AFRL-AFOSR-VA-TR-2024-0235

Separation Dynamics: The View From the Wall

Tamer Zaki
JOHNS HOPKINS UNIV BALTIMORE MD
3400 NORTH CHARLES STREET
BALTIMORE, MD,
US

05/30/2024
Final Technical Report

Air Force Research Laboratory
Air Force Office of Scientific Research
Arlington, Virginia 22203
Air Force Materiel Command

REPORT DOCUMENTATION PAGE

PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ORGANIZATION.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED	
20240530	Final	START DATE 20200915	END DATE 20230914

4. TITLE AND SUBTITLE
Separation Dynamics: The View From the Wall

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. PROGRAM ELEMENT NUMBER
	FA9550-20-1-0269	61102F
5d. PROJECT NUMBER	**5e. TASK NUMBER**	**5f. WORK UNIT NUMBER**

6. AUTHOR(S)
Tamer Zaki

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
JOHNS HOPKINS UNIV BALTIMORE MD
3400 NORTH CHARLES STREET
BALTIMORE, MD
US

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Air Force Office of Scientific Research 875 N. Randolph St. Room 3112 Arlington, VA 22203	AFRL/AFOSR RTA1	AFRL-AFOSR-VA-TR-2024-0235

12. DISTRIBUTION/AVAILABILITY STATEMENT
A Distribution Unlimited: PB Public Release

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Flow separation and reattachment have a profound impact on the performance of flight vehicles, yet are probed using only a small number of discrete wall sensors. As the flow crosses the onset of separation, the spectra of the incident disturbances change significantly. As a result, the accuracy of interpreting wall-pressure data is sensitive to sensor placement, whether the sensor is positioned upstream of the separation or within the reverse-flow region. This research investigates the challenges of flow estimation within a separated high-speed flow. The impact of separation on the accuracy of flow estimation from wall measurements is first quantified in a compression ramp configuration with a six-degree ramp angle. At free-stream Mach number M = 5.59, this configuration produces a sufficiently strong compression which leads to separation upstream of the corner and downstream reattachment on the ramp. An ensemble variational (EnVar) data assimilation technique is used to perform two flow estimations: the first is conducted with sensor observations taken upstream of the separation, and the second with sensor observations taken from within the separated region. This study adopts numerical observations in lieu of experimental measurements. Whether the sensor data are extracted upstream or within separation, the non-linear optimization improves the error in the initial estimate of the boundary layer instability waves. However, to a lesser extent when the flow estimation utilizes observations from within the separated region. A comparison of the two flow estimations reveals increased errors in the disturbance spectra, as well as in instantaneous wall-pressure observations for the estimate derived from observations in the separated region. Sensor sensitivity to flow disturbances directly impacts to the efficacy of the EnVar procedure. A lack of sensitivity results in an inaccurate or inconclusive assimilation. The difference in the accuracy of the two estimations is interpreted/explained in terms of sensor sensitivity.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	c. THIS PAGE		
U	U	U	UU	22

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER *(Include area code)*
GREGG ABATE	425-1779

Standard Form 298 (Rev.5/2020)
Prescribed by ANSI Std. Z39.18

Separation dynamics: The view from the wall

Abstract

Flow separation and reattachment have a profound impact on the performance of flight vehicles, yet are probed using only a small number of discrete wall sensors. As the flow crosses the onset of separation, the spectra of the incident disturbances change significantly. As a result, the accuracy of interpreting wall-pressure data is sensitive to sensor placement, whether the sensor is positioned upstream of the separation or within the reverse-flow region. This research investigates the challenges of flow estimation within a separated high-speed flow. The impact of separation on the accuracy of flow estimation from wall measurements is first quantified in a compression ramp configuration with a six-degree ramp angle. At free-stream Mach number $M = 5.59$, this configuration produces a sufficiently strong compression which leads to separation upstream of the corner and downstream reattachment on the ramp. An ensemble variational (EnVar) data assimilation technique is used to perform two flow estimations: the first is conducted with sensor observations taken upstream of the separation, and the second with sensor observations taken from within the separated region. This study adopts numerical observations in lieu of experimental measurements. Whether the sensor data are extracted upstream or within separation, the non-linear optimization improves the error in the initial estimate of the boundary layer instability waves. However, to a lesser extent when the flow estimation utilizes observations from within the separated region. A comparison of the two flow estimations reveals increased errors in the disturbance spectra, as well as in instantaneous wall-pressure observations for the estimate derived from observations in the separated region. Sensor sensitivity to flow disturbances directly impacts to the efficacy of the EnVar procedure. A lack of sensitivity results in an inaccurate or inconclusive assimilation. The difference in the accuracy of the two estimations is interpreted/explained in terms of sensor sensitivity.

1. Introduction

The push to expand the flight regime of vehicles leads to increased mechanical and thermal loads on the aircraft structure. Efforts to guard against these limiting factors can be expensive and result in weight penalties. For these reasons, there is a pressing need to understand the flow physics around the vehicle, which requires access to high-fidelity data from experiments and simulations. The most relevant data are arguably from flight, but these are often limited to the signals of a few probes along the wall. These measurements are also susceptible to noise and are complex to interpret. In addition, often the desired quantity cannot be measured and must be computed based on measured ones. For these reasons, high-fidelity simulations are commonly adopted to complement experimental research. While simulations can provide non-intrusive access to the entire flow state, they introduce assumptions that can limit their fidelity. For example, simulations are often performed in truncated domains and must prescribe boundary conditions that may not match the true flight environment. One objective approach to ensure quantitative agreement between experiments and simulations is to infuse experimental measurements into the simulations, through the use of data assimilation algorithms. We will adopt data

assimilation to predict the state of the boundary layer from wall-pressure measurements, and assess the impact of the sensor placement. Specifically, we are concerned with the impact of separation on the accuracy of the state estimation.

Aircraft geometries often exhibit outer mold lines that change abruptly with respect to the oncoming flow. These sudden changes in geometries can result in appreciable adverse pressure gradients, including ones due to compression waves in the supersonic regime. If the pressure gradient is significant, it can separate the flow. This phenomenon is commonly observed in various cases, such as flow over a compression ramp or cone flare, the reflection of a shock from a secondary body, or flow around a blunt fin (Clemens & Narayanaswamy 2014). The region characterized by flow separation can significantly influence the spectral characteristics of boundary-layer disturbances. For example, Butler & Laurence (2021) conducted an experimental investigation that demonstrated the impact of separation on Mack's second mode. The authors found that, at separation, the waves move away from the wall through the detached shear layer, extending over top of the separated region. This behavior leads to a damped second mode when observed by wall-pressure measurements within the separated region. Experiments that probed compression ramps (McClure 1992) and cone-flare configurations (Butler & Laurence 2021) have also shown the manifestation of low-frequency oscillations at the compression-wave foot, along the streamwise direction. The driving mechanism of oscillations in the separated region has significant implications for the accuracy of flow estimation with data assimilation, particularly when measurements are taken along the wall. For successful flow reconstruction, it is essential that the sensors within the separation region maintain sensitivity to the upstream spectra. Without such sensitivity, accurate and reliable flow estimation becomes unattainable. The objective of this work is therefore to assess the impact of flow separation on data assimilation when observations are taken in the separated region. This impact will be quantified by comparing flow reconstructions of an independent simulation of flow over a compression ramp, using sensors placed in the separated region and sensors placed upstream of separation. The errors in the estimated flow will be explained in terms of sensitivity using two different approaches.

The choice of the data assimilation approach depends on the problem of interest. Our group has developed a wide suite of data-assimilation algorithms for a wide range of fluid applications and observations. When there is an abundance of measurements, we can adopt direct forcing methods (Wang & Zaki 2022) and successfully synchronize a simulation to an experiment. In practice, however, the measurements are limited. In such scenario, and for short optimization horizons, adjoint-variational approaches are most efficient and accurate (Wang *et al.* 2019*b*; Wang 2021; Wang & Zaki 2021). This approach requires a dual, or adjoint, model which is not always available, and incur a significant storage cost since the time-dependent forward flow field appears in the adjoint equations. In addition, maintaining accuracy over long observation windows is challenging. For long observation horizons, ensemble-variational strategies can be effective (Mons *et al.* 2019, 2021; Buchta & Zaki 2021; Buchta *et al.* 2022), and have the benefit of being non-intrusive, i.e. they can be adopted with any forward model, and are applicable to long observation horizons. Finally, we also note that machine-learning techniques are gaining popularity (Du *et al.* 2023; Du & Zaki 2021; Mao *et al.* 2021; Cai *et al.* 2021; Di Leoni *et al.* 2023; Hao *et al.* 2023), but they do not strongly enforce the governing equations and hence are not ideally suited for fundamental studies similar to our current interest. In light of these consideration, we focus on ensemble-variational data assimilation (EnVar) which strongly satisfies the Navier-Stokes equations, and which we will be adopt in conjunction with direct numerical simulations in this effort in order to ensure accuracy of our predictions.

In section 2, we summarize the flow configuration that will be adopted to study the

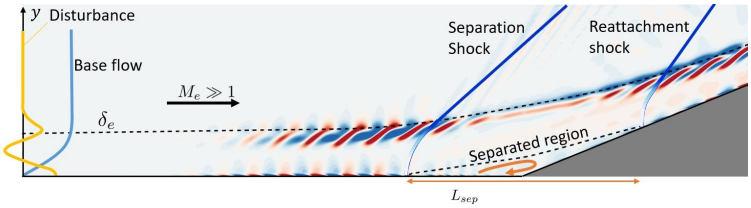

Figure 1: Schematic of flow separation on a compression ramp. The boundary/shear-layer edge δ_e is marked by the dashed line, as well as the separation region with length L_{sep}. The inflow condition is a superposition of a laminar boundary layer and instability waves.

impact of separation on the interpretation of sensors. We then introduce the ensemble-variational data-assimilation procedure (§3), and report on estimation of the flow from wall-pressure sensors upstream of and within the separation region (§4). In section 5, we provide an interpretation of the results by introducing the notion of sensor sensitivity, and we evaluate this quantity for the two classes of sensors upstream and within the separation zone. Conclusions are presented in the final section.

2. Flow configuration

For this investigation, we consider supersonic flow over a two-dimensional compression corner, which is shown schematically in figure 1. As the flow approaches the ramp, the required change in flow direction establishes a compression wave or a shock, with an associated pressure gradient. At relatively small ramp angles, the adverse pressure gradient is not sufficient to initiate flow separation, positioning the compression wave at the corner. However, as the ramp angle increases, a separation bubble emerges and the compression wave is displaced upstream from the corner, as shown in the schematic. Flow disturbances in the upstream flat-plate boundary layer amplify or decay depending on the stability characteristic of the attached flow. Upon reaching the compression wave, some of the disturbances waves detach from the boundary layer and advect along the separated shear layer, while other frequencies amplify within the separation zone (Butler 2021; Butler & Laurence 2021). We will consider measurements that are recorded by sensors situated on the horizontal wall. To quantify the influence of separation on the flow estimation, two assimilation studies are conducted. The first adopts sensors positioned upstream of the separation, and the second utilizes sensors located within the separated region.

