Modeling Strategies for Aerodynamic Flow Reconstruction from partial measurements

Lucas Franceschini

Lucas Franceschini. Modeling Strategies for Aerodynamic Flow Reconstruction from partial measurements. Fluids mechanics [physics.class-ph]. Université Paris Saclay, 2019. English. NNT : . tel-02962789v1

HAL Id: tel-02962789
https://hal.science/tel-02962789v1

Submitted on 14 Oct 2020 (v1), last revised 17 May 2021 (v2)

université
PARIS-SACLAY

ONERA
THE FRENCH AEROSPACE LAB

NNT : 20XXSACLXXXX

Thèse de doctorat

Modeling Strategies for Aerodynamic Flow Reconstruction from Partial Measurements

Thèse de doctorat de l'Université Paris-Saclay
préparée à l'ONERA

Ecole doctorale n°579 Sciences mécaniques et énergétiques, matériaux et géosciences (SMEMAG)
Spécialité de doctorat : voir spécialités par l'ED

Thèse présentée et soutenue à Meudon, le
4 novembre 2019, par

LUCAS FRANCESCHINI

Composition du Jury :

Beverley McKeon Directrice de Recherche, CalTech (Department of Aerospace)	Rapporteur
Etienne Mémin Directeur de Recherche, IRMAR, INRIA (Fluminance)	Rapporteur
Lutz Lesshafft Professeur, École Polytechnique (LadhyX))	Examinateur
Paola Cinnella Directrice de Recherche, ENSAM (DynFluid)	Examinateur
Carlo Cossu Professor, École Centrale de Nantes (LHEEA)	Examinateur
Hervé Bezard Ingénieur de Recherche, ONERA (DAAA)	Examinateur
Denis Sipp Directeur de Recherche, ONERA (DAAA)	Directeur de thèse
Olivier Marquet Ingénieur de Recherche, ONERA (DAAA)	Encadrant

Contents

1 Introduction 1
1 Context and Motivations 1
2 Data-Assimilation Techniques 2
2.1 Kalman Filters (KF) 2
2.2 Variational data-assimilation (VDA) 2
3 Goal of the Present Work 3
3.1 Mean-flow data-assimilation 4
3.2 Beyond mean-flow: stability analysis 5
4 Overview of the Thesis 6

2 Data-Assimilation for minimal Corrections of RANS equations 9
1 Introduction 9
2 Flow Configuration and numerical solutions 11
3 Models and Data assimilation procedure 14
3.1 Model 1: source term in momentum equation 14
3.2 Model 2: source term in $\tilde{\nu}$ equation 15
3.3 Data Assimilation procedure 15
4 Results 17
4.1 Complete velocity field measurements 18
4.2 Sparse velocity measurements 22
4.2.1 Sparse velocity measurement operator 22
4.2.2 Reconstruction results using model 1 and model 2 23
4.2.3 Optimization with gradient penalization of model 1 25
4.3 Understanding the 'rigidity' of model 2 through observability Gramian analysis 25
5 Conclusion 26

3 Wall-measure assimilation for RANS equations 29
1 Introduction 29
2 Wall-measure for laminar RANS equations 30
2.1 Theoretical developments on laminar equations 30
2.2 Physical Description of the Considered Laminar Flow 33
2.3 Results 33
2.3.1 Wall-Pressure Assimilation 34
2.3.2 Wall-Friction Assimilation 35
3 Wall-measure for turbulent RANS 36
3.1 Results for turbulent backward-facing step 36

3.1.1 Wall-Friction Information 36
3.1.2 Wall-Pressure Information 38
3.1.3 Combined Measure: Wall-Pressure and Velocity 42
4 Conclusions 42

4 Reslvent-based Data-Assimilation 45
1 Introduction 45
2 Theory 47
2.1 Time-periodic flows 49
2.2 Broadband Flows 49
2.3 Numerical resolution of models 50
3 An ”Oscillator” Flow - Squared-Section Cylinder 51
3.1 Configuration and numerical implementation 52
3.2 Single point velocity measurement 52
3.3 Other measurements 54
4 Noise Amplifier Flows - Backward-facing step 58
4.1 Configuration 58
4.2 Upstream noise with two incommensurable frequencies: $\omega = 0.48$ and $\omega = 0.70$ 59
4.3 Upstream noise with broadband white noise forcing 62
4.3.1 Objective functional 64
4.3.2 Optimization procedure 66
4.3.3 Reconstruction results 66
5 Conclusions and Perspectives 68

