \hat{p} profiles in figures 7 and 8 indicate, modes $\langle 110, 0 \rangle$, $\langle 100, 0 \rangle$ and $\langle 110, 1 \rangle$ with one zero crossing of $\text{real}(\hat{p})$ in the wall-normal direction are the so-called second-mode instabilities (Mack 1984). These are acoustic waves that reflect back and forth between the wall and the sonic line of the relative flow. Modes $\langle 40, 3 \rangle$ and $\langle 50, 6 \rangle$ are first-mode instabilities which, for $Ma_\infty < 5$, are vorticity waves that can cause strong inflectional instability.

Depending on the wavelength of the modes and the local boundary-layer thickness, the growth rate of the acoustic waves can be much larger than the typical vorticity modes. It is also evident in figures 7 and 8 that, while the second-mode instabilities are large across the boundary layer, the first-mode instabilities have appreciable magnitude only close to its edge. These observations have important implications for the mechanism of transition to turbulence in cases E1N and E2N. For instance, a commonly observed feature in hypersonic and supersonic boundary-layer transition is

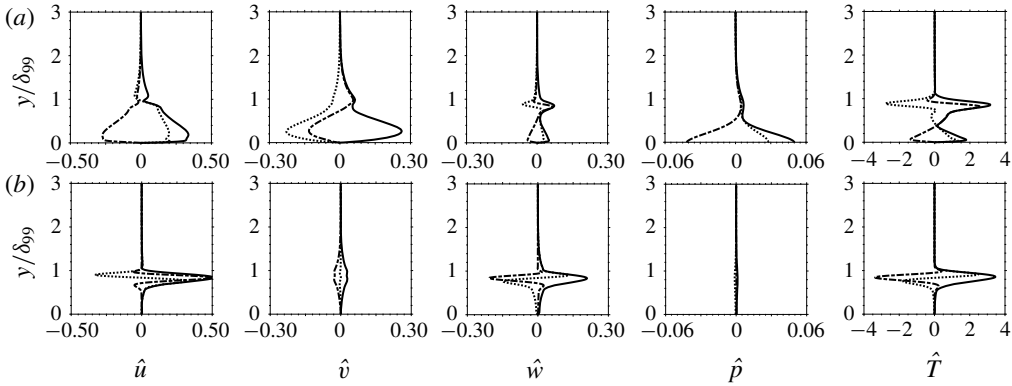


FIGURE 8. The mode shapes of the instability waves of case E2N at the inlet, $\sqrt{Re_x} = 1800$. (a) Mode $\langle 110, 1 \rangle$ and (b) mode $\langle 50, 6 \rangle$. The solid, dashed-dotted and dotted lines are the magnitude, real part and imaginary part of the mode shapes, respectively.

the generation of elongated streaks from the nonlinear interaction of oblique instability waves (Thumm, Wolz & Fasel 1990; Jiang *et al.* 2006; Mayer, Von Terzi & Fasel 2011; Franko & Lele 2013; Sivasubramanian & Fasel 2016). Since streaks are the boundary-layer response to vortical forcing by three-dimensional modes, they are anticipated in case E2N in response to mode $\langle 110, 1 \rangle$. Also note that mode $\langle 110, 1 \rangle$ has strong wall-normal perturbation which is essential for the lift-up process that leads to the streak response (Zaki & Durbin 2005, 2006).

4.2. Transition mechanisms

In this section we discuss the two transition mechanisms due to the nonlinearly most dangerous inflow spectra for cases E1N and E2N (table 2), respectively. In the former case, the inflow is comprised of a pair of two-dimensional second-mode disturbances and an oblique first-mode instability wave. We will demonstrate that the resulting breakdown to turbulence cannot be categorized as fundamental and/or oblique, and is hence a new mechanism. Focus will therefore be placed on this case, followed by a relatively brief discussion of the higher energy level where transition takes place via a second-mode oblique breakdown.

The downstream development of $\mathcal{E}_{(F,k_z)}$ of modes that play a principal role in transition was evaluated using (4.2) and is plotted in figures 9 and 11. For clarity, the curves become thin and faintly coloured beyond the location where their interactions are discussed. Figures 9 and 11 also include instantaneous iso-surfaces of streamwise velocity u coloured by its fluctuation u' . Figures 9(e) and 11(e) feature the downstream development of integrated terms of the TKE equation

$$\underbrace{\frac{1}{2} \frac{\partial \overline{\rho u'_i u'_i}}{\partial t} + \frac{1}{2} \frac{\partial \overline{u'_j \rho u'_i u'_i}}{\partial x_j}}_{\bar{D}k/\bar{D}t} = \underbrace{-\overline{\rho u'_i u'_j}}_P \frac{\partial \overline{u'_i}}{\partial x_j} - \underbrace{\overline{\tau'_{ij}}}_{\epsilon} \frac{\partial \overline{u'_i}}{\partial x_j} - \underbrace{\frac{1}{2} \frac{\partial \overline{\rho u'_i u'_i u'_j}}{\partial x_j}}_{\mathcal{T}_1} - \underbrace{\frac{\partial \overline{p' u'_i \delta_{ij}}}{\partial x_j}}_{\mathcal{T}_2} + \underbrace{\frac{\partial \overline{\tau'_{ij} u'_i}}{\partial x_j}}_{\mathcal{T}_3} + \underbrace{p' \frac{\partial \overline{u'_i}}{\partial x_i}}_{\Pi} + \underbrace{\overline{u'_i} \frac{\partial \overline{\tau'_{ij}}}{\partial x_j}}_{\Sigma_1} - \underbrace{\overline{u'_i} \frac{\partial \overline{p}}{\partial x_i}}_{\Sigma_2}. \quad (4.3)$$

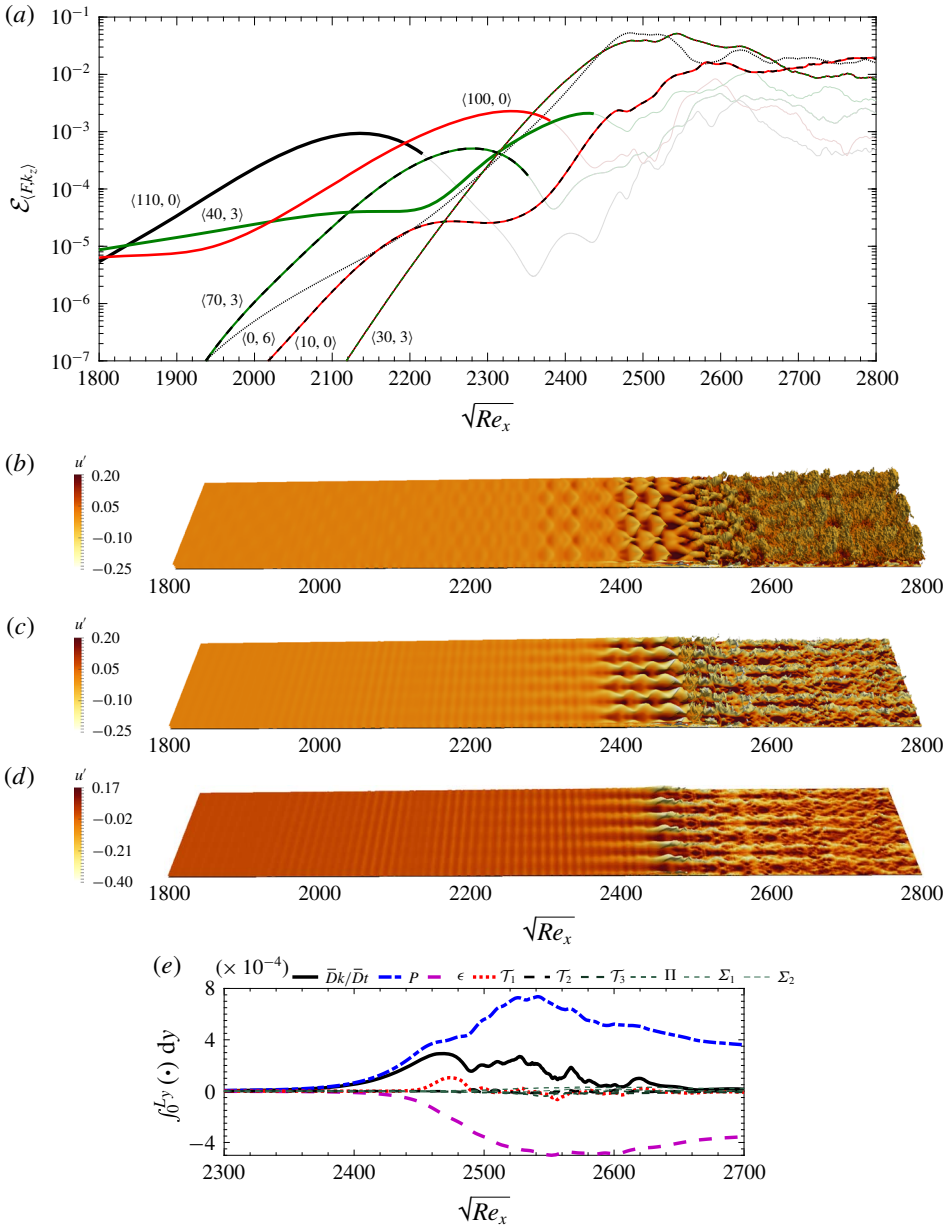


FIGURE 9. (Colour online) Results corresponding to case E1N. (a) Fluctuation energy content for selected instability modes, $\mathcal{E}_{(F,k_z)}$, versus the streamwise coordinate. Frequency F is the normalized frequency and k_z is the integer spanwise wavenumber, defined in (3.1) and (3.2). (b–d) Instantaneous iso-surfaces of streamwise velocity component, (b) $u = 0.9$, (c) $u = 0.6$ and (d) $u = 0.3$, coloured by the streamwise velocity fluctuations u' . (e) The TKE transport terms integrated in the wall-normal direction along the transition zone.