The flow state satisfies the compressible Navier-Stokes equations. Dimensional variables are marked by superscript $(^*)$, and the reference scales adopted for non-dimensionalization are the the length $L^* = 1m$ and boundary-layer edge quantities,

$$\begin{aligned} x_i &= \frac{x_i^*}{L^*}, \qquad t = \frac{t^*}{L^*/c_e^*}, \qquad u_i = \frac{u_i^*}{c_e^*}, \qquad \rho = \frac{\rho^*}{\rho_e^* c_e^*}, \qquad p = \frac{p_i^*}{\rho_e^*(c_e^*)^2} \\ T &= \frac{T^*}{(\gamma-1)T^*} = \frac{T^*}{(c_e^*)^2/C_{p,e}^*}, \qquad \mu = \frac{\mu^*}{\mu_e^*}, \qquad E = \frac{E^*}{\rho_e^*(c_e^*)^2}, \qquad \tau_{ij} = \frac{\tau_{ij}^*}{\mu_e^* c_e^*/L^*} \end{aligned} \tag{2.1}$$

where ρ^* is the density, u_i^* is the velocity vector, p^* is pressure, $E^* = p^*/(\gamma-1)+0.5\rho^* u_i^* u_i^*$ is the total energy, $T^* = p^*/(\rho^* R^*)$ is the temperature with gas constant R^* and ratio of specific heats γ, and c_e^* is the speed of sound. The choice of the edge conditions is motivated by recent experiments by Butler (2021), and the reference values are reported in Table 1. In the non-dimensional form, the Navier-Stokes equations are:

Parameter	Units	Value
T_e^*	K	123
P_e^*	Pa	1321
ρ_e^*	kg/m^3	3.74×10^{-2}
c_e^*	m/s	222
U_e^*	m/s	1242
M_e	-	5.59
μ_e^*	kg/m s	8.41×10^{-6}
Re_e/L^*	m^{-1}	9.89×10^5
Pr	-	0.72

Table 1: Boundary-layer edge conditions.

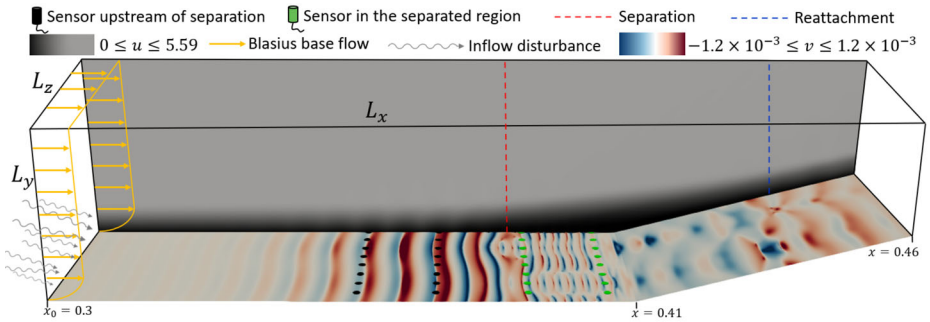

Figure 2: Flow configuration of separated high-speed flow over a ramp. The top of the figure labels all the different features of the flow field. The length of the computational domain is labeled along the surface of the compression ramp in Cartesian x-coordinates.

$$\frac{\partial \rho}{\partial t} + \frac{\partial (\rho u_i)}{\partial x_i} = 0, \tag{2.2}$$

$$\frac{\partial (\rho u_i)}{\partial t} + \frac{\partial}{\partial x_j}(\rho u_i u_j) = -\frac{\partial p}{\partial x_i} + \frac{\partial \tau_{ij}}{\partial x_j}, \tag{2.3}$$

$$\frac{\partial E}{\partial t} + \frac{\partial}{\partial x_i}[u_i(E+p)] = \frac{\partial}{\partial x_i}(\tau_{ij} u_j) + \frac{\partial q_i}{\partial x_i}. \tag{2.4}$$

The non-dimensional stress tensor and heat flux are,

$$\tau_{ij} = \frac{\mu}{Re_e}\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i}\right) + \frac{\lambda}{Re_e}\delta_{ij}\frac{\partial u_k}{\partial x_k}, \qquad q_i = -\frac{\mu}{Re_e Pr}\frac{\partial T}{\partial x_i}, \tag{2.5}$$

where the Reynolds number is $Re_e = c_e^* \rho_e^* L^* / \mu_e^*$, the Prandtl number is $Pr = \mu^* C_{p,e}^* / \kappa^*$, and κ^* is the thermal conductivity. Lastly, the non-dimensional total energy is $E = p/(\gamma - 1) + 0.5\rho u_i u_i$, non-dimensional temperature is $T = P/(\rho R)$ where $R = (\gamma - 1)/\gamma$, and non-dimensional viscosity is $\mu = [(\gamma - 1)T]^n$. The state vector, $\boldsymbol{q} = [\rho, \rho u, \rho v, \rho w, E]^\top$, can be expressed as a the superposition $\boldsymbol{q} = \bar{\boldsymbol{q}} + \boldsymbol{q}'$, where $\bar{\boldsymbol{q}}$ denotes the mean flow and $\boldsymbol{q}'$ is the perturbation field.

The computational setup is illustrated in figure 2. The domain sizes are $L_x = 16 \times 10^{-2}$ in the streamwise, $L_y = 4.43 \times 10^{-2}$ in the vertical, and $L_z = 1.44 \times 10^{-2}$ in the spanwise directions. With a ramp angle of $\phi = 6°$, the resulting laminar separation length is $L_{sep} = 11.7\delta_e$ where δ_e is the inflow boundary layer thickness. This classifies the separation as strong, as defined by Clemens & Narayanaswamy (2014). The ramp angle was chosen to ensure an adequate separation length. This configuration allows the

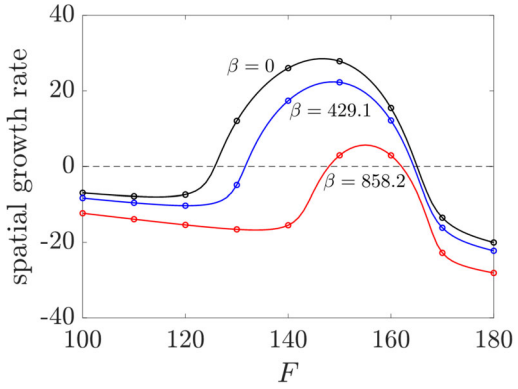

Figure 3: Inflow spatial growth rates versus normalized frequency for spanwise wavenumbers $\beta = \{0,\ 429.1,\ 858.2\}$.

placement of two sensor rows upstream of the separation point but sufficiently downstream of the inflow such that non-linear effects are relevant, and two sensor rows downstream of the laminar separation point and prior to the start of the ramp. Isothermal, no-slip boundary conditions are imposed at the bottom wall, with $T_w^* = 300K$ which matches Butler (2021). The inflow boundary condition is a superposition of the base state and a disturbance field at the inflow $x = x_0$,

$$\boldsymbol{q}_0(x_0, y, z, t) = \boldsymbol{q}_B(x_0, y, z) + \boldsymbol{q}'(x_0, y, z, t) \tag{2.6}$$

where the base state $\boldsymbol{q}_B$ is the compressible Blasius solution, based on the edge conditions in table 1. The inflow disturbance $\boldsymbol{q}'$ is defined a superposition of instability waves

$$\boldsymbol{q}'(x_0, y, z, t) = Re\left(\sum_m \sum_f c_{f,m} \breve{\boldsymbol{q}}_{f,m}(y) \exp\left[i\beta_m z - i\omega_f t + i\theta_{f,m}\right]\right), \tag{2.7}$$

which includes a range slow second modes (Mack 1984; Fedorov 2011; Park & Zaki 2019) with amplitudes $c_{f,m}$, phases $\theta_{f,m}$, frequencies ω_f, spanwise wavenumbers β_m, and mode shapes $\breve{\boldsymbol{q}}_{f,m}(y) = [\breve{\rho}, \breve{u}, \breve{v}, \breve{w}, \breve{T}]_{f,m}(y)$.

Linear stability analysis was performed on the Blasius base state, for a range of non-dimensional frequencies,

$$F\left(\equiv \frac{\omega_f L_B}{M Re_{L_B}} \times 10^6\right) = \{100, 110, \ldots, 180\}, \tag{2.8}$$

and spanwise wavenumbers specified as $\beta = \{0,\ 429.1,\ 858.2\}$, where the Blasius length is $L_B^* = \sqrt{\mu_e^* x_0^* / \rho_e^* U_e^*}$ and the Reynolds number at the inflow is $Re_{L_B} = \rho_e^* U_e^* L_B^* / \mu_e^* = 1287.3$. The results of the linear stability analysis are shown in figure 3, which motivated the choice of the frequencies and wavenumbers of interest. The most unstable mode at each (F, β) pair was adopted as a member of the superposition of inflow disturbances. The eigenfunctions were normalized to unit energy, prior to specifying their inflow amplitudes.

The computational grid was comprised of $[N_x, N_y, N_z] = [1641, 201, 60]$ grid points. Uniform spacing was utilized in both the streamwise and spanwise directions, while hyperbolic tangent stretching was applied in the wall-normal direction. The stretching in the wall-normal direction strategically positions points near the inflow boundary-layer edge, ensuring the accuracy of the high gradients present in the disturbances in this region. The near-wall grid spacing is given by $(\Delta x_w, \Delta y_w, \Delta z_w) = (9.76 \times 10^{-5}, 8.94 \times 10^{-6}, 2.44 \times 10^{-4})$.

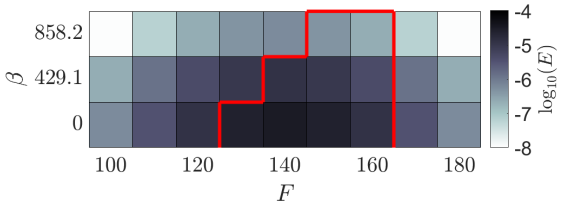

Figure 4: Inflow energy spectra of the hidden truth. Modes that are linearly unstable at the inflow are circumscribed by the red lines.

2.1. *The hidden truth: Reference simulation and observations*

Data assimilation requires measurements which can be obtained from experiments or flight tests. However, such data would involve uncertainties that can not be fully eliminated, and hence it becomes difficult to distinguish their impact from the influence of sensor placement outside or within separation. To circumvent this limitation, we utilize synthetic observations, from an independent simulation. This methodology enables exact error quantification, thus ensuring a precise evaluation of the accuracy of the flow estimation. Even when using synthetic measurements from a simulation, the observations are treated analogously to those obtained from experiments or flight tests. The "truth" is concealed until the conclusion of the data assimilation. Only after the assimilation process is concluded are the predicted results compared to the hidden truth.