5 Mean-flow analysis for the mixed periodic/broadband fluctuations on turbulent flows 71
1 Introduction 71
2 Triple Decomposition and Mean-Flow analysis 73
2.1 Modelling of quasi-periodic component 74
2.1.1 Modelling with mean-flow marginal global modes 75
2.1.2 How to obtain ν_t 76
2.2 Resolvent analysis for broadband component 76
2.2.1 Steady flow approximation: $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$ 78
2.2.2 Quasi-static Approximation 78
2.3 Remarks on the choice of averaging procedure and eddy-viscosity 79
3 Mean-Flow analysis on DNS data 79
3.1 Direct Numerical Simulation solution 80
3.2 Mean-Flow Stability Analysis: periodic component 81
3.3 Resolvent analysis: broadband component 83
4 Conclusion 87

6 Resolvent-based Data-Assimilation for the turbulent flow around a squared-section cylinder 89
1 Introduction 89
2 Development of the model for Data-Assimilation 90
2.1 RANS and URANS Spalart-Allmaras Solutions 90

- 2.2 URANS Mean-Flow resolvent Analysis 92
- 2.3 Data-Assimilation Model 93
- 3 Results 94
- 4 Conclusion 96

7 Conclusions and Perspectives 99

- 1 RANS-based Data-Assimilation 99
- 2 Mean-flow analysis and Data-Assimilation 100

A RANS - Spalart-Allmaras Numerical Resolution 103

- 1 Incompressible RANS 103
  - 1.1 Original model 103
  - 1.2 Modification in S 104
  - 1.3 Negative $\tilde{\nu}$ modifications 104
  - 1.4 Negative r modifications 105
- 2 Stabilization procedure - SUPG 107
- 3 Numerical tests and Mesh Convergence 109
  - 3.1 RANS-SA Solution 109
  - 3.2 Resolvent mode analysis 110
  - 3.3 Data-Assimilation model resolution 111

B BFGS Algorithm and its use with a Finite-Element Solver 113

- 1 BFGS and Line search Methods 113
- 2 BFGS with Finite-Element Solver 114
  - 2.1 A toy model for EDP solver - Laplace equation 114
  - 2.2 Optimization Problem 115
  - 2.3 Optimization Results 116
- 3 Remarks on Navies-Stokes equations: Incompressibility of $\tilde{\mathbf{f}}_\mathbf{x}$ 118
- 4 Adjoint Discretization in FEM 119

Chapter 1

Introduction

1 Context and Motivations

Experimental and numerical fluid dynamics have been used for several years as tools to investigate and understand the physics of fluid flow phenomena. Separately, they provide both useful and different insights on those phenomena. From the point of view of the numerical scientist or engineer, we usually have complete control on the numerical simulation we run in the sense that we can design the geometry of the physical settings, apply different boundary or initial conditions and investigate several regimes of non-dimensional numbers, such as the Reynolds and Mach numbers, etc. However, all this freedom has a drawback since, in order to produce a solution that represents a real physical experiment, we need to precisely know those actual quantities, such as boundary and initial conditions. Furthermore, apart from those aspects, the numerical model itself is often seen as a major source of discrepancy between the resulting numerical solution and what happens in real flow phenomena, especially in turbulent flow simulations, as it will be discussed later. On the other hand, experimental scientists and engineers have developed several tools to probe fluid flows, generating data from real experiments that are often regarded as more valuable than numerical ones. However, this experimental approach suffers from several drawbacks as well. For example, the data extracted this way are often insufficient for describing, in a complete manner, the fluid flow. For example, having the wall-pressure only does not provide direct and precise information of what goes on the bulk of the flow. Another related issue is the fact that the extracted data is often contaminated by external noise, leading to an uncertainty on those quantities. For those reasons, those two techniques are regarded as complementary and have been extensively used in parallel as cross-validation of one another. For example, having a numerical solution that matches well with the experimental data gives us confidence to extrapolate this solution to regions where the probes could not reach, enlarging this way our understanding of the physical phenomenon. Although this approach provides several insights, it becomes less useful when the numerical solution does not match well the experimental data. In those cases, the numerical simulation (including the model, boundary and initial conditions, etc), becomes questionable and may lose its predictive power. It is interesting then to modify the model such that its solution matches the experimental data, improving this way its quality and, hopefully, its predictability in regions far from the probe locations. For this purpose, the data-assimilation procedure has been designed as a mathematically well-defined manner to include the experimental data in the model, correcting it. In the literature, several approaches in several different contexts have been developed in order to solve this problem. In what follows, we provide a brief description of the most common approaches, identified in two (maybe overlapping) groups, Kalman Filters and Variation data-assimilation, followed by some of their applications in the context of Fluid Mechanics. We will then precise what is the goal of the present thesis, for which we present a brief state-of-the-art, followed by an overview of the thesis.