For an arbitrary quantity \mathcal{X} , the Reynolds average is denoted by $\overline{\mathcal{X}}$, the Favre (density-weighted) average is identified by $\overline{\mathcal{X}^f}$ and fluctuations with respect to the Reynolds and Favre averages are represented as \mathcal{X}' and \mathcal{X}'' , respectively. The terms on the

right-hand side of (4.3) correspond to production rate P , dissipation rate ϵ , transport by velocity fluctuations \mathcal{T}_1 , transport by pressure fluctuation \mathcal{T}_2 , transport by viscous diffusion \mathcal{T}_3 , pressure dilatation (pressure–strain) Π , viscous mass flux coupling Σ_1 and pressure mass flux coupling Σ_2 .

In case E1N, transition involves a series of instability waves that are activated at various downstream locations (figure 9a): the inflow modes amplify first and, downstream, spur other instabilities which were not prominently featured in the inlet condition. Specifically, the inlet pair $\langle 100, 0 \rangle$ and $\langle 110, 0 \rangle$ activate mode $\langle 10, 0 \rangle$ – an interaction that we will denote $\mathcal{I}^{(1)}$. The resulting mode $\langle 10, 0 \rangle$ is manifest as a low-frequency modulation of the near-wall waves in figure 9(d). In a subsequent interaction $\mathcal{I}^{(2)}$, the pair $\langle 40, 3 \rangle$ and $\langle 110, 0 \rangle$ give rise to mode $\langle 70, 3 \rangle$, which is visible in figure 9(b) near $\sqrt{Re_x} = 2300$. The next interaction $\mathcal{I}^{(3)}$ involves the newly formed three-dimensional mode and the inflow wave $\langle 100, 0 \rangle$. It spurs the instability $\langle 30, 3 \rangle$ which grows at the highest rate, as evident in figure 9(b) in the range $2400 < \sqrt{Re_x} < 2550$, and is the seat of breakdown to turbulence (also see supplementary movie 1). Note that a triad $\mathcal{I}^{(4)}$ is established between the first two nonlinearly generated waves, $\langle 10, 0 \rangle$ and $\langle 40, 3 \rangle$, and the fastest amplifying instability $\langle 30, 3 \rangle$. Flow structures resembling interweaving rope-shaped waves evolve far from the wall prior to breakdown to turbulence in supplementary movie 1, and bear resemblance to those observed experimentally in second-mode transition at Mach 5–8 (Casper *et al.* 2016; Laurence *et al.* 2016).

Streaks feature in the instantaneous field (figures 9c,d). In the spectra, they correspond to mode $\langle 0, 6 \rangle$, which can be generated by the nonlinear interaction of modes $\langle \pm F, 3 \rangle$; their spanwise size is approximately equal to the thickness of the boundary layer at $\sqrt{Re_x} \approx 2600$. The energy of the streaks, or mode $\langle 0, 6 \rangle$, shadows that of mode $\langle 30, 3 \rangle$ which was discussed in connection with the outer Λ -shaped structure and breakdown to turbulence. The breakdown of the streaks follows almost immediately after that of the outer Λ shapes.

The interactions \mathcal{I} described above can only be hypothesized based on the spectra. For example, whether the streaks $\langle 0, 6 \rangle$ are due to the self interactions of $\langle \pm 30, 3 \rangle$, $\langle \pm 40, 3 \rangle$ or $\langle \pm 70, 3 \rangle$ cannot be conclusively determined; the similarity between $\langle \pm 30, 3 \rangle$ and $\langle \pm 0, 6 \rangle$ suggests a connection but does not guarantee it. In order to quantify the nonlinear interactions, we compute the energy transfer terms among wavenumber triads. Starting from \mathcal{T}_1 in (4.3), the following expression for \mathcal{I} can be derived (see appendix B for details):

$$\mathcal{I}_{\langle F, k_z \rangle} = \sum_{\pm F, \pm k_z} \int_0^{L_y} |\hat{\mathcal{A}}_{\langle F_1, k_{z,1} \rangle}^* \hat{\mathcal{B}}_{\langle F_2, k_{z,2} \rangle} + \hat{\mathcal{A}}_{\langle -F_2, -k_{z,2} \rangle}^* \hat{\mathcal{B}}_{\langle -F_1, -k_{z,1} \rangle}| dy, \tag{4.4}$$

where $F = F_2 - F_1$ and $k_z = k_{z,2} - k_{z,1}$. In equation (4.4), $\hat{\mathcal{A}}$ and $\hat{\mathcal{B}}$ are vector quantities containing the Fourier coefficients of $\mathcal{A} = [\rho u'' u'' \quad \rho u'' v'' \quad \rho u'' w'']^{tr}$ and of $\mathcal{B} = [u'' \quad v'' \quad w'']^{tr}$. The quantity $\mathcal{I}_{\langle F, k_z \rangle}$ thus represents the nonlinear energy transfer among modes $\langle F, k_z \rangle$, $\langle F_1, k_{z,1} \rangle$ and $\langle F_2, k_{z,2} \rangle$ (Cheung & Zaki 2010). Since $\mathcal{I}_{\langle F, k_z \rangle}$ is independent of the ordering of modes within the triad, it only measures the energy transfer among the three modes but not the direction of the transfer.

In case E1N, nonlinear interactions feature prominently and, therefore, $\mathcal{I}_{\langle F, k_z \rangle}$ for the seven most important interactions is plotted in figure 10. Note that the designation $\langle F_1, k_{z,1} \rangle + \langle F_2, k_{z,2} \rangle \Rightarrow \langle F, k_z \rangle$ in the figure is only intended to identify the triad, and bears no physical significance. However, together with the energy spectra, $\mathcal{I}_{\langle F, k_z \rangle}$

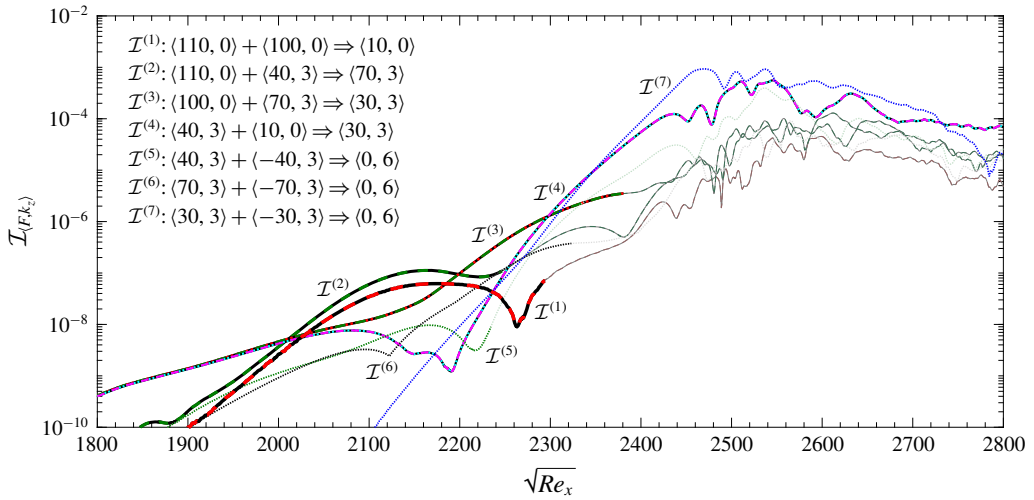


FIGURE 10. (Colour online) The modal nonlinear energy transfer coefficient, defined in (4.4), computed for the most important modal interactions corresponding to transition scenario of case EIN.

can provide a clearer view of the nonlinear interactions that spur new instability waves. The interaction $\mathcal{I}^{(1)}$ takes place between inflow modes $\langle 100, 0 \rangle$ and $\langle 110, 0 \rangle$ and generates $\langle 10, 0 \rangle$. Similarly, $\mathcal{I}^{(2)}$ takes place between inlet modes $\langle 40, 3 \rangle$ and $\langle 110, 0 \rangle$ and gives rise to $\langle 70, 3 \rangle$. This emergent mode has a strong interaction $\mathcal{I}^{(3)}$ with $\langle 100, 0 \rangle$ and leads to $\langle 30, 3 \rangle$ which amplifies at the highest rate. Note also that near $\sqrt{Re_x} \approx 2300$, this interaction is overtaken by another triad, $\mathcal{I}^{(4)}$, that involves $\langle 30, 3 \rangle$. The source of the streaks $\langle 0, 6 \rangle$ is not a single interaction, but rather three, $\mathcal{I}^{(5,6,7)}$, and each of them is dominant over a different region in the streamwise direction: $\mathcal{I}^{(5)}$ initiates the streaks in the region $\sqrt{Re_x} < 2150$; $\mathcal{I}^{(6)}$ takes over in the range $2150 < \sqrt{Re_x} < 2270$; and finally the interaction $\mathcal{I}^{(7)}$ is dominant beyond $\sqrt{Re_x} \approx 2270$, where the streaks become visible in the instantaneous visualizations (figure 9c,d), and where mode $\langle 0, 6 \rangle$ shadows $\langle 30, 3 \rangle$ in the spectra.