The reference simulation for the generation of observations satisfied the compressible Navier-Stokes equations, and are performed in the flow configuration in figure 2. The inflow condition was comprised of a superposition of twenty-seven modes: nine-planar and eighteen-oblique instability waves ($F = \{100, 110, \ldots, 180\}$ with $\Delta F = 10$ and $\beta = \{0, 429.1, 858.2\}$. The inflow modal energy distribution of the hidden truth is shown in figure 4 with the unstable modes outlined in red.

During the data assimilation, the primary objective is to discover, or estimate, the upstream disturbance spectra (figure 4) that reproduces the available observations. These observations are sourced from the reference simulation using an array of 32 probes, akin to PCB sensors, arranged in a $(N_x \times N_z) = (4 \times 8)$ grid configuration, as illustrated in figure 2. The sensors are placed at four streamwise positions: $\{x_{s1}, x_{s2}, x_{s3}, x_{s4}\} = \{0.358, 0.374, 0.389, 0.405\}$ and uniformly distributed in the span, starting at $z_{s1} = 0$ with an interval of $\Delta z_s = 1.83 \times 10^{-3}$. As noted earlier, two of the four streamwise positions are upstream of the laminar separation point, and the subsequent two rows are positioned within separation. This arrangement facilitates a comparative analysis of flow estimations using observations from either the upstream sensors or those within the separated region.

The simulations were initialized with a laminar base flow. After two flow-through times were complete, determined by the non-dimensional edge velocity U_e and streamwise domain length L_x, observations were collected. The observation window, $T_a = 2.04 \times 10^{-2}$, corresponds to a single period of the minimum inflow frequency $F = 10$, and the measurement resolution corresponds to a Nyquist frequency of $F = 2.275 \times 10^4$. In the spanwise direction, the longest resolved wavelength has a wavenumber of $\beta = 429.1$; the sensor spacing yields the largest spanwise wavenumber $\beta = 1.717 \times 10^3$ resolved by the measurements. Figure 5 displays select instantaneous wall-pressure observations from sensors located both upstream and downstream of separation. From sensor x_{s1} to x_{s2}, the signal intensifies due to the amplification of the second unstable modes within the boundary layer. In the separated region, the wall-pressure signal attenuates as the

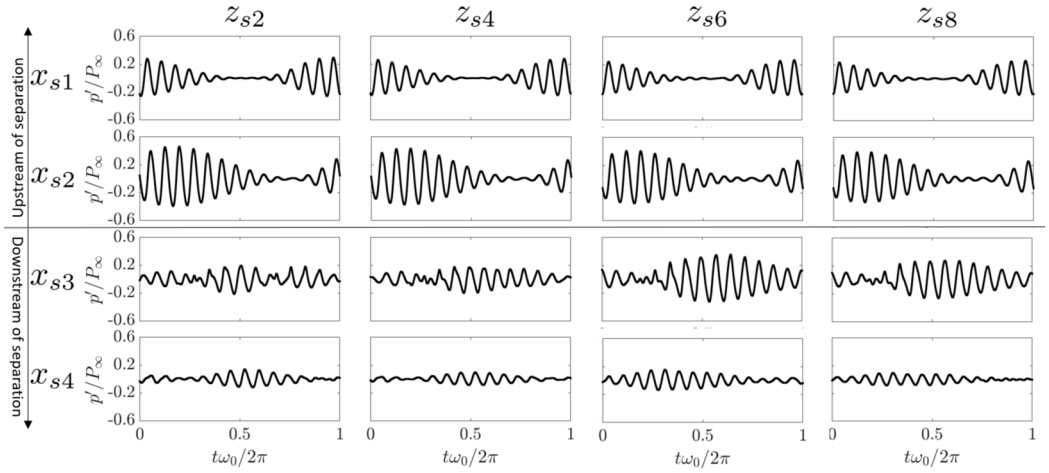

Figure 5: Synthetic measurements from the reference simulation, taken from spanwise sensor locations $\{z_{s2},\, z_{s4},\, z_{s6},\, z_{s8}\}$ upstream of separation $\{x_{s1},\, x_{s2}\}$, and downstream of separation, $\{x_{s3},\, x_{s4}\}$. Time is normalized by the fundamental frequency ω_0 corresponding to $F = 10$.

disturbance separates from the boundary layer and propagates through the shear layer. Observations within this region also reveal increased spanwise variation.

3. Ensemble-variational data assimilation (EnVar)

Data assimilation relies on the availability of measurements $\boldsymbol{m}$, which in the present case are temporal wall-pressure fluctuations at the sensor locations. The objective of the assimilation is to predict the optimal value of the unknown parameters $\boldsymbol{c}$, such that the flow estimation aligns with the measurements. The control parameters $\boldsymbol{c}$ in our case are the amplitudes of the inflow boundary-layer instability waves, and the flow model is direct simulations of the compressible Navier-Stokes equations which is expressed compactly in operator form as $\boldsymbol{q} = \mathcal{N}(\boldsymbol{q}_0)$. The cost functional for the optimization is,

$$\mathcal{J} = \frac{1}{2} \left\| \boldsymbol{m} - \boldsymbol{h} \right\|^2_{\Sigma_{\boldsymbol{m}}^{-1}} + \frac{1}{2} \left\| \boldsymbol{c} - \boldsymbol{c}_i \right\|^2_{\Sigma_{\boldsymbol{c}}^{-1}} . \tag{3.1}$$

The first term in the cost is the difference between the available observations $\boldsymbol{m}$ and their estimate in the assimilation $\boldsymbol{h} = \mathcal{M}(\boldsymbol{q}) = \mathcal{M}(\mathcal{N}(\boldsymbol{q}_0))$ where $\mathcal{M}$ is the measurement kernel that extracts the wall-pressure observations from the flow state $\mathcal{M}(\boldsymbol{q}) = \int_{\mathcal{T}} \int_{\mathcal{V}} p\delta(\boldsymbol{x} - \boldsymbol{x}_s, t - t_m) d\mathcal{V} dt$. The second term is a regularization that limits the update to the control vector from its prior estimate $\boldsymbol{c}_i$. The two terms are normalized by their respective covariance matrices, $\Sigma_{\boldsymbol{m}}$ and $\Sigma_{\boldsymbol{c}}$, providing appropriate weighting.

Minimizing the cost function is achieved using a gradient descent method. Various strategies can be used to compute the gradient of the cost function, $\nabla\mathcal{J}$. While adjoint methods provide precise gradients and boast high efficiency, they can face difficulties in maintaining forward-adjoint duality over the long observation horizons, and they also demand significant data storage for each iteration (Wang *et al.* 2019*b*). We adopt the EnVar method, wherein the local gradient is extracted from an ensemble of control vectors $\boldsymbol{c}_i^{(j)}$ centered around the mean vector $\boldsymbol{c}_i$ (Buchta & Zaki 2021; Buchta *et al.* 2022). The EnVar approach updates the mean control vector $\boldsymbol{c}_i$ with a weighted superposition

$$\boldsymbol{c}_{i+1} = \boldsymbol{c}_i + \boldsymbol{P}\boldsymbol{w}, \tag{3.2}$$

where the perturbation matrix is $\boldsymbol{P} = \left[\boldsymbol{c}_i^{(1)} - \boldsymbol{c}_i | \, \boldsymbol{c}_i^{(2)} - \boldsymbol{c}_i | \, \ldots | \, \boldsymbol{c}_i^{(N_e)} - \boldsymbol{c}_i\right]$ and $\boldsymbol{w}$ are the weights associated with each ensemble control vector, dictating the direction of the steepest descent. In the form (3.2), $\boldsymbol{w}$ can be considered the new control vector.

Considering the ensemble members are small perturbations around the mean, the measurements $\boldsymbol{h}$ can be approximated as,

$$\boldsymbol{h} \approx \boldsymbol{h}_i + \boldsymbol{H}\boldsymbol{w}, \tag{3.3}$$

where the observation matrix is $\boldsymbol{H} = \left[\boldsymbol{h}_i^{(1)} - \boldsymbol{h}_i | \, \boldsymbol{h}_i^{(2)} - \boldsymbol{h}_i | \, \ldots | \, \boldsymbol{h}_i^{(N_e)} - \boldsymbol{h}_i\right]$. The updated approximate cost function then becomes,

$$\tilde{\mathcal{J}} = \frac{1}{2} \left\| \boldsymbol{m} - \boldsymbol{h}_i - \boldsymbol{H}\boldsymbol{w} \right\|^2_{\Sigma_{\boldsymbol{m}}^{-1}} + \frac{1}{2} \left\| \boldsymbol{P}\boldsymbol{w} \right\|^2_{\Sigma_{\boldsymbol{c}}^{-1}} . \tag{3.4}$$

A Newton step can then be performed to optimize $\tilde{\mathcal{J}}$ by setting the gradient of the cost function with respect to the weights,

$$\boldsymbol{g} = \frac{\partial \tilde{\mathcal{J}}}{\partial \boldsymbol{w}} = \boldsymbol{H}^\top \Sigma_{\boldsymbol{m}}^{-1} \left(\boldsymbol{h}_i + \boldsymbol{H}\boldsymbol{w} - \boldsymbol{m} \right) + \boldsymbol{P}^\top \Sigma_{\boldsymbol{c}}^{-1} \boldsymbol{P}\boldsymbol{w}, \tag{3.5}$$

equal to zero, and then solving for the optimal weights,

$$\boldsymbol{w} = \underbrace{- \left(\boldsymbol{H}^\top \Sigma_{\boldsymbol{m}}^{-1} \boldsymbol{H} + \boldsymbol{P}^\top \Sigma_{\boldsymbol{c}}^{-1} \boldsymbol{P} \right)^{-1}}_{-\mathcal{H}^{-1}} \underbrace{\left(\boldsymbol{H}^\top \Sigma_{\boldsymbol{m}}^{-1} \left(\boldsymbol{h}_i - \boldsymbol{m} \right) \right)}_{\boldsymbol{g}|_{w=0}} . \tag{3.6}$$

In the expression (3.6), the first term on the right-hand side represents the inverse of the Hessian $\mathcal{H}$ of the cost function, or its curvature. The second term, $g|_{\boldsymbol{w}=0} = \nabla \tilde{\mathcal{J}}(\boldsymbol{w} = 0)$, is the gradient at the current control vector location. Once computed, the optimal weights establish the search direction and are employed to update the control vector. To optimize the step size, the update step equation (3.7) is adapted to incorporate a parameter α,

$$\boldsymbol{c}_{i+1} = \boldsymbol{c}_i + \alpha \boldsymbol{P}\boldsymbol{w} \tag{3.7}$$

We applied Jarratt method (Jarratt 1967; Brent 2013) to determine α at each iteration. After each iteration, the ensemble members and covariance matrix are updated according to,

$$\boldsymbol{P}_{i+1} = \sqrt{N-1} \boldsymbol{P}_i \mathcal{H}^{-1/2} \boldsymbol{U}, \tag{3.8}$$

$$\Sigma_{\boldsymbol{c}}^{i+1} = \frac{1}{N-1} \boldsymbol{P}_{i+1} \boldsymbol{P}_{i+1}^\top. \tag{3.9}$$

As the iterations progress and the cost decreases, the confidence in the control vector increases and hence the covariance matrix $\Sigma_{\boldsymbol{c}}$ reduces. Excessive reduction can over-emphasize the prior term in the cost function. To counteract this effect, Anderson & Anderson (1999) advocate for the inflation of the covariance matrix. This is implemented using an inflation coefficient A, which results in the updated perturbation matrix $\boldsymbol{P}_{i+1} = \sqrt{A}\boldsymbol{P}_i$ and covariance matrix $\Sigma_c^{i+1} = A/(N_e - 1)\boldsymbol{P}_i \boldsymbol{P}_i^\top$. Practically, this adjustment is invoked when the cost stagnates before meeting the convergence criteria.