2 Data-Assimilation Techniques

2.1 Kalman Filters (KF)

The data-assimilation concept goes back to the theory of modern control of linear dynamical systems (see Kalman [1960]), where an estimation of the whole state of the system from limited partial measurements is needed in order to act on the system with a control law, stabilizing it. One can show that, for deterministic linear systems (under the condition of observability), one can always design a (not unique) dynamical system, called observer, mimicking the dynamics of the original one, such that its output tends towards the true state as time progresses. However, if noise is present, the performance of such observer can be drastically deteriorated and the error committed in the prediction of the state can be large. The Kalman filter is the design of such observer so that its prediction is as close as possible to the actual state. The final form of this filter can be seen as a prediction step, where the model of the dynamical system is applied on the *a priori* state, and a correction step, where the external data (or measure) is taken into account through the filter, which is constructed based on the covariance matrix of the state vector. We remark that, in this procedure, not only the state vector is propagated in time, but also its corresponding covariance matrix. For this reason, this approach is mainly used when the system has only a few degrees of freedom, as it is often the case in many engineering application, not involving Partial Differential Equations (PDE). However, in cases where the dimension of the state vector is prohibitively large, the propagation of the covariance matrix becomes unfeasible. For this purpose, the Kalman filter has been modified with an approximation of this matrix, consisting in the propagation of a cloud of points (or ensemble) in the phase space, instead of only one point. The evolution of this cloud provides, if the number of points is sufficiently large, an estimation of the statistics. In those circumstances, this filter is called Ensemble Kalman Filter (EnKF, see Evensen [2009]). This concept was first introduced by Evensen [1994] in the context of oceanography.

In fluid mechanics, this technique has been applied for several purposes: the estimation of velocity field from Particle-Tracking Velocimetry snapshots in a planar-jet configuration at $Re = 2000$ (Suzuki [2012]), the recovery of low-Reynolds wall-bounded turbulent flows from wall measure (Colburn et al. [2011]), estimation of Mach number, angle of attack and eddy-viscosity field at high Reynolds number flows in complex flow configuration Kato et al. [2015].

2.2 Variational data-assimilation (VDA)

The variational data-assimilation technique is based on the theory of optimal control (see, for example Lions [1971]) where we act on a system such that its output satisfies as closely as possible some target task. In the data-assimilation context we dispose in general of some tuning parameter, that plays the role of the control variables, which is related to the model that we effectively solve (which can be initial/boundary conditions, the spatial distribution of an unknown forcing term, etc). This tuning parameter will be found such that the model's solution is as close as possible to the external reference data. One of the earliest applications of this technique was on the weather forecast problem (Talagrand and Courtier [1987], Liu et al. [2008]) under the name of 3D/4D-Var where the tuning parameters were the initial condition of the dynamical model such that some cost, related to the error between the model's prediction and the external data is minimal. In this context, we can see that the main difference between this approach and Kalman Filters is that, VDA can be posed as an optimal problem on the initial condition where EnKF can be view as a way to correct an *a priori* condition with the information of the external data. Indeed, in general, VDA problems can be viewed as optimization under constraint ones and, when the number of tuning parameters becomes large, gradient-descent methods need to be employed so that the optimization procedure reaches a local minimum efficiently. This gradient is often obtained with the aid of the Lagrangian formalism where the adjoint problem needs to be solved. This adjoint problem represents in general a supplementary numerical tool (other than the model's resolution) and may pose some difficulties.