The role of every interaction, including those that precede the amplification of $\langle 30, 3 \rangle$, cannot be discounted. We performed exhaustive tests that involved assigning a significant portion of the inflow energy to mode $\langle 30, 3 \rangle$ but transition was delayed. Thus, all the events, including nonlinear interactions and base-flow distortion, that precede the formation of this mode must take place for transition to set in as upstream as recorded in our simulations.

Terms in the TKE equation are shown versus downstream Reynolds number in figure 9(e). Three regions can be distinguished, based on the behaviour of the rate of production P , which represents the energy exchange between the base state and the instability waves. (i) In the range $2400 < \sqrt{Re_x} < 2500$, the production increases faster than dissipation; in this region, the instability wave $\langle 30, 3 \rangle$ dominates the spectra and reaches saturation. (ii) In the region $2500 < \sqrt{Re_x} < 2550$, the rate of production increases as the secondary instability sets in and leads to the formation of the Λ -structures and their local breakdown. (iii) Finally, in the range $2550 < \sqrt{Re_x} < 2650$, the flow starts to approach the statistical state of a self-similar turbulent boundary layer.

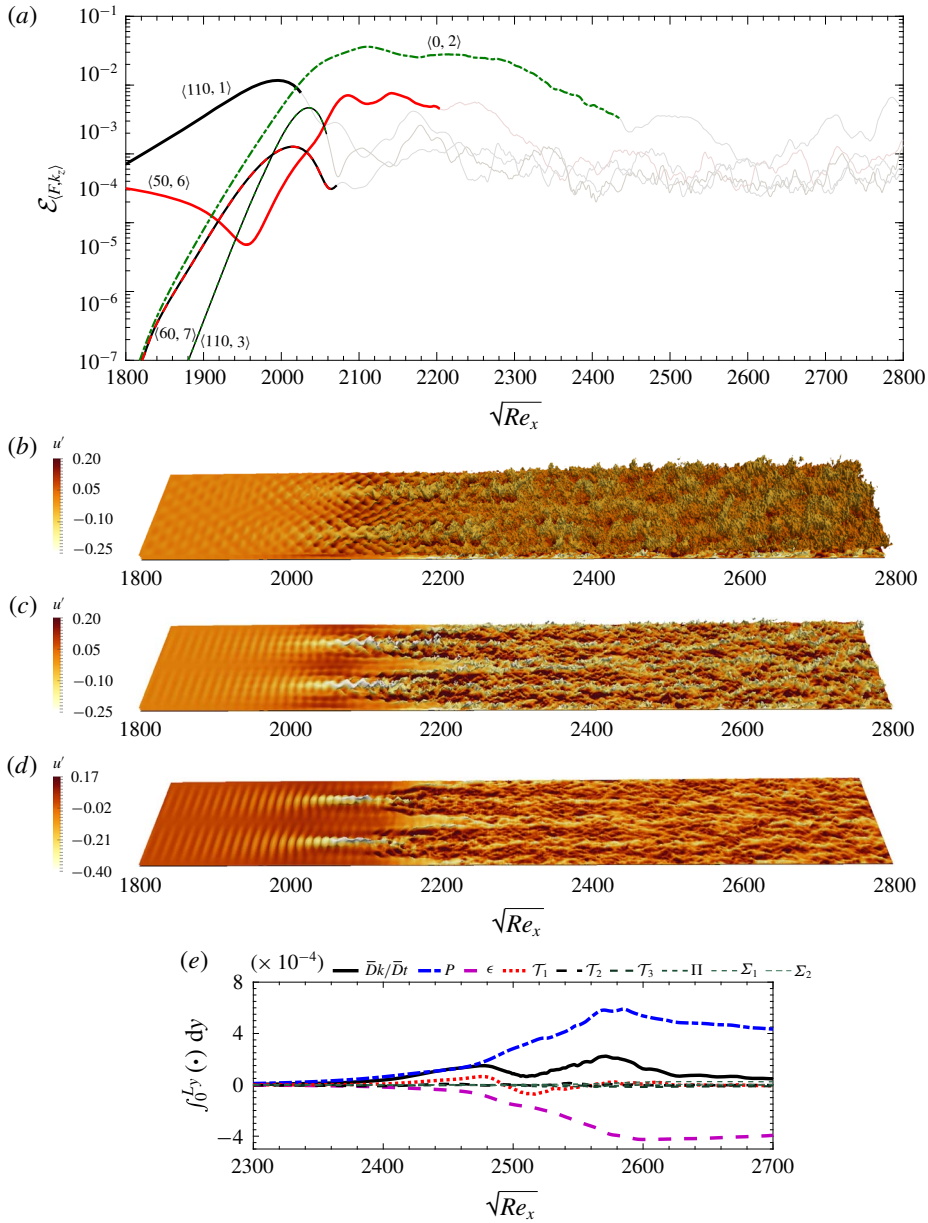


FIGURE 11. (Colour online) Results corresponding to case E2N. (a) Fluctuation energy content for selected instability modes, $\mathcal{E}_{(F,k_z)}$, versus the streamwise coordinate. Frequency F is the normalized frequency and k_z is the integer spanwise wavenumber, defined in (3.1) and (3.2). (b–d) Instantaneous iso-surfaces of streamwise velocity component, (b) $u = 0.9$, (c) $u = 0.6$ and (d) $u = 0.3$, coloured by the streamwise velocity fluctuations u' . (e) The TKE transport terms integrated in the wall-normal direction along the transition zone.

The transition scenario for case E2N is of different ilk. Figure 11(a) captures the fast emergence and amplification of a streamwise elongated streak mode $(0, 2)$. This wave becomes the seat of secondary instability and onset of turbulence.

Its instantaneous form is evident in the visualization of the streamwise velocity iso-surfaces (figure 11*b–d*). The spanwise size of the structure is nearly three times the boundary-layer thickness at $\sqrt{Re_x} \approx 2250$, which is the approximate location where the flow has reached a self-similar turbulent state. Streaks with commensurate sizes were reported in transition at $Ma_\infty = 3$ (Mayer *et al.* 2011) and also at $Ma_\infty = 6$ (Franko & Lele 2013).

The streak $\langle 0, 2 \rangle$ is generated by modes $\langle \pm 110, 1 \rangle$, which initially amplify linearly and peak near $\sqrt{Re_x} \approx 2000$ (cf. figure 11*a*) – the location where the skin-friction coefficient starts to rise (cf. figure 3*b*). At this stage, the mode $\langle 50, 6 \rangle$ starts to amplify, and reaches its highest energy level at $\sqrt{Re_x} \approx 2100$. In the instantaneous realization, it appears to affect the secondary instability of the streak mode $\langle 0, 2 \rangle$, which meanders and breaks down to turbulence (figure 11*b–d* and supplementary movie 2). This interpretation was verified by examining the spectra and nonlinear energy transfer terms (not shown here); we have also performed numerous linear parabolized stability equation studies of the evolution of instability waves atop the Blasius similarity profile and also the mean-flow profile from case E2N.

Despite describing case E2N as a second-mode oblique transition, it is not without differences from the typical configuration where nearly all the disturbance energy is assigned to the oblique mode and a very small amount is included in broadband noise. For E2N, high levels of energy are included in both modes $\langle 110, 1 \rangle$ and $\langle 50, 6 \rangle$ at the inlet. As a result, their nonlinear interaction is active early upstream and generates mode $\langle 60, 7 \rangle$ (figure 11*a*). This and other spurred oblique waves, e.g. $\langle 110, 3 \rangle$, promote breakdown of the streaky boundary layer to turbulence.

The upstream stages of transition in supplementary movie 2 exhibit alternation of light and dark regions within the bulk of the boundary layer. These regions amplify downstream and ultimately break down into turbulence. While rope-like structures appear near the boundary-layer edge close to transition onset, they do not undergo the spatial growth observed in case E1N. The qualitative change in transition at different levels of disturbance energy is important. A parallel, although not a one-to-one correspondence, can be drawn to the qualitative changes observed in the recent experiments of transition on a slender 7° half-angle cone at Mach 6–7, conducted at low- and high-enthalpy conditions (Laurence *et al.* 2016).