3.1. *Initial estimate of the control vector*

The EnVar method procedure starts from an initial control vector, denoted as $\boldsymbol{c}_0$. A well formulated initial guess can significantly bolster both the accuracy and computational efficiency of the non-linear optimization. This is particularly true for scenarios with complex cost functions that may be challenging to navigate (Evensen *et al.* 2009). To

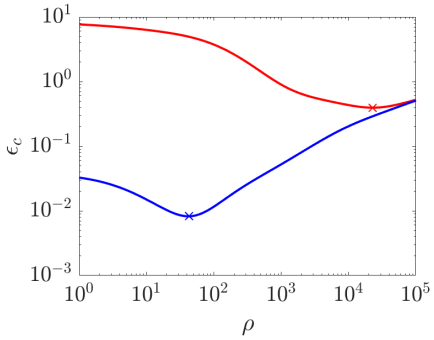

Figure 6: Error in the initial estimate ϵ_c versus penalty term ρ. The blue and red lines represent the initial estimates from observations upstream of separation and within the separated region, respectively. An (x) marks the optimum$-\rho$ value on both curves.

guide our initial estimate, we adopt a physics-based approach, specifically using the linearized Navier-Stokes equations. Assuming linearized dynamics, the cost function becomes,

$$\mathcal{J}_l = \frac{1}{2}\left\|\boldsymbol{m} - \boldsymbol{L}\boldsymbol{v}\right\|^2 + \frac{\rho}{2}\left\|\boldsymbol{v}\right\|^2. \tag{3.10}$$

In this expression, $\boldsymbol{L}\boldsymbol{v}$ is the linear estimate of the measurements. Specifically, we perform a linearized Navier-Stokes solution $\boldsymbol{q} = \mathcal{L}(\boldsymbol{c})$, and collect model observations that we Fourier decompose in order to isolate the various components of the wall pressure, $\hat{p}_{(f,m)}$. The measurements are then approximated using the superposition where $\boldsymbol{L} = \left[\ldots \,|\, Re\left(\hat{p}_{(f,m)}\right) \,|\, Im\left(\hat{p}_{(f,m)}\right) \,|\, \ldots\right]$ is the observation matrix and $\boldsymbol{v} = \left[\ldots, v_{f,m},\ w_{f,m}, \ldots\right]^\top$ is a vector of the optimal Fourier coefficients. In the language of EnVar, we adopt an ensemble of $(f \times m)$ members each targeting one of the search directions of the control vector. The error committed by adopting a linear model, relative to a non-linear simulation, increases as a function of downstream distance as disturbances amplify. Therefore, we restrict the measurements adopted in (3.10) to the most upstream sensor location. For this reason, the initial estimates for the sensors upstream and downstream of separation were evaluated from the sensors at $x_{s1} = 0.358$ and $x_{s3} = 0.389$, respectively (without x_{s2} in the former and x_{s4} in the later case). Newton method was once again employed to optimize equation (3.10), yielding the optimal modal coefficients

$$\boldsymbol{v} = -(\boldsymbol{L}^\top\boldsymbol{L} + \rho\boldsymbol{I})^{-1}(\boldsymbol{L}^\top\boldsymbol{m}_l), \tag{3.11}$$

with the optimal inflow modal amplitudes given by $\boldsymbol{c}_{f,m} = \sqrt{v_{f,m}^2 + w_{f,m}^2}$.

The cost $\mathcal{J}_l$ in (3.10) also incorporates regulation term with coefficient, ρ, which serves to mitigate over-fitting of the Fourier components, effectively constraining the total inflow energy. The most optimal value for the penalty term ρ was chosen to minimize the error in the inflow modal energy of the initial estimates compared to the hidden truth. This choice, although not possible in practice, provides the best possible initial estimate for the reconstruction. Figure 6 shows the error in the initial estimate, $\epsilon_c = \|\boldsymbol{c}_{\text{truth}} - \boldsymbol{c}\|/\|\boldsymbol{c}_{\text{truth}}\|$, versus the penalty value ρ, with optimal values denoted by an (x). The figure indicates that initial estimates from observations in the separated region consistently have higher errors than those generated from upstream observations. This points to diminished sensor sensitivity to upstream spectra in the separated region, generating an inaccurate gradient calculated by linear EnVar.

The accuracy of reproducing the measurement is shown in figure 7, which compares the

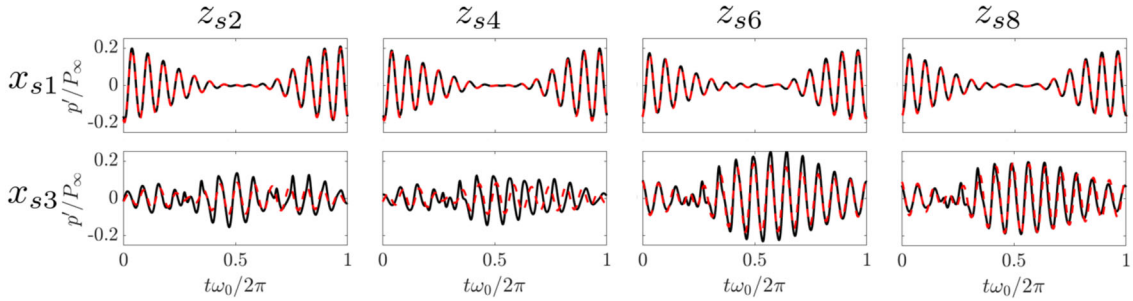

Figure 7: Linear reconstruction ($\boldsymbol{Lv}$) at sensor locations x_{s1} (upstream of separation) and x_{s3} (within separation) for spanwise sensors $\{z_{s2},\, z_{s4},\, z_{s6},\, z_{s8}\}$ in dashed-red with model observations ($\boldsymbol{m}$) in black.

linear reconstruction $\boldsymbol{Lv}$ with the observations $\boldsymbol{m}$ at the sensor locations x_{s1} upstream of separation and x_{s3} within the separated region. The linear reconstruction at x_{s1} exhibits minimal error compared to the reference observations. In contrast, the reconstruction at x_{s3} displays increased error, highlighting a lack of sensor sensitivity in this region.

Having formed the initial estimate, an ensemble is generated around this mean-control vector (see Buchta & Zaki 2021, for details). All elements are then in place to perform the nonlinear EnVar optimization, which pursues the inflow condition that most precisely reproduces observations from the reference simulation and, in doing so, estimates the complete time-dependent flow field.

4. Results

4.1. *Flow estimation using sensors upstream of separation*

We start by evaluating the accuracy of the flow estimation using sensors positioned upstream of the separation against the hidden truth. Figure 8 shows the reduction in the cost during the EnVar iteration, where $\mathcal{J}$ is normalized by the norm of the observations $\mathcal{J}_r = 0.5\,\|\boldsymbol{m}\|^2_{\Sigma_{\boldsymbol{m}}^{-1}}$. As noted in §3.1, the initial linear estimate of $\boldsymbol{c}$ was able to reproduce the wall pressure signal at the sensors using a linear model. However, even with such an accurate initial estimate, the non-linear optimization further reduced the cost function by nearly an order of magnitude.

Figure 9 displays the instantaneous, time-dependent wall-pressure observations compared to the observation from the reference simulation for the initial and optimal control vector. It is noteworthy that the differences between the observations from the initial and the optimal control vectors are indistinguishable, and nearly identical to the reference observations.

A comparison between the reference and estimated RMS pressure perturbations is presented in figure 10*a*, which also includes the associated error (figure 10b). The RMS pressure signal is accurately estimated throughout the domain. The error at the inflow is on the order of 10^{-3} and decreases by an order of magnitude near sensor location x_{s2}. The errors amplify in the separated region, and reach approximately an order of magnitude greater than that just upstream of separation. The decay and subsequent amplification of errors hints to a difficulty in estimating the state at the inflow from measurements either upstream, or downstream of separation alone. In the former case, reducing errors in estimating the measurements at the sensor may not guarantee accurate prediction downstream. In the latter case, the error pattern suggests that the downstream sensors, within separation, may lack sensitivity to the inflow.

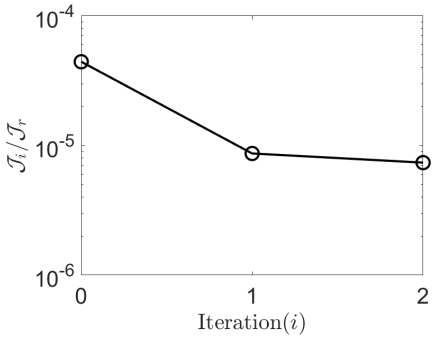

Figure 8: Cost reduction in flow estimation using sensors upstream of separation.