Several examples can be presented in the context of fluid mechanics and aerodynamics, such as estimation of boundary and/or initial conditions from PIV data using DNS in low Reynolds numbers configurations (Gronskis A. [2013], Yegavian et al. [2015]), estimation of fluid flow from a sequence of sparse snapshots (Heitz and Mémin [2010]), as a tool for determining pollutant sources in urban areas, where Large Eddy Simulations (LES), coupled with Proper Orthogonal Decomposition (POD), are used in the data-assimilation scheme (Mons et al. [2017b]), and to determine the fluid flow on a POD-based lower order system (D'Adamo et al. [2007]).

We remark that, although those two techniques differ in nature, some hybrid ones may exist as well. For example, in order to bypass the need for an adjoint solver, necessary for VDA, we may consider instead the idea of ensembles, where, in order to compute the sensitivity of the cost functional with respect to the control parameters, a cloud of points is propagated, each one of those points presenting a different outcome. This technique is called Ensemble-based Variational assimilation (EnVar). It was first designed by Liu et al. [2008] in the context of weather prediction. More recently, Mons et al. [2016]), performed some studies on this technique, in comparison with EnKF and VDA in the context of Fluid Mechanics.

3 Goal of the Present Work

At this point, data-assimilation has been presented in a very general manner, together with a few applications it may have in Fluid Mechanics. In what follows, we intent to present a more specific overview of the precise problem we will consider in this work. This particularization will come mainly from the type of problem we will consider as numerical model and on the measures that will be the most compatible with the problem.

In most of the engineering applications, for example in aerodynamics, the time-averaged (mean) flow is sufficient for the global understanding of the fluid flow, allowing us to design mechanical components, to evaluate loads on structures, etc. Indeed, in most of industrial wind tunnel facilities, the nature of the data extracted is usually time-averaged. Examples of such data are the mean wall-pressure or friction distributions or mean-velocity fields, obtained with either point-wise probes (such as Pitot) or Particle Image Velocimetry (PIV) techniques. From the numerical point of view, a very common tool used in industry is the Reynolds-Averaged Navier-Stokes (RANS) equations, whose solution represents the mean-flow. Since those equations are steady in nature, their solution does not require a fine time-stepping procedure, minimizing the computational effort to solve them. This makes them very popular in industry, where the procedure of engineering design often demands easily obtainable solutions. The RANS equations are formally derived from the Navier-Stokes ones, by applying the Reynolds decomposition of the fluid flow variables, separating them into a mean-flow component and a fluctuation around it. Due to nonlinearities, the mean-flow variables in RANS equations will depend on the statistics of the fluctuation, given by the Reynolds-stress tensor. Those statistics cannot be explicitly computed since the fluctuation is not known, once we solve only the mean-flow component. Several models exist that attempt to provide an expression of this tensor in terms of mean-flow quantities and possibly of supplementary variables. This is the so-called closure problem. Generally those models are build using physical considerations such as eddy-viscosity, mixing-lengths, scaling laws, boundary layer theory, etc. This physical reasoning, in general, is capable of fixing somewhat the structure of the model, but may leave a set of parameters or coefficients to be specified. The specification of those parameters is done by considering some benchmark physical configurations, providing a calibration for them. This means that, for those configurations, the models are designed to work well. However, if we consider other configurations far away from those design points, typically with strong separation regions, the model may lose its accuracy when compared with experimental or higher-fidelity numerical data.

The main goal here is then to modify those models (or propose others) in a way that their solution matches the best some higher-fidelity data. We can see that this kind of problem can easily be posed in the logic of Variational Data-Assimilation (VDA) by identifying a tuning parameter, related with the Reynolds-stress tensor itself or with the closure model. This is the main goal of this work. In what follows, we present the brief state-of-the-art of mean-flow data-assimilation, together with a common application of it, the mean-flow stability analysis (in a broad sense), which will also be explored in the present work. This analysis is capable to provide an approximation for the nonlinear fluctuation around this mean-flow.