We can further examine the transition zone of case E2N by considering the evolution of terms in the TKE equation in various subregions (figure 11*e*). In a similar manner to the discussion of case E1N, three regions can be identified. (i) In the range $2000 < \sqrt{Re_x} < 2100$, the increase in the rate of production, P , is moderate but P itself exceeds the integrated ϵ . The energy of the streaks overtakes all other modes at the start of this region, and their secondary instability is amplifying. (ii) In the region $2100 < \sqrt{Re_x} < 2200$, the secondary instability $\langle 50, 6 \rangle$ is nearly saturated (cf. figure 11*a*), and localized patches of turbulence emerge; here the increases in the rate of production and dissipation become steeper, the latter enhanced by the emergence of small turbulent scales. Within the localized patches of turbulence, production exceeds the levels of fully turbulent boundary layers (Marxen & Zaki 2019). (iii) In the region $2200 < \sqrt{Re_x} < 2300$, the turbulence spread and fills the domain, and terms in the TKE equation start to relax towards their equilibrium values for self-similar turbulent boundary layers.

The transition processes in cases E1N and E2N are contrasted in figures 12 and 13. The former figure shows wall-normal profiles of the mean streamwise velocity and temperature, evaluated at different streamwise positions within the transition zone. And the latter figure shows three-dimensional views of the vortical structures,

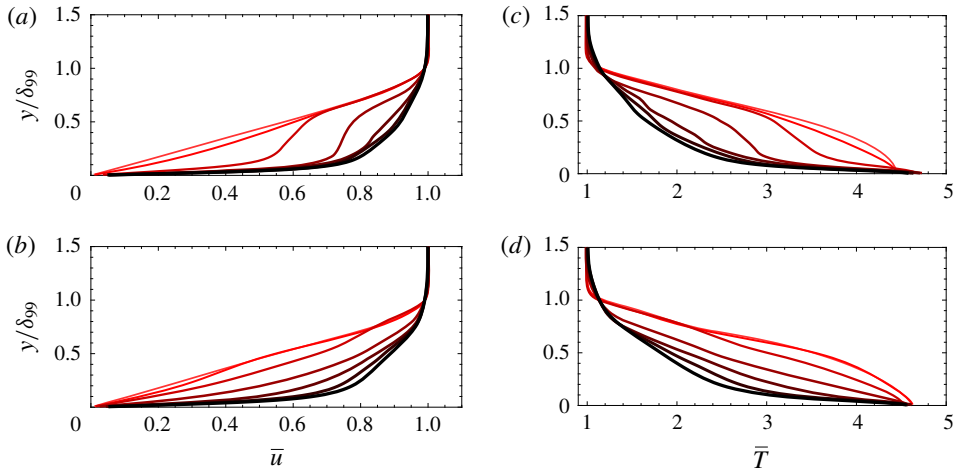


FIGURE 12. (Colour online) Wall-normal profiles of time- and spanwise-averaged streamwise velocity \bar{u} (a,b) and temperature \bar{T} (c,d) at different locations along the transition zone. (a,c) The profiles of case E1N correspond to $2400 \leq \sqrt{Re_x} \leq 2700$ with increment of 50 from light to dark colours (also thin to thick lines). (b,d) The profiles of case E2N correspond to $2000 \leq \sqrt{Re_x} \leq 2300$ with increment of 50 from light to dark colours (also thin to thick lines).

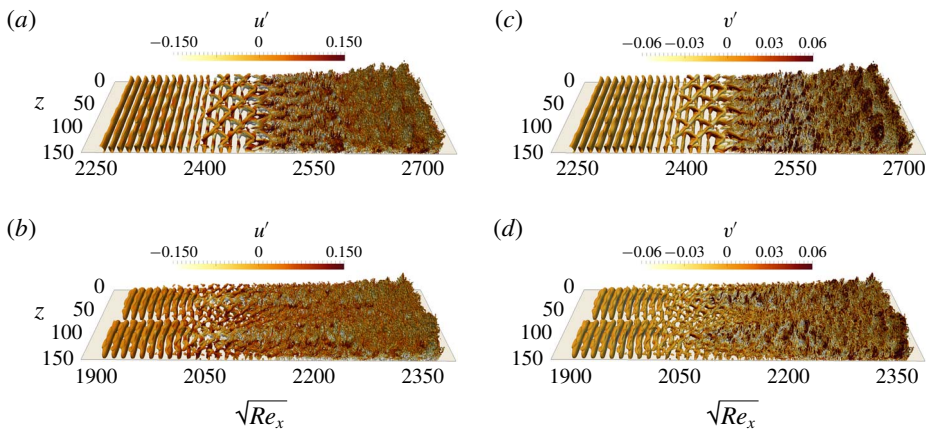


FIGURE 13. (Colour online) Iso-surfaces of $Q = 1 \times 10^{-4}$ coloured by u' (a,b) and v' (c,d) corresponding to cases (a,c) E1N and (b,d) E2N.

visualized using the Q -criterion, within the transition regions of both cases. Transition in case E1N takes place via the formation and breakdown of Λ -shaped vortices. This process takes place over a relatively short streamwise distance as shown qualitatively in figure 13 and demonstrated by the skin-friction coefficient along the plate (figure 3). The relatively abrupt transition is accompanied by a distortion of the mean streamwise velocity profile in the near-wall region, followed by a subsequent relaxation of the outer part of the profile towards the fully turbulent curve. The same trend is captured in the mean temperature profile. In case E2N, transition is due to the secondary instability of streamwise-elongated streaks. While the streaks themselves are not

captured in the visualization of the Q -criterion, their secondary instability is seen in figure 13(b,d). The transition length is longer in this case, and therefore the mean velocity and temperature profiles gradually approach the turbulent curve.

For case E1N, figure 13(a,c) shows that the Λ -structures, which are generated as a result of the amplification of mode $\langle 30, 3 \rangle$, are located close to the edge of the boundary layer. They lie above the near-wall streaks shown in figure 9(d), and successive rows of Λ -structures are staggered in the spanwise direction. Typical of hairpin structures in wall-bounded flows (see e.g. Adrian 2007; Farano *et al.* 2015), their legs lead to ejections in the plane of symmetry while the head generates a strong sweep motion. The wall-normal and streamwise velocity fluctuations therefore have opposite signs in those regions. Ejections due to positive wall-normal velocity fluctuations are accompanied by transport of low streamwise momentum upward and a negative streamwise fluctuation; sweep due to negative v' leads to downward transport of high-momentum fluid and a positive u' perturbation. The late stages in the evolution of these Λ -structures resemble natural transition to turbulence, where trains of hairpin vortices are formed and break down over a short streamwise distance. Since the structures are staggered in the span, a fully turbulent flow is established quickly downstream thus leading to a short transition length.

Figure 13(b,d) for case E2N shows that Λ -structures form in this case as well, and successive rows are similarly staggered in the spanwise direction. However, breakdown to turbulence does not commence at the tips of individual structures; instead it is initiated due to the instability of the underlying steady streaks. These streaks are clearly captured in figure 11(d), and at the same locations the breakdown pattern is clear in figure 13(b,d). In this region, the streaks are elevated low-speed structures (negative u') straddled by Λ -structures that induce their ejection (positive v'). Since breakdown takes place on the low-speed streaks, it directly impacts every other row of the Λ -structures and subsequently spreads to fill the domain; the transition length is therefore longer in this case relative to E1N.

4.3. Prediction from linear stability theory

In order to highlight the importance of the present nonlinear approach in determining the inflow disturbance spectrum, we examine other possible inflow conditions that are selected based on linear theory alone. The starting point is to evaluate the linear evolution of all the Orr–Sommerfeld and Squire instability waves that are part of the inlet condition, using the linear parabolized instability equations (Park & Zaki 2019). The N -factor associated with every $\langle F, k_z \rangle$ wave was obtained as follows:

$$N\text{-factor} = \frac{1}{2} \ln \left(\frac{\mathcal{E}_{(F,k_z)}}{\mathcal{E}_0} \right), \quad (4.5)$$

where \mathcal{E}_0 is the energy of the mode at the inlet. The results are shown for three downstream locations in figure 14(a). Depending on the point at which the growth rates are computed, different instability waves attain the highest N -factor value. Figure 14(b) shows the change of N -factor versus streamwise location for the four instability waves that are each most amplified within a subregion of the computational domain: mode $\langle 110, 0 \rangle$ for $\sqrt{Re_x} < 2190$, mode $\langle 100, 0 \rangle$ along $2190 < \sqrt{Re_x} < 2470$, mode $\langle 90, 0 \rangle$ along $2470 < \sqrt{Re_x} < 2650$ and mode $\langle 20, 2 \rangle$ for $\sqrt{Re_x} > 2650$. In order to make use of these results in selecting the inflow condition, prior knowledge of the transition location is needed; linear theory does not provide this information.