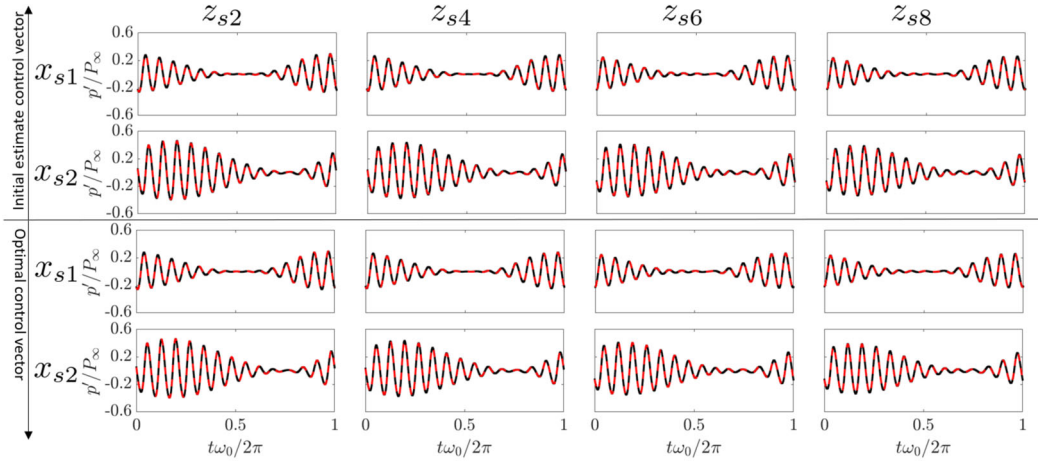

Figure 9: Synthetic observations (black line) and their estimate from the optimal control vector (red-dashed line), assimilated using data from sensors x_{s1} and x_{s2}.

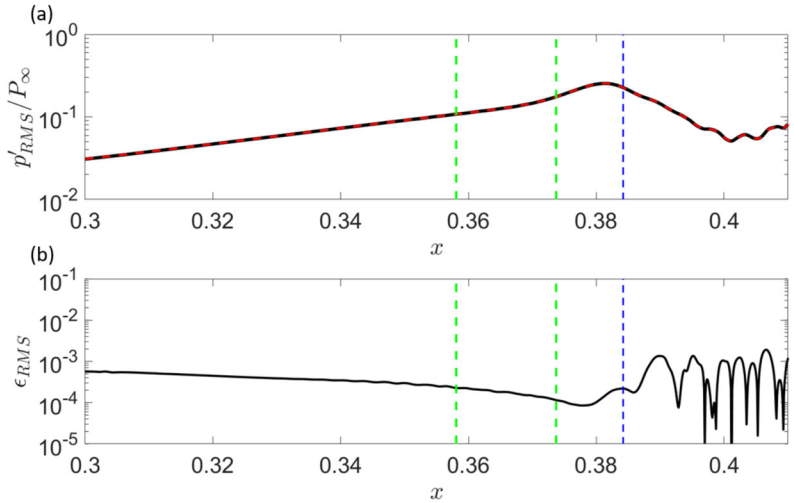

Figure 10: (a) RMS pressure perturbation and (b) their errors in the estimate by the optimal control vector. Flow estimation is performed using sensors upstream of separation. Black line depicts the truth; red-dashed lines denote the estimated state. Separation is indicated by a blue-dashed line, and sensor locations are marked with green-dashed lines.

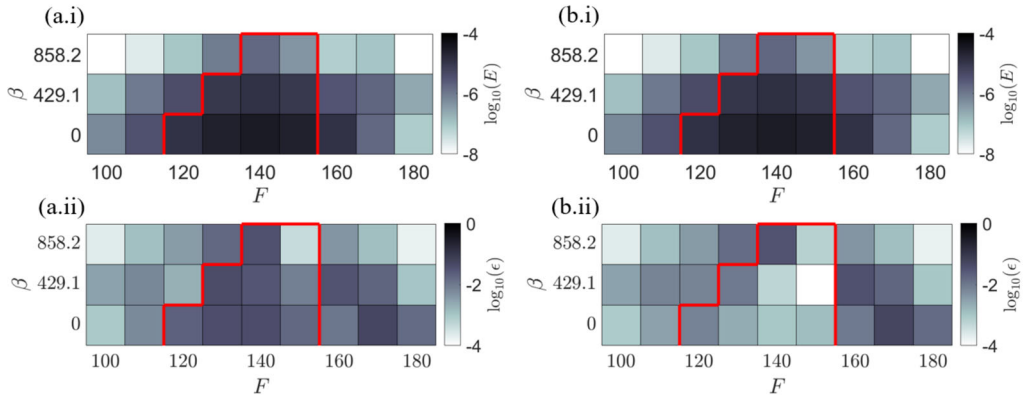

Figure 11: Modal energy and error in the estimation from sensors positioned upstream of separation: (a) Initial estimate; (b) Final mean-control vector; (i) inflow spectra; (ii) error relative to the hidden truth. The most amplified modes according to the N-factor within the computational domain are circumscribed by the red line.

In figure 11, we report the control vector, which is the inflow energy spectra E. Both the initial linear guess and the converged estimate are plotted, as well as their errors relative to the truth,

$$\epsilon(F,\beta) = \frac{|E(F,\beta) - E_t(F,\beta)|}{E_t(140,0)}. \tag{4.1}$$

True values are denoted with subscript t, and the normalization by mode $(140, 0)$ is adopted because that mode has the largest N-factor for a laminar base state. Modes near the upper boundary of the neutral curve, such as $F = 160$, very quickly become stable within the domain. In contrast, modes near the lower neutral curve, specifically $\{(120, 0), (130, 429.1), (140, 858.2)\}$, are unstable within the domain, amplify appreciably and are interpretable from the sensor measurements. A visual comparison of the converged estimate with the initial guess does not immediately show significant differences. Nonetheless, the errors show the improvement of roughly an order of magnitude, across a majority of the most amplified modes. The total energy predicted by the final estimation is $E = 1.59 \times 10^{-4}$, which is within 7.6% from the true value.

The amplitudes of the stable inflow modes appear unchanged by the optimization process. Within the domain, the wall-pressure signals of these modes decay with downstream distance. Due to the separation of the inflow plane and the sensor locations, these modes are not detected.

The most amplified planar and oblique waves at the different spanwise wavenumbers are reported in figure 12. The results compare the reference state, or hidden truth, to the assimilated state. There is good agreement for the planar mode and for the oblique wave with $\beta = 429.1$. Notable discrepancies, however, arise upstream of the sensor locations in the oblique mode with $\beta = 858.2$. The early errors in $(140, 858.2)$, while visible in the figure, are insignificant with respect to the total inflow energy. In addition, as this mode amplifies and reaches a relevant amplitude near separation, the initial errors while still present become insignificant relative to the attained $\hat{p}\hat{p}^*$. This result underscores that the amplification of this mode as it approaches separation is not linear, or else the errors would have grown in proportion to the signal. This interpretation is consistent with separation being the key driver of the dynamics in this region.

The results demonstrate that the EnVar assimilation was able to improve the initial estimate of the upstream disturbance spectra within two iterations. The amplification of

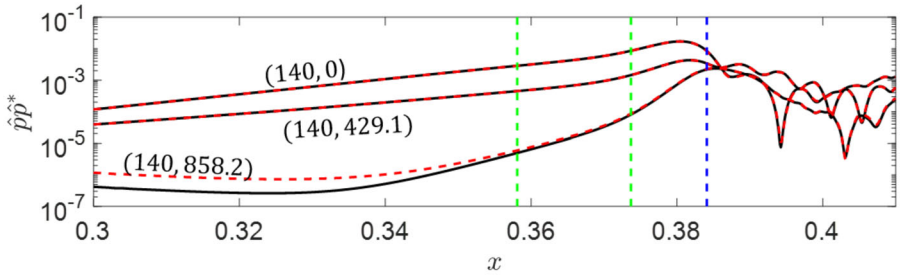

Figure 12: Comparison of the wall-pressure spectra from the (solid) reference simulation and (dashed) the assimilated flow, at $F = 140$. Separation is indicated by a blue-dashed line, and sensor locations are marked by green-dashed lines.

originally stable modes, specifically $(140, 858.2)$, within the separated region may motivate use of downstream sensors, although connecting these observations to the inflow, through the region of appreciable modal decay, may be challenging.

4.2. *Flow estimation using sensors within the separated region*

We now consider results from the assimilation of sensor data from within the separated region, and compare the accuracy of predictions to the true flow. Figure 13 reports the convergence of the cost function normalized by the norm of the measurements. Because the initial estimate is farther from the observations than for upstream sensors, the associated normalized cost is higher and is on the order of 10^{-2}. Through five iterations of the EnVar algorithm, this cost is reduced by an order of magnitude, but remains larger than the case of upstream sensors.

Figure 14 shows the improvement in matching the measurements at four of the eight spanwise sensors located at x_{s3} and x_{s4}. The large errors in the initial estimate (top) are appreciably reduced for the optimal control vector (bottom) at the fifth iteration. Despite the improvement, there are still discernible differences between the model and reference observations, specifically near the peaks of the signals.

Figure 15 shows the RMS pressure perturbation for all the EnVar iterations. A progressive improvement is observed, as the curves approach the truth. Despite this improvement, the results are less accurate than the same figure in the previous section where the sensors were upstream of separation (figure 10). It is evident from the present results in figure 15a that the EnVar optimization in this present case is essential for improving the estimation. Errors nonetheless persist, approximately on the order of 10^{-2}, which is two orders of magnitude higher than those for the estimation using upstream sensors.

In figure 16, we report the energy spectra in the initial estimate (a.i) and the final control vector (b.i), as well as the errors relative to the hidden truth (a.ii, b.ii). The most amplified modes according to the N-factor within the computational domain are circumscribed by the red line. The result show an appreciable difference between the initial and optimal control vectors. Specifically, there is notable error reduction in the majority of the amplified modes, with the exception of $(140, 0)$, $(130, 0)$, and $(130, 429.1)$. Where errors are reduced, this is an indication that the linear estimate was inaccurate because the dynamics are nonlinear and, in addition, the sensors within the separation possess the necessary sensitivity to improve the accuracy of the initial estimate through the non-linear optimization procedure. However, it is important to note that the total inflow energy for the final estimate is $E_i = 1.29 \times 10^{-4}$, which has a 22.6% error compared to the hidden truth. In addition, the largest errors are roughly two orders of magnitude

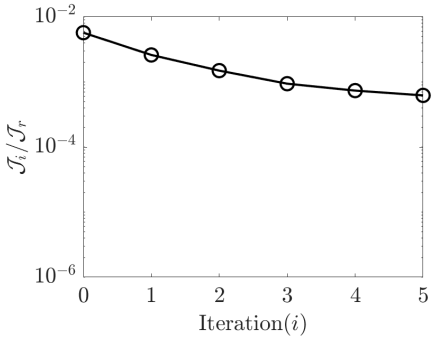

Figure 13: Cost reduction in flow estimation using sensors within the separated region.