3.1 Mean-flow data-assimilation

The mean-flow data-assimilation, as presented in the previous section, has been considered in the past few years mainly from the point of view of variational data-assimilation, where some parameters related to the Reynolds-stress tensor are tuned such that the solution of RANS equations matches the best with external data. Foures et al. [2014] proposed a model consisting in the laminar, steady Navier-Stokes, where the correction parameter is a volume-force modeling directly the divergence of the Reynolds-stress tensor. They showed that, in a laminar / transitional context ($Re = O(10^2)$) that this model performs well to assimilate synthetic Particle Image Velocimetry (PIV) data. The same model was applied in a higher-Reynolds number context ($Re = O(10^4)$, Symon et al. [2017]). They reported some difficulties related to the use of the laminar equations in a higher

Reynolds number configuration. To circumvent those difficulties, a series of increasing Reynolds number data-assimilation needed to be performed, which can be costly. Other approaches, more appropriate for high-Reynolds number flows, rely on a RANS turbulence model. For example Li et al. [2017] tuned the $\kappa - \omega$ model's parameters such that its solution is the closest to reference ones. Singh and Duraisamy [2016], on the other hand, proposed a field inversion method where a spatially-dependent parameter related to the eddy-viscosity production term is tuned. In their work, the output solution was used for machine learning techniques such that the considered model becomes improved by learning the modifications caused by the data-assimilated procedure. We remark that, although those techniques are based on Variational Data-Assimilation, ensemble-based assimilation can also be used in a steady (mean) flow simulations contexts. Kato and Obayashi [2013] used ensemble techniques by varying the parameters (constants) in the Spalart-Allmmaras model, and a further Kalman filter, where the external measure was applied, leading to the converged optimal set of parameters. We remark that this kind of approach is somewhat equivalent to variational assimilation, where if the statistics are generated in an iterative procedure, with a large enough ensemble, VDA and KF solutions will be similar (Iglesias et al. [2013]).

3.2 Beyond mean-flow: stability analysis

In this paragraph we make a few remarks on the mean-flow and how it can be used to investigate aspects of the nonlinear fluid flow other than mean velocity-field, mean wall-pressure, etc. It was shown that, stability analysis (in a broad sense) using the mean-flow, instead of the steady-state solution, could provide an approximation of the unsteady nonlinear fluctuation around this mean-flow. Those studies have been formalized by Barkley [2006], who applied this technique to the flow around a circular cylinder at low Reynolds number. He showed that the nonlinear solution corresponds to a periodic limit-cycle, that can be decomposed in a mean-flow and on a fluctuation around it. The stability analysis of this mean-flow lead to an eigenvalue whose real part was zero and whose imaginary part matched very well with the nonlinear frequency. Later, Sipp and Lebedev [2007a] and Mezić [2013] provided conditions for the validity of this kind of analysis in the context of oscillator flows (flows whose dynamics do note depend on external forces) and showed that, under this conditions, the eigenmode provided an approximation for the first harmonics of the fluctuation, represented by the periodic limit-cycle. An extension of this analysis for flows whose eigenspectrum does not indicate unstable or neutral mode (for example, ”noise amplifiers”) was investigated by McKeon and Sharma [2010], Beneddine et al. [2016]. They proposed the use of Singular-Value decomposition of the Resolvent operator (transfer function of the mean-flow linearized system) rather than the eigenvalue/eigenvector analysis. They showed that, at a given frequency, if the leading Singular Value is much higher that the others, the nonlinear fluctuation could be approximated by the leading (left) Singular Vector at that frequency. They argued that this condition is verified in most flows presenting a strong instability mechanism. We remark however that, since this analysis is linear, it is not capable to determine alone the amplitudes (or energies) of those modes at each frequency. For this reason, most of the studies aiming to obtain the actual fluctuation needed time-resolved data, which is usually point-wise information, from which we can extract, for each frequency, the energy of the mode, together with its phase. Furthermore, if this low-rank representation of the

Resolvent operator holds, in principle only one (or a few, for robustness) time-resolved probe is sufficient for determining those amplitudes. We remark that even if this kind of measure is not time-averaged ones, the very low quantity of those probes still leaves this kind of technique feasible. For this reason, in the past few years, the determination of the mean-flow by variational data-assimilation and its further Resolvent-based analysis is increasingly popular technique, since it allows one to obtain an approximation of the full flow field from mean-flow measurements and very few time-resolved probes (see Symon et al. [2019a], Symon et al. [2019b], He et al. [2019a]).