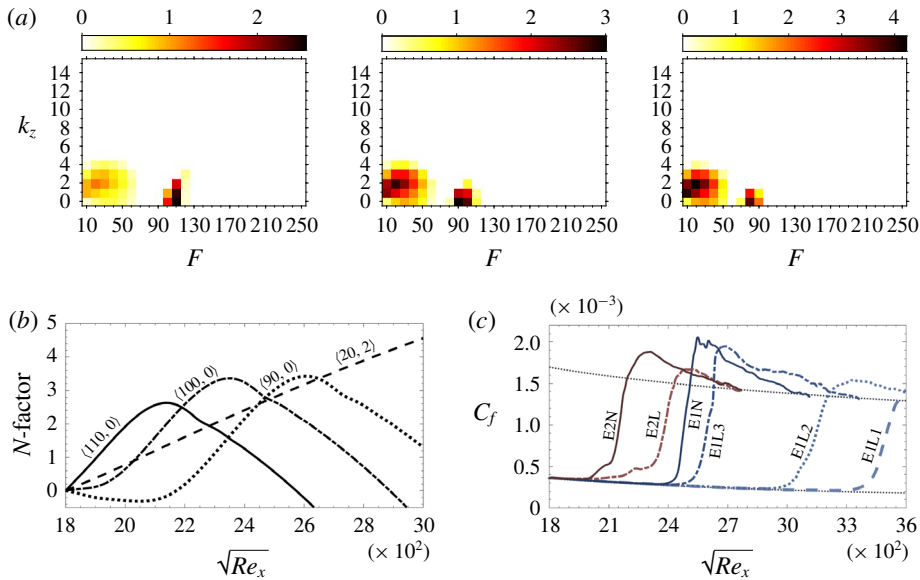


FIGURE 14. (Colour online) (a) Contours of N -factor computed from (4.5) from left to right for $\sqrt{Re_{x_f}} = 2100$, 2500 and 2900, respectively. (b) The N -factor of the important modes versus streamwise location. (c) Skin friction curves corresponding to the nonlinearly most dangerous cases (E1N and E2N) and the linearly most unstable cases (E1L1, E1L2, E1L3 and E2L) presented in table 2.

We examined four inflow conditions that were selected based on the linear results, and computed their evolution using direct numerical simulations in order to compare the outcome to the nonlinearly most dangerous disturbance. These new cases are designated E1L1, E1L2, E1L3 and E2L, and their parameters are reported in table 2. The first three are at the lower inflow energy level and, since transition is expected in the second half of the domain, 99% of the inflow energy is allocated to modes $(100, 0)$, $(90, 0)$ and $(20, 2)$, respectively. The fourth case, E2L, is at the higher inflow energy level and, therefore, 99% of the inflow energy is assigned to mode $(110, 0)$. For all four cases, the remaining 1% of the inflow energy was in the form of broadband forcing, in order to enable possible nonlinear interactions and secondary instabilities. The outcome of these simulations is contrasted to the nonlinearly most dangerous cases in figure 14(c), where C_f is plotted versus downstream distance. Note that, for some of these simulations, the domain was extended in the streamwise direction in order to capture the delayed transition to turbulence, in particular in E1L1 and E1L2.

The transition scenarios in the new simulations are very different from those due to the nonlinearly optimal inflow conditions. The streamwise development of the energy of associated key modes is reported in figure 2 of the supplementary material. Cases E1L1 and E1L2 are akin to classical two-dimensional second-mode transition, with a region of oblique secondary instability prior to breakdown. Case E1L3 is a typical first-mode oblique breakdown, in which the primary three-dimensional wave develops oblique instabilities, Λ -structures and ultimately hairpin vortices that become the seat for the onset of turbulence. Transition in case E2L is a typical second-mode fundamental scenario, where the oblique secondary instability has the same frequency as the primary two-dimensional wave.

In all four cases, the inflow primary instability is the key determinant of the path to turbulence. And while the linearly most unstable waves guarantee fastest exponential growth based on linear theory, they do not necessarily guarantee largest energy in the nonlinear regime or earliest secondary instability and transition. In contrast, when the nonlinearly most dangerous inflow condition was adopted, the resulting transition mechanism was not due to any one particular inflow wave; instead, the nonlinear interactions of the inflow instabilities led to the generation of new waves, the distortion of the base flow and the earliest possible transition location.

5. Summary and conclusions

An EnVar algorithm is introduced to evaluate the nonlinearly most dangerous inflow disturbance that results in the earliest location of laminar-to-turbulence transition in a high-speed boundary layer. This disturbance (i) has a prescribed amount of total energy, (ii) satisfies the full Navier–Stokes equations and (iii) undergoes the fastest breakdown to turbulence. The first two elements are constraints, and the third is the objective of the optimization procedure and was modelled using a cost function that is proportional to the integral of the skin friction along the plate. The EnVar algorithm starts from an estimate of the optimal inflow condition, which in the present work is initialized to be a broadband disturbance. An ensemble of possible solutions is then constructed around the current estimate, and all members are advanced using the Navier–Stokes equations. The associated outcomes are used to evaluate the gradient of the cost function which, in turn, is used to update the estimate of the optimal inflow condition. A new ensemble is then formed around the estimated solution, and the process is repeated until convergence.

The algorithm was applied in the case of ZPG boundary-layer flow at free-stream Mach number $M_\infty = 4.5$, and for two levels of the inflow disturbance energy. The level of energy of the first case is of the same order as the kinetic energy of stratospheric turbulent layers, measured during a recent experimental campaign in an attempt to replicate the environmental conditions of high-altitude flight tests. The second case had energy 50 times larger, in order to highlight the effect of this parameter on the outcome of the algorithm. The laminar-to-turbulence transition scenarios due to the nonlinearly most dangerous disturbances from each case were examined in detail. While at the higher energy level, transition displays symptoms of second-mode oblique breakdown, the lower energy case cannot be ascribed to any classical route. Instead, transition is initiated due to nonlinear interactions of a couple of normal acoustic second modes and an oblique vortical first mode. The interactions spur new instabilities and distort the mean flow, ultimately leading to very rapid growth of three-dimensional disturbances that form Λ -structures and break down to turbulence.

A number of other inflow conditions, all selected based on the amplification rates from linear theory, were also examined using direct numerical simulations. These efforts invariably led to classical breakdown scenarios and, for the same level of inflow energy, delayed transition onset relative to the nonlinearly most dangerous disturbance. The results thus highlight that a nonlinear approach to the design of the inflow condition is required in order to obtain the minimum transition Reynolds number. We expressed the inflow disturbance as a superposition of instability waves, and optimized their amplitudes and relative phases using the full Navier–Stokes equations. This approach bridges the gap between LST which excludes the nonlinear terms and transition which must involve nonlinearity in order to effect the energy cascade that sustains turbulence.

The EnVar approach is applicable with other choices of the cost function, for example based on the downstream development of the disturbance energy or wall temperature. The behaviours of these two quantities were examined when the cost function was based on skin friction. The results demonstrated that the optimal inflow disturbance is likely to differ, in particular if the cost function is based on wall temperature. Despite the present use of direct numerical simulations, the EnVar approach is compatible with any forward model of the Navier–Stokes equations (e.g. nonlinear parabolized stability equations, large-eddy simulations). It is therefore attractive to adopt EnVar with a lower fidelity method, in order to reduce the computational cost and to explore new directions, e.g. search for global optima, the dependence of transition Reynolds number on the energy of the inflow disturbance and the impact of the cost function.

The nonlinearly most dangerous disturbance provides an objective measure for evaluating design performance, in particular when environmental disturbances are uncertain. If a design is modified, for instance for the purpose of delaying transition, the spectral content of what constitutes the most dangerous disturbance will, however, change and must be re-evaluated. If transition is delayed in the new configuration, design surety is guaranteed.

Acknowledgements

This work was supported in part by the US Air Force Office of Scientific Research (grant no. FA9550-16-1-0103) and by the Office of Naval Research (grant no. N00014-17-1-2339). Computational resources were provided by the Maryland Advanced Research Computing Center (MARCC). The authors are grateful to Professor J. Larsson for sharing the Hybrid code.

Supplementary data

Supplementary data are available at <https://doi.org/10.1017/jfm.2019.527>.

Appendix A. Ensemble generation algorithm

In this appendix, the algorithm for constructing the ensemble members that are used in the optimization procedure (2.26) is described. The starting point is the mean control vector, $\mathbf{c}^{(e)}$, which is either the initial guess or the outcome of the optimization at the end of the previous iteration, and satisfies the energy constraint. The objective is to generate an ensemble of control vectors, $\mathbf{c}^{(r)}$ for $r = 1$ to N_{en} , around this mean; the relationship between $\mathbf{c}^{(e)}$ and $\mathbf{c}^{(r)}$ is given by (2.19), and therefore $\mathbf{c}^{(e)}$ is not the arithmetic average of the members. The latter must not deviate appreciably from the mean, in a sense that will be quantified using the variance of the ensemble, and must each satisfy the energy constraint as well. Another important consideration is that, for a particular ensemble size N_{en} , the members must span the control vector subspace as best as possible and be well-conditioned. The ensemble generation algorithm has two main steps: first, a very large random ensemble is formed that satisfies specific criteria; second, a smaller ensemble is determined from the larger one, and whose members are better conditioned. A detailed description of the first step is presented here, while for the second step we only quote the procedure and refer the reader to Evensen (2009, chap. 11) for its theoretical foundation and proof of concept.

In the first step of the algorithm, our concern is to enforce the constraints that the ensemble members should satisfy rather than how well the ensemble is conditioned.

Therefore, in order to ensure that the members span the control-vector subspace, we generate a very large ensemble. Since the members will be randomly generated, and assuming a very large size of the ensemble, it is very likely that any possible control vector can be expressed as a linear superposition of the ensemble members. There are three constraints that must be satisfied, related to the mean, covariance and energy. As discussed in § 2.2, satisfying the energy constraint by the mean and ensemble members is not possible if the arithmetic average is adopted. In order to address this challenge, we performed two change of variables during the algorithm, referred to as \mathbf{e} and $\hat{\mathbf{e}}$ later in this appendix. And the covariance constraint is to ensure that the members of the ensemble are close to their mean, which is an essential condition of the EnVar optimization procedure.