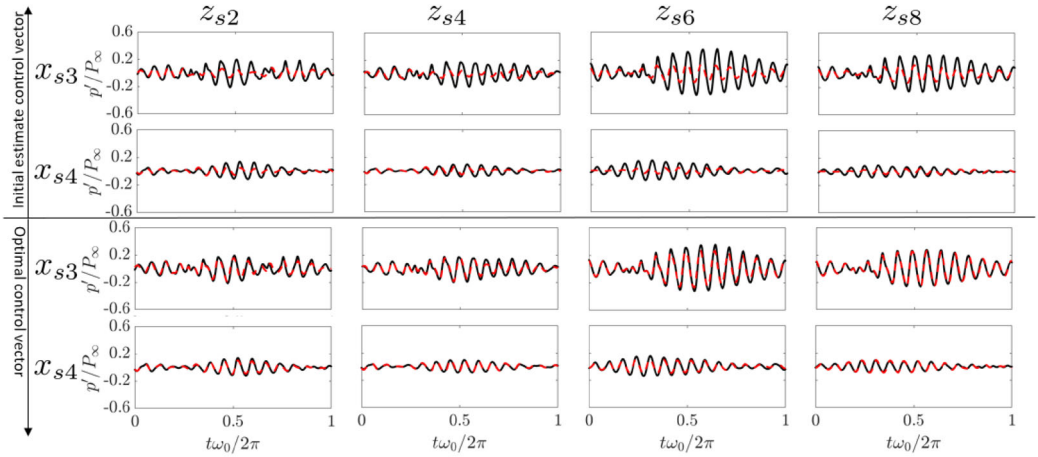

Figure 14: Synthetic observations (black line) and their estimate from the optimal control vector (red-dashed line), assimilated using data from sensors within separation.

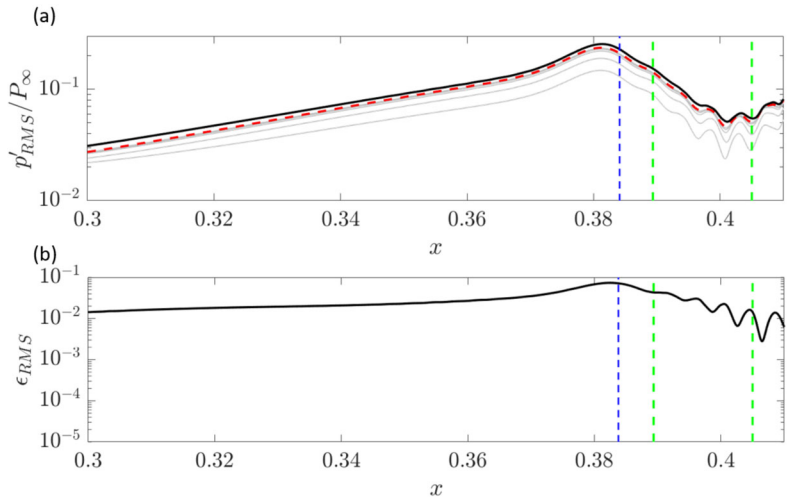

Figure 15: (a) RMS pressure perturbation and (b) their errors in the estimate by the optimal control vector. Flow estimation is performed using sensors within separation. Black line depicts the truth; red-dashed lines denote the estimated state. Separation is indicated by a blue-dashed line, and sensor locations are marked with green-dashed lines.

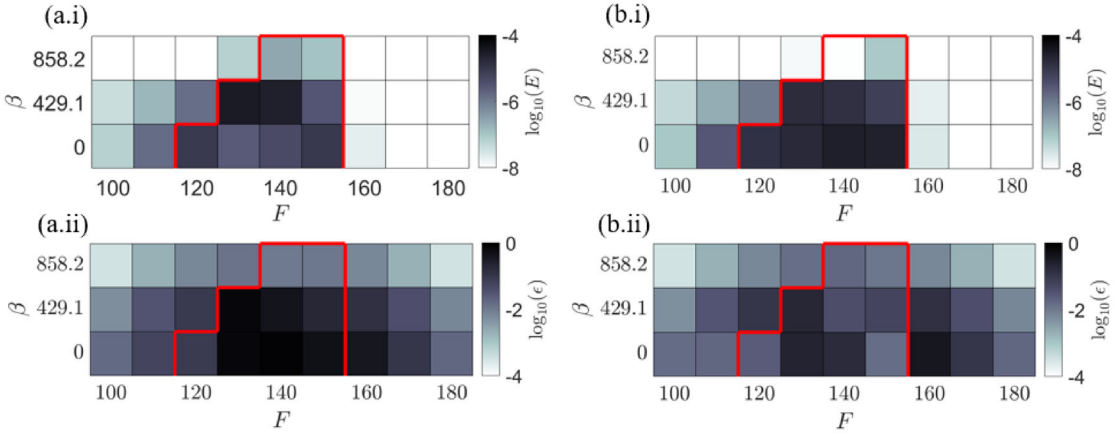

Figure 16: Modal energy and error in the estimation from sensors positioned within separation: (a) Initial estimate; (b) Final mean-control vector; (i) inflow spectra; (ii) error relative to the hidden truth. The most amplified modes according to the N-factor within the computational domain are circumscribed by the red line.

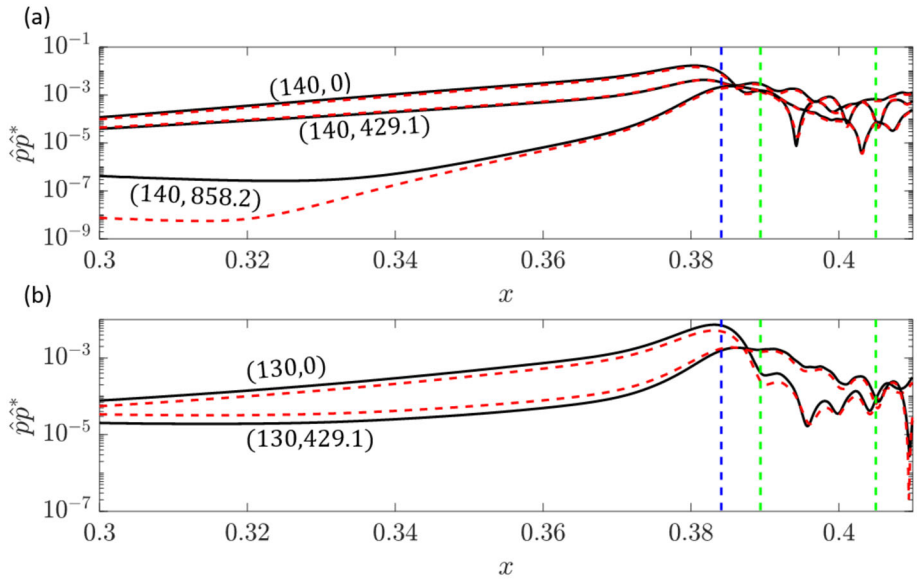

Figure 17: Comparison of the wall-pressure spectra from the (solid) reference simulation and (dashed) the assimilated flow, at (a) $F = 140$ and (b) $F = 130$. Separation is indicated by a blue-dashed line, and sensor locations are marked by green-dashed lines.

greater than those in figure 11 for upstream sensors. The inability to further improve the accuracy of the initial estimate suggests reduced sensor sensitivity in the separated region or a more difficult cost landscape to navigate. We will explore the former possibility in the next section.

The disparities observed between the true and estimated observations (figure 14) can be examined in further detail by considering the evolution of the spectra, which is reported in figure 17. While the measurements are only available at discrete sensor locations, here we take advantage of the fact that we have access to the hidden truth; we therefore plot the evolution of the instability waves from the inflow to the end of the horizontal section, and compare the estimated flow to the hidden truth. Figure 17 focuses on key frequencies, namely $F = \{130, 140\}$. Each of these modes reveals discernible errors at the

inflow, which appear relatively small closer to the sensor location within the separated region. Specifically, mode $(140, 858.2)$ shows similar behavior to earlier results with sensors upstream of separation: significant relative errors at the inflow when the modal energy is inappreciable, and relatively small relative errors near the sensor locations. Once again these results highlight that modal amplification is nonlinear, and that separation plays an important role.

The magnitude of the errors near separation is less noticeable for $F = 140$ than for $F = 130$. Modes at the latter frequency serve to highlight the pronounced discrepancies at the sensor locations. Notable errors are evident at the inflow for both the planar and oblique modes, and persist downstream. Careful inspection of the errors in the estimated inflow modal energy (figure 16) shows that the most pronounced errors among the amplified modes (those delineated in red) are predominantly in $(130, 0)$ and $(130, 429.1)$. This amplified error, observable in both the inflow energy and spectral evolution, culminates in deviations between the reference and model observations, as portrayed in figure 14.

These results thus far demonstrate that flow estimation can be performed using sensors situated within a separated flow region. Nevertheless, while the non-linear optimization was able to reduce the error in the initial estimate, discrepancies observed in the inflow energy, time-dependent pressure, and spectral content across the domain indicate that the flow estimation is less accurate in comparison to the assimilation of sensor data upstream of separation. Furthermore, the estimation showed considerable errors associated with specific modes, particularly $(130, 0)$ and $(130, 429.1)$. These errors suggest diminished sensor sensitivity within the separated flow region. It should be stressed, however, that sensors within separation may still be essential as part of a complete measurement campaign. For example, such sensors may have sensitivity to certain frequencies that only amplify upon flow reversal and that are too weak to detect upstream. One such scenario was noted by Butler & Laurence (2021) for low frequencies that were only recorded in the separation region.

4.3. *Comparison of estimation accuracy*

It is beneficial to directly compare the estimation accuracy for sensors upstream of and within separation. Figure 18 shows the error in the estimated inflow spectra of the most energetic planar waves, reported as a function of EnVar iteration. At a glance, it is clear that the estimation of the inflow from the upstream sensors is two orders of magnitude more accurate than when the sensors are within the separated region. The error reduction in the latter has an interesting behaviour, where the energetic two-dimensional modes $(140, 0)$ and $(130, 0)$ appear immune to the EnVar optimization while the errors in $(120, 0)$ and $(150, 0)$ reduce.

The same visualization of the errors in the estimation of the oblique modes ($\beta = 429.1$) is shown in figure 19, with consistent interpretation. Specifically, the errors in the estimate of the inflow spectra is appreciably reduced when the assimilation is performed using sensors upstream of separation. However, when the sensor data is obtained from within the separated region, the errors remain roughly two orders of magnitude greater. In addition, similar to the behavior of the planar modes $(130, 0)$ and $(140, 0)$, the error for mode $(130, 429.1)$ persists at elevated levels throughout the non-linear optimization, seemingly unaffected by the optimization process. The reasoning for the apparent insensitivity of these modes to the assimilation process will be explored using the idea of sensor sensitivity in the next section.