4 Overview of the Thesis

In this work, we will investigate a few techniques for mean-flow data-assimilation, together with mean-flow analysis. More precisely, in the chapters 2 and 3, we will consider the modification of the RANS equations for mean-flow data-assimilation purposes, by tuning some parameters of the model. In the chapters 4,5 and 6, we will deal with mean-flow analysis and Resolvent-based data-assimilation. A detailed description of each chapter individually is provided:

- **Chapter 2**: In this chapter, we will discuss the mean-flow data-assimilation with the aid of the RANS equations, completed by a turbulence model, the Spalart-Allmaras. More precisely, we will introduce two different tuning parameters related to a correction of those equations. The first will be a volume force, following Foures et al. [2014], modeling the corrective force supplementary to the one coming from the eddy-viscosity provided by the SA model. We note that the high-Reynolds number issues reported by Symon et al. [2017] will not occur in our case since we dispose of a backup model (SA), providing enough eddy-viscosity such that the unforced solution (the original RANS-SA one) is well posed. This way, we avoid the need for iterative assimilation for increasing Reynolds numbers. The second one is a corrective source term acting in the SA equation (similar to Singh and Duraisamy [2016]), modifying how the production or destruction of the turbulent eddy-viscosity was designed. The data to be assimilated will be the mean-velocity field, either in the whole domain, to investigate controlability aspects of the proposed model, or at partial/sparse velocity measurements, as we would have using point-wise Pitot probes. By doing so we assess how those models perform in function on the quantity of information. The flow configuration will be a Backward-Facing Step (BFS) at $Re = 28275$, investigated in Dandois et al. [2007] with Direct Numerical Simulations (DNS), which will be used as ground-truth.

- **Chapter 3**: In this chapter we essentially investigate the same models as in the Chapter 2 but changing now the nature of the data to be assimilated. Here, we consider wall-information (pressure and friction). In a first time, we derive the adjoint equations for the laminar regime (without turbulence model), as their boundary conditions will be non standard, due to the wall-measurements. This formalism will also be, at first, tested in a laminar configuration. Then, we perform those data-assimilation on the turbulent case presented on Chapter 2, studying how the different models manage to assimilate this type of data, together with a few studies on the number of wall-probes necessary for convergence.

- **Chapter 4**: The mean-flow Resolvent analysis, as presented in §3.2, has been extensively used for the recovering of the nonlinear fluctuation. Most of the studies consider the mean-flow as a given quantity, both directly from experiments (or higher-precision numerical simulations) or from previous mean-flow data-assimilation. However, we remark that, if we know an approximation for the fluctuation, we can also have an approximation of the mean-flow. Indeed, from this fluctuations, we can compute the Reynolds-stress tensor and thus the forcing term on the RANS equations, that we could use to compute the mean-flow itself. For this reason, we propose here to skip the *a priori* knowledge of the mean-flow, proposing instead a nonlinear model composed by the mean-flow equation coupled with the Resolvent modes through the Reynolds-stress tensor. This model is not closed since, as mentioned before, the amplitudes of each mode (for each frequency) is unknown and will be tuned such that some quantity of the reconstructed flow matches the best with some external ones. Note that this measure can, in principle, be the same few time-resolved point-wise probes used by Beneddine et al. [2017] or Symon et al. [2019a], since their frequency spectra provide, individually for each frequency, the point-wise energy of the fluctuation. Moreover, those quantities may provide the information about which frequencies our model should have. For this reason, from the point of view of the input information needed, this as an extremely efficient data-assimilation technique. This technique will be applied on transitional flows: a squared-section cylinder, an example of flows whose instabilities saturate into a periodic limit cycle (oscillator flows), and a backward-facing step, an example of flows driven by external noise (noise amplifiers).

- **Chapter 5**: In this chapter we investigate the mean-flow stability analysis on flows whose fluctuation is composed by a quasi-periodic component and a broadband one. The application case is a squared-section cylinder at $Re = 22000$, whose quasi-periodic component is in fact represented by the periodic vortex-shedding and its broadband fluctuation is partially composed by Kelvin-Helmholtz (KH) structures located at the shear layer generated from the detachment point. To obtain both structures from mean-flow analysis (obtained from DNS), we decompose the flow variables using the triple decomposition as proposed by Mezić [2013], where the Koopman formalism is used, leading to two different equations, one for each one of them. The periodic vortex-shedding mode is obtained with a classical mean-flow stability. We investigate here the need of taking in consideration an eddy-viscosity, modeling the broadband fluctuation's nonlinear interactions and computed from the DNS statistics. The Kelvin-Helmholtz structures, on the other hand were obtained with a Resolvent analysis based on the linearized equations around the periodic component, instead of the mean-flow, similarly to what is done in Floquet stability theory. By using this formalism, we are able to capture the oscillation of those KH structures with the periodic vortex-shedding, which is also observed in DNS.