We start by constructing the very large ensemble of size $\Upsilon \times N_{en}$. In analogy to (2.19), the ensemble members are related to the mean by

$$\mathbf{A}_1 \mathbf{c}^{(e)} \circ \mathbf{A}_1 \mathbf{c}^{(e)} = \frac{1}{\Upsilon N_{en}} \sum_{r=1}^{\Upsilon N_{en}} \mathbf{A}_1 \mathbf{c}^{(r)} \circ \mathbf{A}_1 \mathbf{c}^{(r)}, \quad (\text{A } 1a)$$

and

$$\mathbf{A}_2 \mathbf{c}^{(e)} = \frac{1}{\Upsilon N_{en}} \sum_{r=1}^{\Upsilon N_{en}} \mathbf{A}_2 \mathbf{c}^{(r)}, \quad (\text{A } 1b)$$

where \mathbf{A}_1 and \mathbf{A}_2 are defined in equations (2.12) and (2.20), respectively.

The above two equalities can be encoded into one expression by defining $\mathbf{e}^{(r)} \in \mathbb{R}^{2M \times 1}$:

$$\mathbf{e}^{(r)} \equiv \mathbf{A}_1 \mathbf{c}^{(r)} \circ \mathbf{A}_1 \mathbf{c}^{(r)} + \mathbf{A}_2 \mathbf{c}^{(r)}, \quad (\text{A } 2)$$

whose arithmetic average is $\mathbf{e}^{(e)}$. A further change of variables is introduced in order to simplify the algorithm

$$\hat{\mathbf{e}}^{(r)} \equiv \mathbf{S}^{-1}(\mathbf{e}^{(r)} - \mathbf{e}^{(e)}), \quad (\text{A } 3)$$

where the diagonal matrix \mathbf{S} is equal to $\text{diag}(\boldsymbol{\sigma})$ and $\boldsymbol{\sigma}$ is a vector containing the desired standard deviations of the original ensemble members from their mean. By construction, $\hat{\mathbf{e}}^{(r)}$ has a zero mean and unit variance. Furthermore, as long as the members of the original ensemble and their mean, $\mathbf{e}^{(r)}$ and $\mathbf{e}^{(e)}$, satisfy the energy constraint, the energy of $\hat{\mathbf{e}}^{(r)}$ is zero. Thus the ensemble that we are seeking satisfies the following conditions:

$$\hat{\mathbf{P}} \mathbf{b}_1 = 0, \quad (\text{A } 4a)$$

$$\hat{\mathbf{P}} \hat{\mathbf{P}}^{tr} = (\Upsilon N_{en} - 1) \hat{\mathbf{C}}, \quad (\text{A } 4b)$$

and

$$\mathbf{b}_1^{tr} \mathbf{A}_1 \hat{\mathbf{P}} = 0, \quad (\text{A } 4c)$$

where

$$\hat{\mathbf{P}} = [\hat{\mathbf{e}}^{(1)} \quad \hat{\mathbf{e}}^{(2)} \quad \dots \quad \hat{\mathbf{e}}^{(\Upsilon N_{en})}]_{2M \times \Upsilon N_{en}}, \quad (\text{A } 5)$$

$$\hat{\mathbf{C}} = \begin{bmatrix} [\mathcal{C}_{ij}]_{M \times M} & \mathbf{O}_{M \times M} \\ \mathbf{O}_{M \times M} & \mathbf{I}_{M \times M} \end{bmatrix}_{2M \times 2M}, \tag{A 6}$$

$$\mathbf{b}_1 = [1 \quad 1 \quad \dots \quad 1]_{1 \times 2M}^T. \tag{A 7}$$

In the above, $\mathcal{C}_{ij} = \mathcal{C}_{ji}$ are the correlation coefficients and $\mathcal{C}_{ii} = 1$, \mathbf{I} is the identity matrix, \mathbf{O} is a zero matrix and \mathbf{A}_1 is defined in (2.12). Equation (A 4a) ensures that the mean of the ensemble is zero, equation (A 4b) is the covariance constraint and (A 4c) enforces the energy constraint on all members of the ensemble. Note that the condition $\mathcal{C}_{ii} = 1$ is a consequence of the normalization in (A 3).

The problem can be simplified further by combining the covariance and the energy constraints. From (A 4b) and (A 4c), we obtain $\mathbf{b}_1^T \mathbf{A}_1 \hat{\mathbf{C}} = 0$, and therefore the energy constraint is equivalent to $\sum_{j=1}^M \mathcal{C}_{ij} = 0$. Thus, we can redefine the problem as evaluating a set of $\hat{\mathbf{e}}$ whose arithmetic average is zero, and that are correlated with one another such that the components of their covariance matrix satisfy $\sum_{j=1}^M \mathcal{C}_{ij} = 0$. The former condition is satisfied by randomly generating uncorrelated control vectors, $\mathbf{y}^{(r)}$, that have a zero mean; and the latter condition is enforced by performing Cholesky factorization of the covariance matrix and setting $\hat{\mathbf{e}}^{(r)} = \mathbf{L}\mathbf{y}^{(r)}$, where \mathbf{L} is the lower triangular matrix of the factorization. By definition, $\hat{\mathbf{e}}^{(r)}$ satisfy both conditions. Using the symmetry of $\hat{\mathbf{C}}$, we can obtain $\hat{\mathbf{C}}\mathbf{A}_1\mathbf{b}_1 = 0$. Hence, zero is an eigenvalue of $\hat{\mathbf{C}}$ and $\mathbf{A}_1\mathbf{b}_1$ is an eigenvector; the dimension of the eigenspace corresponding to this eigenvalue is unity, and the rank of $\hat{\mathbf{C}}$ is equal to $2M - 1$. A new covariance matrix is defined whose rank is equal to $2M$ as

$$\hat{\hat{\mathbf{C}}} = \begin{bmatrix} 1 & \mathbf{o}^T \\ \mathbf{o} & \hat{\mathbf{C}} \end{bmatrix}_{(2M+1) \times (2M+1)}, \tag{A 8}$$

where $\mathbf{o} \in \mathbb{R}^{2M \times 1}$ is a vector of all zeros. Note that zero is still an eigenvalue of $\hat{\hat{\mathbf{C}}}$. Therefore, a reconstruction from its non-zero singular values and associated singular vectors is used in the Cholesky factorization, in order to obtain the lower triangular matrix $\mathbf{L} \in \mathbb{R}^{2M \times 2M}$.

The above procedure completes the first step of the ensemble generation, and provides a set of $\hat{\mathbf{e}}^{(r)} \in \mathbb{R}^{2M \times 1}$ which satisfy (A 4). The second step of the algorithm is to identify, from this large set of members, the best possible ensemble with size N_{en} . This task involves singular-value decomposition of the large ensemble and choosing the N_{en} dominant singular vectors and values as the new ensemble. In the limit of $\gamma N_{en} \rightarrow \infty$, the singular vectors and values of the large ensemble will converge to eigenvectors and eigenvalues of the full-rank ensemble realizations (an ensemble that represents the physical problem exactly). Using the singular-value decomposition of the larger-sized ensemble to approximate its smaller-sized representation, therefore, ensures that for a given size the new ensemble provides the best possible rendition. The complete procedure for the ensemble generation is summarized in algorithm 2.

Appendix B. Nonlinear energy transfer

In this appendix, the derivation of the nonlinear energy transfer term $\mathcal{I}_{(F, k_z)}$ is presented, where

$$\mathcal{I}_{(F, k_z)} = \sum_{\pm F, \pm k_z} \int_0^{L_y} \left| \hat{\mathcal{A}}_{(F_1, k_{z,1})}^* \hat{\mathcal{B}}_{(F_2, k_{z,2})} + \hat{\mathcal{A}}_{(-F_2, -k_{z,2})}^* \hat{\mathcal{B}}_{(-F_1, -k_{z,1})} \right| dy. \tag{B 1}$$

Algorithm 2: Generation of the ensemble.

Step 1:

- Generate $\mathcal{Y}N_{en}$ random vectors $\mathbf{y}^{(r)}$ whose mean and covariance are 0 and \mathbf{I} .
- Find $-1 \leq \mathcal{C}_{ij} \leq 1$ that satisfies $\sum_{j=1}^M \mathcal{C}_{ij} = 0$ for all $1 \leq i \leq M$.
- From non-zero singular values of $\hat{\mathbf{C}}$ (A8) and associated vectors, construct a full rank covariance matrix. By Cholesky factorization, obtain $\mathbf{L} \in \mathbb{R}^{2M \times 2M}$.
- Compute $\hat{\mathbf{e}}^{(r)} \in \mathbb{R}^{2M \times 1}$ as $\hat{\mathbf{e}}^{(r)} = \mathbf{L}\mathbf{y}^{(r)}$, and store them in $\hat{\mathbf{P}} \in \mathbb{R}^{2M \times \mathcal{Y}N_{en}}$.