Table 2 reports the separation and reattachment locations, along with the total inflow energy for the true flow and the results from the two data assimilation attempts. Consistent with the earlier discussion, the estimation using sensors within the attached boundary layer

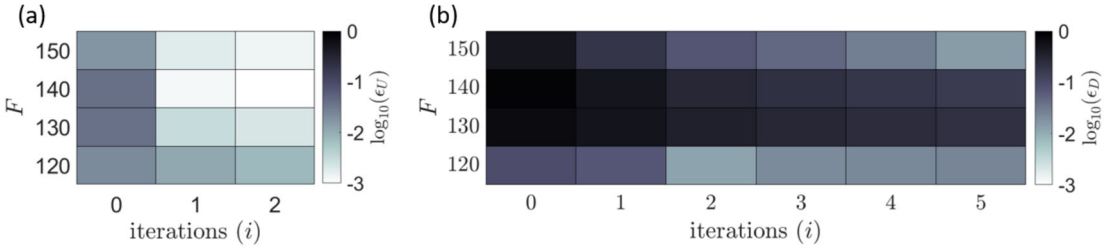

Figure 18: Errors in the estimation of the most dominant inflow planar waves versus EnVar iteration: (a) sensors upstream of separation, (b) sensors within separation.

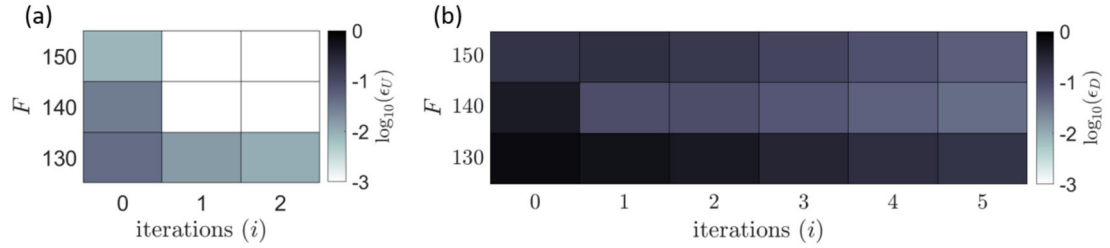

Figure 19: Errors in the estimation of the most dominant inflow oblique waves versus EnVar iteration: (a) sensors upstream of separation, (b) sensors within separation.

Case	$x_{\text{separation}}$	$x_{\text{reattachment}}$	$\boldsymbol{E_t}$
True flow	0.3841	0.4374	1.6×10^{-4}
Estimate from sensors upstream of separation	0.3842	0.4373	1.59×10^{-4}
Estimate from sensors within separation	0.3837	0.4373	1.29×10^{-4}

Table 2: Comparison of true and estimated flow quantities: streamwise location of separation onset, streamwise location of reattachment, and total inflow energy.

most accurately reproduces these parameters, in particular the total energy. When the sensors are within the reverse-flow region, the estimate of separation onset is shifted slightly upstream, but reattachment is accurately predicted. The inflow energy is under-predicted by approximately 22.6%, which is consistent with the upstream shift in separation onset.

5. Sensor sensitivity

Sensor sensitivity is an insightful approach to interpret the data-assimilation performance. Multiple methodologies exist for obtaining the sensitivity, and they vary significantly in terms of computational cost. For instance, the data from the ensemble of solutions in the EnVar algorithm can be used for statistical analysis of the sensor correlation with the inputs. Here we adopt an adjoint approach, which is an effective and computationally efficient technique to acquire sensor sensitivity (Wang *et al.* 2022, 2019*a*; Zaki & Wang 2021).

We focus on the difference between a single measurement data m and its model observation h. The latter is obtained by applying an observation kernel onto the flow state, represented as $h = \mathcal{M}(\boldsymbol{q}) = \mathcal{M}(\mathcal{N}(\boldsymbol{q}_0)) = \int_{\mathcal{T}} \int_{\mathcal{V}} p\delta(\boldsymbol{x} - \boldsymbol{x}_m, t - t_m) d\mathcal{V} dt$. The cost

function is then,

$$\mathcal{J}_a = \frac{1}{2}\left(m-h\right)^2. \tag{5.1}$$

Taking the variation of the cost shows the impact of small deviations in the state, and hence in the model observation, to the overall cost,

$$\delta\mathcal{J}_a = (m-h)\underbrace{\left\langle \frac{\partial\mathcal{M}}{\partial\boldsymbol{q}}, \frac{\partial\boldsymbol{q}}{\partial\boldsymbol{q}_0}\delta\boldsymbol{q}_0 \right\rangle}_{\equiv\delta h} = (m-h)\underbrace{\left\langle \frac{\partial\mathcal{M}}{\partial\boldsymbol{q}}, \mathcal{L}_{\boldsymbol{q}}\delta\boldsymbol{q}_0 \right\rangle}_{\equiv\delta h}. \tag{5.2}$$

The first term $(m-h)$ is the difference between the measurements and the model observation, the second term δh is the variation in the model observation, and angle brackets denote Chu's energy norm (Chu 1965; Hanifi *et al.* 1996). The operator $\mathcal{L}_{\boldsymbol{q}}$ is the linearized Navier-Stokes equations. Using integration by parts, we can re-express the variation in the observation as,

$$\delta h = \left\langle \frac{\partial\mathcal{M}}{\partial\boldsymbol{q}}, \mathcal{L}_{\boldsymbol{q}}\delta\boldsymbol{q}_0 \right\rangle = \left\langle \underbrace{\mathcal{L}^\dagger_{\boldsymbol{q}}\frac{\partial\mathcal{M}}{\partial\boldsymbol{q}}}_{=\boldsymbol{q}^\dagger_0}, \delta\boldsymbol{q}_0 \right\rangle. \tag{5.3}$$

The adjoint operator $\mathcal{L}^\dagger_{\boldsymbol{q}}$ acting on $\partial\mathcal{M}/\partial\boldsymbol{q}$ is advanced in backward time to obtain the adjoint field $\boldsymbol{q}^\dagger_0$. As such, forward-adjoint duality can be expressed as,

$$\delta h = \left\langle \frac{\partial\mathcal{M}}{\partial\boldsymbol{q}}, \mathcal{L}_{\boldsymbol{q}}\delta\boldsymbol{q}_0 \right\rangle = \langle \boldsymbol{q}^\dagger_0, \delta\boldsymbol{q}_0 \rangle. \tag{5.4}$$

The implication of the above equation is profound: The adjoint field $\boldsymbol{q}^\dagger_0$ in equation (5.4) represents sensitivity in the sensor observation $\boldsymbol{h}$ to any perturbation of the initial state $\delta\boldsymbol{q}_0$. Figure 20 presents an example of the adjoint field for the two-dimensional compression ramp, illustrating sensitivity of the sensor at x_{s4} within the separation region; the red-dashed line marks the edge of the boundary layer (δ_e), while the green-dashed line marks the separated region. The contours show the adjoint pressure $p^\dagger$, or the sensivity of the sensor to an earlier-in-time and space pressure perturbations. Given that the adjoint equations operate in reverse time, the adjoint pulse $\partial\mathcal{M}/\partial\boldsymbol{q}$ emanates from the sensor location x_{s4}, at the measurement time t_m. This pulse then advects upstream, culminating at the position and temporal point displayed in figure 20. Regions in the figure with non-zero values signify areas to which the sensor is sensitive.

An examination of the adjoint field reveals distinct zones of concentrated sensitivity: the separated region, the shear layer directly above this region, the upstream boundary layer leading up to the point of separation, and the free stream. Increased sensitivity is evident at the upper boundary of the shear layer, situated directly above the separated region, and also along the edge of the boundary layer upstream of separation.

The adjoint field is initiated by the measurement kernel at the sensor location $\boldsymbol{x}_m$ and time t_m, which is a delta function $\delta(\boldsymbol{x}-\boldsymbol{x}_m, t-t_m)$. The field thus encompasses the sensitivity of the sensor to all resolved frequencies and wavenumbers. Here we will focus on the sensitivity to planar waves ($\beta = 0$), which we evaluate using a laminar base state and computing the adjoint in a two-dimensional domain. We extract the frequency spectra by Fourier transforming the time-dependent adjoint fields. We then evaluate the wall-normal integrated energy of the adjoint field at every frequency F, and report these results in figure 21. The sensor at x_{s1} has a distinct peak in its sensitivity to the unstable second-modes waves, at all streamwise locations upstream of the sensor. In contrast, the sensitivity of the sensor at x_{s4}, within the separation region, oscillates as a function

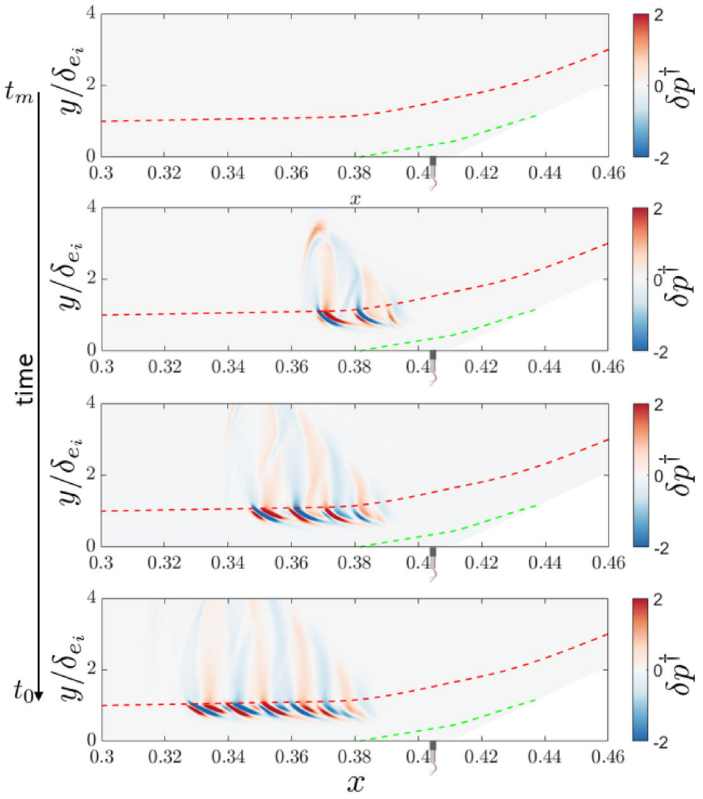

Figure 20: Adjoint simulation for sensor location x_{s4}, showing contours of adjoint pressure $p^\dagger$ throughout the domain. The time starts at the measurement time t_m and progresses to the initial time t_0.

of frequency. This sensitivity to the second-mode waves in this case is relatively low compared to the upstream sensor, and the sensitivity to lower frequencies is relatively enhanced.

Within the scope of this study, we are primarily concerned with the sensitivity to the inflow, where we optimize the modal amplitude in order to reproduce the sensor measurements. Figure 22 shows this sensitivity as a function for non-dimensional frequency F. The results at sensor x_{s1}, located upstream of the separation, predominantly show sensitivity to the second mode frequencies. For the downstream sensor x_{s4}, the sensitivity to the second mode frequencies is reduced and hence the relatively less accurate estimation of the inflow using these sensors. The sensor at x_{s4} also has heightened sensitivity in the lower frequency range, specifically within $F = [55, 80]$. It is important to consider this increased sensitivity when conducting data assimilation in a separated flow, especially when the objective is to reconstruct experimental observations where low-frequency oscillations may be recorded within the separation region. These modes were not part of the inflow condition in the synthetic experiment that generated the present measurements, and hence this effect was not clear in the assimilation. Including these low frequencies in the synthetic experiment would highlight the importance of sensors within the separated region for predicting these low frequencies at the inflow.