Step 2:

- Compute the singular value decomposition $\hat{\mathbf{P}} = \hat{\mathbf{U}}\hat{\mathbf{\Sigma}}\hat{\mathbf{V}}^T$.
- Retain the first N_{en} singular vectors in $\hat{\mathbf{U}}$, and store them in $\mathbf{U} \in \mathbb{R}^{2M \times N_{en}}$.
- Retain the first $N_{en} \times N_{en}$ quadrant of $\hat{\mathbf{\Sigma}}$, and store it in $\mathbf{\Sigma} \in \mathbb{R}^{N_{en} \times N_{en}}$.
- Retain the first $N_{en} \times N_{en}$ quadrant of $\hat{\mathbf{V}}$, and store it in $\mathbf{V} \in \mathbb{R}^{N_{en} \times N_{en}}$.
- Generate the new ensemble perturbations matrix, $\mathbf{P} = \mathbf{U} \frac{1}{\sqrt{\mathcal{Y}}} \mathbf{\Sigma} \mathbf{V}^T$.

Normalization of $\mathbf{\Sigma}$ by $\sqrt{\mathcal{Y}}$ ensures the new ensemble has the correct variance.

- Compute $\mathbf{e}^{(r)} = \mathbf{S}\mathbf{p}^{(r)} + \mathbf{e}^{(e)}$, where $\mathbf{p}^{(r)}$ are the columns of $\mathbf{P} \in \mathbb{R}^{2M \times N_{en}}$.
 - Compute the control vectors $\mathbf{c}^{(r)}$ from the set of $\mathbf{e}^{(r)}$ using equation (A2).
-

Superscript $*$ denotes complex-conjugate transpose and $|\cdot|$ is the absolute value. The starting point is the transport term \mathcal{T}_1 in the TKE equation (4.3), which is a nonlinear redistribution of energy between different perturbations. While all the other contributions in the energy equation originate from linear terms of the momentum balance, \mathcal{T}_1 arises from the nonlinear advection term.

By integrating \mathcal{T}_1 in the wall-normal direction, we obtain

$$\int_y \mathcal{T}_1 dy = -\frac{1}{2} \frac{\partial}{\partial x} \left(\underbrace{\int_0^{L_y} (\mathcal{A} \cdot \mathcal{B}) dy}_{\mathcal{I}} \right) = -\frac{1}{2} \frac{\partial}{\partial x} \left(\underbrace{\int_0^{L_y} (\mathcal{A}^* \mathcal{B}) dy}_{\mathcal{I}} \right), \quad (\text{B } 2)$$

where the vector quantities are defined as $\mathcal{A}_i = \rho(u_j'' \delta_{1j})u_i''$ and $\mathcal{B}_i = u_i''$, and the variable \mathcal{I} represents the nonlinear interaction. In the above expression, it was assumed that the fluctuations vanish at the top and bottom boundaries of the domain and, as such, only the streamwise advection is retained. Using the two-dimensional Fourier representations of the vectors \mathcal{A} and \mathcal{B} in temporal frequency and spanwise wavenumber, \mathcal{I} can be rewritten as

$$\mathcal{I} = \int_0^{L_y} \left(\sum_{a,p} \sum_{b,q} (\hat{\mathcal{A}}_{\langle F_a, k_{z,p} \rangle}^* \hat{\mathcal{B}}_{\langle F_b, k_{z,q} \rangle}) e^{-i \left(\frac{2\pi(-k_{z,p} + k_{z,q})}{L_z} z + \frac{\sqrt{Re_{x_0}}(-F_a + F_b)}{10^6} t \right)} \right) dy. \quad (\text{B } 3)$$

A new variable is defined as $\hat{\mathcal{L}}_{\langle F, k_z \rangle} = \sum_{a,p} \sum_{b,q} \hat{\mathcal{A}}_{\langle F_a, k_{z,p} \rangle}^* \hat{\mathcal{B}}_{\langle F_b, k_{z,q} \rangle}$, where the summation is performed over all $\langle F_a, k_{z,p} \rangle$ and $\langle F_b, k_{z,q} \rangle$ that satisfy $F = F_b - F_a$ and $k_z = k_{z,q} - k_{z,p}$.

In terms of $\hat{\mathcal{I}}_{\langle F, k_z \rangle}$, equation (B 3) becomes

$$\mathcal{I} = \sum_{F, k_z} \left(\int_0^{L_y} \hat{\mathcal{I}}_{\langle F, k_z \rangle} dy \right) e^{-i \left(\frac{2\pi k_z}{L_z} z + \frac{\sqrt{Re_{x0}} F}{10^6} t \right)}. \quad (\text{B } 4)$$

Therefore the quantity $\hat{\mathcal{I}}_{\langle F, k_z \rangle}$ represents the total amount of energy transferred into mode $\langle F, k_z \rangle$ due to the interaction of all relevant modal pairs (Cheung & Zaki 2010). We here focus our attention on the energy transferred into $\langle +F, +k_z \rangle$ due to the specific pair $\langle F_1, k_{z,1} \rangle$ and $\langle F_2, k_{z,2} \rangle$; the associated nonlinear modal energy transfer coefficient can be defined as

$$\mathcal{I}_{\langle +F, +k_z \rangle} = \int_0^{L_y} \left| \hat{\mathcal{A}}_{\langle F_1, k_{z,1} \rangle}^* \hat{\mathcal{B}}_{\langle F_2, k_{z,2} \rangle} + \hat{\mathcal{A}}_{\langle -F_2, -k_{z,2} \rangle}^* \hat{\mathcal{B}}_{\langle -F_1, -k_{z,1} \rangle} \right| dy. \quad (\text{B } 5)$$

Ultimately, we add the contributions from all four quadrants, and adopt the shorthand

$$\mathcal{I}_{\langle F, k_z \rangle} = \sum \mathcal{I}_{\langle \pm F, \pm k_z \rangle} = \sum_{\pm F, \pm k_z} \int_0^{L_y} \left| \hat{\mathcal{A}}_{\langle F_1, k_{z,1} \rangle}^* \hat{\mathcal{B}}_{\langle F_2, k_{z,2} \rangle} + \hat{\mathcal{A}}_{\langle -F_2, -k_{z,2} \rangle}^* \hat{\mathcal{B}}_{\langle -F_1, -k_{z,1} \rangle} \right| dy. \quad (\text{B } 6)$$

REFERENCES

- ADRIAN, R. J. 2007 Hairpin vortex organization in wall turbulence. *Phys. Fluids* **19** (4), 041301.
- BERTOLOTI, F. P., HERBERT, T. & SPALART, P. R. 1992 Linear and nonlinear stability of the Blasius boundary layer. *J. Fluid Mech.* **242**, 441–474.
- BEWLEY, T. R., MOIN, P. & TEMAM, R. 2001 DNS-based predictive control of turbulence: an optimal benchmark for feedback algorithms. *J. Fluid Mech.* **447**, 179–225.
- CASPER, K. M., BERESH, S. J., HENFLING, J. F., SPILLERS, R. W., PRUETT, B. O. M. & SCHNEIDER, S. P. 2016 Hypersonic wind-tunnel measurements of boundary-layer transition on a slender cone. *AIAA J.* **54** (1), 1250–1263.
- CHANG, C.-L. & MALIK, M. R. 1994 Oblique-mode breakdown and secondary instability in supersonic boundary layers. *J. Fluid Mech.* **273**, 323–360.
- CHERUBINI, S., DE PALMA, P., ROBINET, J.-C. & BOTTARO, A. 2011 The minimal seed of turbulent transition in the boundary layer. *J. Fluid Mech.* **689**, 221–253.
- CHERUBINI, S., ROBINET, J.-C. & DE PALMA, P. 2013 Nonlinear control of unsteady finite-amplitude perturbations in the Blasius boundary-layer flow. *J. Fluid Mech.* **737**, 440–465.
- CHEUNG, L. C. & ZAKI, T. A. 2010 Linear and nonlinear instability waves in spatially developing two-phase mixing layers. *Phys. Fluids* **22** (5), 052103.
- COLBURN, C. H., CESSNA, J. B. & BEWLEY, T. R. 2011 State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter. *J. Fluid Mech.* **682**, 289–303.
- EVENSEN, G. 2009 *Data Assimilation: The Ensemble Kalman Filter*. Springer.
- FARANO, M., CHERUBINI, S., ROBINET, J.-C. & DE PALMA, P. 2015 Hairpin-like optimal perturbations in plane poiseuille flow. *J. Fluid Mech.* **775**, R2.
- FEDOROV, A. 2011 Transition and stability of high-speed boundary layers. *Annu. Rev. Fluid Mech.* **43**, 79–95.
- FRANKO, K. J. & LELE, S. 2014 Effect of adverse pressure gradient on high speed boundary layer transition. *Phys. Fluids* **26** (2), 024106.
- FRANKO, K. J. & LELE, S. K. 2013 Breakdown mechanisms and heat transfer overshoot in hypersonic zero pressure gradient boundary layers. *J. Fluid Mech.* **730**, 491–532.