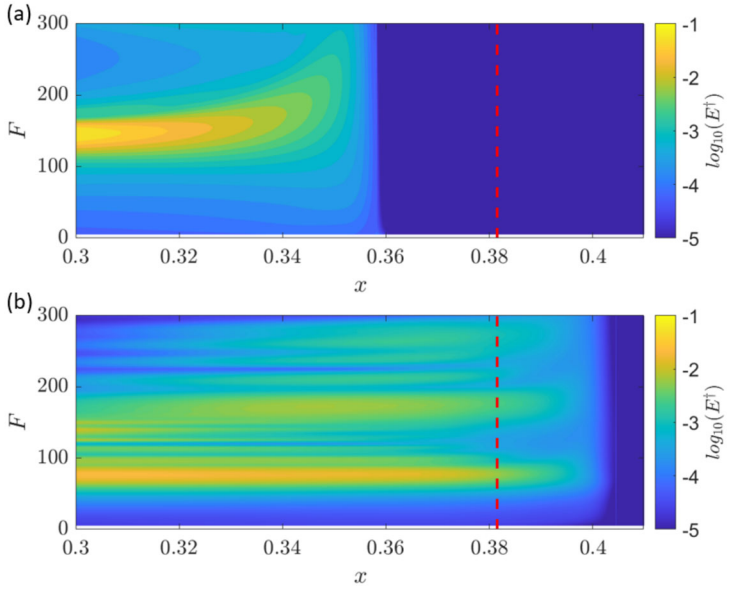

Figure 21: Adjoint sensitivity for sensors at (a) x_{s1} and (b) x_{s4}, as a function of frequency F and streamwise position x. The contours shows the adjoint energy integrated in the wall-normal direction, $E^\dagger$. The red dashed line marks the point of separation.

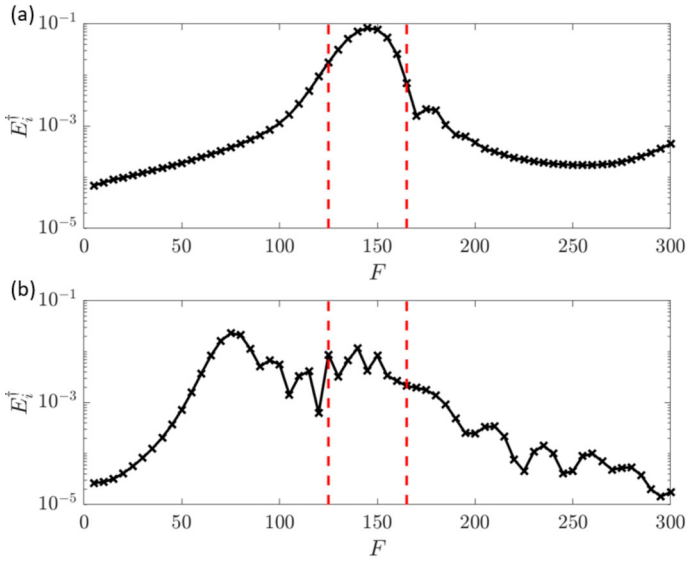

Figure 22: Sensitivity of sensors at (a) x_{s1} and (b) x_{s4} to the inflow. The adjoint energy $E^\dagger$ at the inflow plane is plotted versus the non-dimensional frequency F. The red-dashed lines highlighting the frequencies associated with the unstable second modes.

6. Summary

Estimating separated boundary-layer flows from wall-pressure measurements is challenging due to the qualitative change in the flow state across the separation point, and the associated change in the sensor sensitivity. In this work, we considered separation on a compression ramp, and used synthetic measurements collected from an independent direct numerical simulations. This choice enabled us to precisely place our measurements, obtain noise-free sensor signals, and assess the accuracy of our predicted flow state using

data assimilation against the true flow. In order to investigate the influence of separation on the accuracy of the data assimilation, two estimations were performed: one with measurements from sensors upstream of separation, and the other with observations from sensors placed within the separated region. We analyzed our results and explained them through the lens of sensor sensitivity.

The flow estimation using the sensors upstream of the separated region accurately reconstructed the flow, with just 7.6% error in total inflow energy. Furthermore, errors in the inflow spectra for the most unstable modes relative to the truth were minimal. In contrast, the flow estimation using sensor data from within the separated region displayed was significantly less accurate. The error in the total inflow energy was approximately 22.6%, and the errors in the inflow spectra for the most unstable modes were also appreciable. In particular, the assimilation was unable to reduce the errors in modes $(F, \beta) = (130, 0)$, $(140, 0)$, and $(130, 429.1)$.

The sensor sensitivity analysis for sensors within the separation region highlighted a diminished sensitivity to second-mode frequencies. However, there was also heightened sensitivity to lower-frequency content, which is relevant to prediction of the low-frequency modes that are often observed in separated flows.

REFERENCES

Anderson, Jeffrey L & Anderson, Stephen L 1999 A monte carlo implementation of the nonlinear filtering problem to produce ensemble assimilations and forecasts. *Mon. Wea. Rev.* **127** (12), 2741–2758.

Brent, Richard P 2013 *Algorithms for minimization without derivatives*. Courier Corporation.

Buchta, David A, Laurence, Stuart J & Zaki, Tamer A 2022 Assimilation of wall-pressure measurements in high-speed flow over a cone. *J. Fluid Mech.* **947**, R2.

Buchta, David A & Zaki, Tamer A 2021 Observation-infused simulations of high-speed boundary-layer transition. *J. Fluid Mech.* **916**, A44.

Butler, Cameron 2021 Response of hypersonic boundary-layer disturbances to compression and expansion corners. PhD thesis, University of Maryland, College Park.

Butler, C.S. & Laurence, S.J. 2021 Interaction of second-mode disturbances with an incipiently separated compression-corner flow. *J. Fluid Mech.* **913**, R4.

Cai, Shengze, Wang, Zhicheng, Lu, Lu, Zaki, Tamer A & Karniadakis, George Em 2021 Deepm&mnet: Inferring the electroconvection multiphysics fields based on operator approximation by neural networks. *J. Comput. Phys.* **436**, 110296.

Chu, Boa-Teh 1965 On the energy transfer to small disturbances in fluid flow (part i). *Acta Mechanica* **1** (3), 215–234.

Clemens, N. T. & Narayanaswamy, V. 2014 Low-frequency unsteadiness of shock wave/turbulent boundary layer interactions. *Annu. Rev. Fluid Mech.* **46**, 469–492.

Di Leoni, Patricio Clark, Lu, Lu, Meneveau, Charles, Karniadakis, George Em & Zaki, Tamer A 2023 Neural operator prediction of linear instability waves in high-speed boundary layers. *J. Comput. Phys.* **474**, 111793.

Du, Yifan, Wang, Mengze & Zaki, Tamer A 2023 State estimation in minimal turbulent channel flow: a comparative study of 4dvar and pinn. *Int. J. Heat Fluid Flow* **99**, 109073.

Du, Yifan & Zaki, Tamer A 2021 Evolutional deep neural network. *Physical Review E* **104** (4), 045303.

Evensen, Geir & others 2009 *Data assimilation: the ensemble Kalman filter*, , vol. 2. Springer.

Fedorov, A. 2011 Transition and stability of high-speed boundary layers. *Annu. Rev. Fluid Mech.* **43**, 79–95.

Hanifi, A., Schmid, P. J. & Henningson, D. S. 1996 Transient growth in compressible boundary layer flow. *Phys. Fluids* **8** (3), 826–837.

Hao, Yue, Di Leoni, Patricio Clark, Marxen, Olaf, Meneveau, Charles, Karniadakis, George Em & Zaki, Tamer A 2023 Instability-wave prediction in hypersonic boundary layers with physics-informed neural operators. *J. Comput. Sci.* **73**, 102120.

Jarratt, P 1967 An iterative method for locating turning points. *Comput. J.* **10** (1), 82–84.

Mack, Leslie M 1984 Boundary-layer linear stability theory. *Tech. Rep.*. California Inst of Tech Pasadena Jet Propulsion Lab.

Mao, Zhiping, Lu, Lu, Marxen, Olaf, Zaki, Tamer A & Karniadakis, George Em 2021 Deepm&mnet for hypersonics: Predicting the coupled flow and finite-rate chemistry behind a normal shock using neural-network approximation of operators. *J. Comput. Phys.* **447**, 110698.

McClure, William Berton 1992 *An experimental study of the driving mechanism and control of the unsteady shock-induced turbulent separation in a Mach 5 compression corner flow*. The University of Texas at Austin.

Mons, V., Du, Y. & Zaki, T. A. 2021 Ensemble-variational assimilation of statistical data in large-eddy simulation. *Phys. Rev. Fluids* **6** (10), 104607.

Mons, Vincent, Wang, Qi & Zaki, Tamer A 2019 Kriging-enhanced ensemble variational data assimilation for scalar-source identification in turbulent environments. *J. Comput. Phys.* **398**, 108856.

Park, Junho & Zaki, Tamer A 2019 Sensitivity of high-speed boundary-layer stability to base-flow distortion. *J. Fluid Mech.* **859**, 476–515.

Wang, Mengze 2021 Data assimilation in wall turbulence. PhD thesis, Johns Hopkins University.

Wang, Mengze, Wang, Qi & Zaki, Tamer A. 2019*a* Discrete adjoint of fractional-step incompressible navier-stokes solver in curvilinear coordinates and application to data assimilation. *Journal of Computational Physics* **396**, 427–450.

Wang, Mengze & Zaki, Tamer A 2021 State estimation in turbulent channel flow from limited observations. *J. Fluid Mech.* **917**, A9.

Wang, Mengze & Zaki, Tamer A 2022 Synchronization of turbulence in channel flow. *J. Fluid Mech.* **943**, A4.

Wang, Q., Hasegawa, Y. & Zaki, T. A. 2019*b* Spatial reconstruction of steady scalar sources from remote measurements in turbulent flow. *J. Fluid Mech.* **870**, 316–352.

Wang, Qi, Wang, Mengze & Zaki, Tamer A. 2022 What is observable from wall data in turbulent channel flow? *J. Fluid Mech.* **941**, A48.

Zaki, Tamer A. & Wang, Mengze 2021 From limited observations to the state of turbulence: Fundamental difficulties of flow reconstruction. *Phys. Rev. Fluids* **6**, 100501.