- GAO, X., WANG, Y., OVERTON, N., ZUPANSKI, M. & TU, X. 2017 Data-assimilated computational fluid dynamics modeling of convection-diffusion-reaction problems. *J. Comput. Sci.* **21**, 38–59.
- HAACK, A., GERDING, M. & LÜBKEN, F.-J. 2014 Characteristics of stratospheric turbulent layers measured by LITOS and their relation to the Richardson number. *J. Geophys. Res.* **119** (18).
- HARVEY, W. D. 1978 Influence of free-stream disturbances on boundary-layer transition, *NASA Tech. Memorandum* 78635.
- HERBERT, T. 1988 Secondary instability of boundary layers. *Annu. Rev. Fluid Mech.* **20** (1), 487–526.
- HERBERT, T. 1997 Parabolized stability equations. *Annu. Rev. Fluid Mech.* **29** (1), 245–283.
- JIANG, L., CHOUDHARI, M., CHANG, C.-L. & LIU, C. 2006 Numerical simulations of laminar-turbulent transition in supersonic boundary layer. In *36th AIAA Fluid Dynamics Conference and Exhibit held on 05–08 June 2006 in San Francisco, California*, p. 3224. AIAA Aerospace Research Central.
- JOHNSEN, E., LARSSON, J., BHAGATWALA, A. V., CABOT, W. H., MOIN, P., OLSON, B. J., RAWAT, P. S., SHANKAR, S. K., SJÖGREEN, B., YEE, H. C. *et al.* 2010 Assessment of high-resolution methods for numerical simulations of compressible turbulence with shock waves. *J. Comput. Phys.* **229** (4), 1213–1237.
- JULIANO, T. J., ADAMCZAK, D. & KIMMEL, R. L. 2015 HiFiRE-5 flight test results. *J. Spacecr. Rockets* **52** (3), 650–663.
- KATO, H., YOSHIKAWA, A., UENO, G. & OBAYASHI, S. 2015 A data assimilation methodology for reconstructing turbulent flows around aircraft. *J. Comput. Phys.* **283**, 559–581.
- KAWAI, S. & LARSSON, J. 2012 Wall-modeling in large eddy simulation: length scales, grid resolution, and accuracy. *Phys. Fluids* **24** (1), 015105.
- KENNEDY, R. E., LAURENCE, S. J., SMITH, M. S. & MARINEAU, E. C. 2018 Investigation of the second-mode instability at Mach 14 using calibrated Schlieren. *J. Fluid Mech.* **845**.
- KIMMEL, R. L., ADAMCZAK, D. W., HARTLEY, D., ALESI, H., FROST, M. A., PIETSCH, R., SHANNON, J. & SILVESTER, T. 2018 Hypersonic international flight research experimentation-5b flight overview. *J. Spacecr. Rockets* 1–12.
- LARSSON, J. & LELE, S. K. 2009 Direct numerical simulation of canonical shock/turbulence interaction. *Phys. Fluids* **21** (12), 126101.
- LAURENCE, S. J., WAGNER, A. & HANNEMANN, K. 2016 Experimental study of second-mode instability growth and breakdown in a hypersonic boundary layer using high-speed Schlieren visualization. *J. Fluid Mech.* **797**, 471–503.
- LEYVA, I. A. 2017 The relentless pursuit of hypersonic flight. *Phys. Today* **70** (11), 30–36.
- LUCHINI, P. & BOTTARO, A. 2014 Adjoint equations in stability analysis. *Annu. Rev. Fluid Mech.* **46**, 493–517.
- MACK, L. M. 1984 Boundary-layer linear stability theory. *Tech. Rep.*, California Institute of Technology, Jet Propulsion Laboratory.
- MARXEN, O. & ZAKI, T. A. 2019 Turbulence in intermittent transitional boundary layers and in turbulence spots. *J. Fluid Mech.* **860**, 350–383.
- MAYER, C. S. J., VON TERZI, D. A. & FASEL, H. F. 2011 Direct numerical simulation of complete transition to turbulence via oblique breakdown at Mach 3. *J. Fluid Mech.* **674**, 5–42.
- MONS, V., CHASSAING, J.-C., GOMEZ, T. & SAGAUT, P. 2016 Reconstruction of unsteady viscous flows using data assimilation schemes. *J. Comput. Phys.* **316**, 255–280.
- NARASIMHA, R. 1985 The laminar-turbulent transition zone in the boundary layer. *Prog. Aerosp. Sci.* **22** (1), 29–80.
- NOCEDAL, J. & WRIGHT, S. 2006 *Numerical Optimization*. Springer.
- NOVIKOV, A. V. & EGOROV, I. 2016 Direct numerical simulations of transitional boundary layer over a flat plate in hypersonic free-stream. In *46th AIAA Fluid Dynamics Conference*, p. 3952.
- PARK, J. & ZAKI, T. A. 2019 Sensitivity of high-speed boundary-layer stability to base-flow distortion. *J. Fluid Mech.* **859**, 476–515.
- PRINGLE, C. C. T. & KERSWELL, R. R. 2010 Using nonlinear transient growth to construct the minimal seed for shear flow turbulence. *Phys. Rev. Lett.* **105** (15), 154502.
- RABIN, S. M. E., CAULFIELD, C. P. & KERSWELL, R. R. 2012 Triggering turbulence efficiently in plane Couette flow. *J. Fluid Mech.* **712**, 244–272.

- SCHNEIDER, S. P. 1999 Flight data for boundary-layer transition at hypersonic and supersonic speeds. *J. Spacecr. Rockets* **36** (1), 8–20.
- SCHNEIDER, S. P. 2015 Developing mechanism-based methods for estimating hypersonic boundary-layer transition in flight: the role of quiet tunnels. *Prog. Aerosp. Sci.* **72**, 17–29.
- SIVASUBRAMANIAN, J. & FASEL, H. F. 2015 Direct numerical simulation of transition in a sharp cone boundary layer at Mach 6: fundamental breakdown. *J. Fluid Mech.* **768**, 175–218.
- SIVASUBRAMANIAN, J. & FASEL, H. F. 2016 Direct numerical simulation of laminar-turbulent transition in a flared cone boundary layer at Mach 6. In *54th AIAA Aerospace Sciences Meeting*, p. 0846.
- STANFIELD, S. A., KIMMEL, R. L., ADAMCZAK, D. & JULIANO, T. J. 2015 Boundary-layer transition experiment during reentry of HIFiRE-1. *J. Spacecr. Rockets* **52** (3), 637–649.
- SUTHERLAND, W. 1893 LII. The viscosity of gases and molecular force. *The London, Edinburgh, Dublin Philos. Mag. J. Sci.* **36** (223), 507–531.
- SUZUKI, T. 2012 Reduced-order Kalman-filtered hybrid simulation combining particle tracking velocimetry and direct numerical simulation. *J. Fluid Mech.* **709**, 249–288.
- THUMM, A., WOLZ, W. & FASEL, H. 1990 Numerical simulation of spatially growing three-dimensional disturbance waves in compressible boundary layers. In *IUTAM Symposium on Laminar-Turbulent Transition*, pp. 303–308. Springer.
- WÄCHTER, A. & BIEGLER, L. T. 2006 On the implementation of an interior-point filter line-search algorithm for large-scale nonlinear programming. *Math. Program.* **106** (1), 25–57.
- XIAO, D. & PAPADAKIS, G. 2017 Nonlinear optimal control of bypass transition in a boundary layer flow. *Phys. Fluids* **29** (5), 054103.
- YANG, Y., ROBINSON, C., HEITZ, D. & MÉMIN, E. 2015 Enhanced ensemble-based 4DVar scheme for data assimilation. *Comput. Fluids* **115**, 201–210.
- ZAKI, T. A. & DURBIN, P. A. 2005 Mode interaction and the bypass route to transition. *J. Fluid Mech.* **531**, 85–111.
- ZAKI, T. A. & DURBIN, P. A. 2006 Continuous mode transition and the effects of pressure gradient. *J. Fluid Mech.* **563**, 357–388.
- ZHANG, C., TANG, Q. & LEE, C. 2013 Hypersonic boundary-layer transition on a flared cone. *Acta Mechanica Sin.* **29** (1), 48–54.
- ZHANG, C., ZHU, Y., CHEN, X., YUAN, H., WU, J., CHEN, S., LEE, C. & GAD-EL HAK, M. 2015 Transition in hypersonic boundary layers. *AIP Advances* **5** (10), 107137.
- ZHAO, R., WEN, C. Y., TIAN, X. D., LONG, T. H. & YUAN, W. 2018 Numerical simulation of local wall heating and cooling effect on the stability of a hypersonic boundary layer. *Intl J. Heat Mass Transfer* **121**, 986–998.
- ZHONG, X. & WANG, X. 2012 Direct numerical simulation on the receptivity, instability, and transition of hypersonic boundary layers. *Annu. Rev. Fluid Mech.* **44**, 527–561.
- ZHU, Y., CHEN, X., WU, J., CHEN, S., LEE, C. & GAD-EL HAK, M. 2018 Aerodynamic heating in transitional hypersonic boundary layers: role of second-mode instability. *Phys. Fluids* **30** (1), 011701.
- ZHU, Y., ZHANG, C., CHEN, X., YUAN, H., WU, J., CHEN, S., LEE, C. & GAD-EL HAK, M. 2016 Transition in hypersonic boundary layers: Role of dilatational waves. *AIAA J.* 3039–3049.