- **Chapter 6**: In this chapter, we intend to extend the Resolvent-based data-assimilation technique (presented in chapter 4) to the turbulent flow around the square cylinder, same configuration presented in chapter 5. This is a much more challenging task since now we have both periodic and broadband components to be taken into account. Numerically, this is a costly task and will be circumvented here with the use of a URANS model. The chosen one is the Spalart-Allmaras. This model provides naturally an eddy-viscosity, necessary for the resolution of the vortex-shedding mode and for the mean-flow equation, modeling on part of the Reynolds-stress tensor. The final model is composed by the vortex-shedding mode, whose amplitude is tuned similarly to what was done in the third chapter, coupled with the mean-flow equations.

Chapter 2

Data-Assimilation for minimal Corrections of RANS equations

1 Introduction

Experimental measurements are commonly used to validate numerical simulations. Yet, experimental measurements and numerical simulations can also be more intimately combined to generate an improved picture that makes a compromise between numerical modeling and experimental measurements. In this view, data-assimilation aims at establishing a mathematically well-defined procedure to take advantage of the knowledge of few experimental measurements to complete or correct an incomplete or imprecise numerical model. Data assimilation historically emerged in the field of meteorological forecasts (Lorenc [1986], Liu et al. [2008]), where the lack of accurate models, uncertain initial and boundary conditions yield poor predictions. Several data-assimilation algorithms were designed that can be classified into two categories: direct-adjoint optimization techniques, also called 3D/4D Var, where correction terms or boundary conditions in the numerical model are tuned to minimize the measurement mismatch between the experiment and the numerical solution Le Dimet and Talagrand [1986], and ensemble based-techniques, where a state estimate is corrected using the measurement mismatch Evensen [2009]. In other areas of fluid mechanics, the use of these kind of techniques has been increasing in the past years. The motivations for this were multifold: estimation of initial or inlet boundary conditions in open flows Gronskis et al. [2013], Mons et al. [2016], interpolation of velocity fields between a sequence of images Heitz and Mémin [2010], Yang et al. [2015], identification of pollutant release in urban area Mons et al. [2017b] and even as a theoretical tool to investigate aspects of the decay of large scales in homogeneous isotropic turbulence Mons et al. [2014]. In aerodynamics, the recovery of time-averaged quantities is in many applications sufficient. For example, in industrial wind-tunnels, most of the measurements consist in time-averaged pressure distributions and forces and from a numerical point of view, engineers commonly solve Reynolds-Averaged Navier-Stokes equations with turbulence models to predict the mean features of the flow around a body. In so far, a time-averaged approach is more suited and the data assimilation problem should rather be considered in the framework of inverse problems (see Foures et al. [2014] for the direct-adjoint optimization approach and Iglesias et al. [2013], Kato and Obayashi [2013] for the ensemble-based framework). Concerning the direct-adjoint optimization approach, Foures et al. [2014] started with a low-Reynolds cylinder flow ($Re = O(10^2)$)

exhibiting vortex shedding. They tuned a volume-force (modeling the force associated to the Reynolds-stress) acting in the steady Navier-Stokes equations such that its corresponding solution best matches velocity measurements, mimicking a real experimental situation where such measurements are provided by a Particle Image Velocimetry setup. The same procedure was applied at a higher Reynolds number ($Re = O(10^4)$) by Symon et al. [2017] on a flow around an idealized airfoil. This time, additional difficulties related to the well-posedness of the steady Navier-Stokes equations at such high Reynolds number had to be faced. They overcame this problem by performing several optimization steps for increasing Reynolds numbers, which can be costly. Another possible approach at high Reynolds-numbers is to consider turbulence models. This was, for example, employed by Li et al. [2017], where a set of model-related coefficients (in their work, RANS-$k-\omega$) have been tuned such that the solution of the model is as close as possible to some given partial higher-fidelity-data. Yet, such an approach is strongly constrained by the structure of the model and does not allow a high-enough flexibility to adjust the model to regions exhibiting different behaviors. Concerning the ensemble approach, Kato and Obayashi [2013], Kato et al. [2015] generated an ensemble by varying abritrarily the coefficients of the Spalart-Allmaras model and a Kalman filtering method than led to the tuning of optimal values of these coefficients with respect to given measurements. The difficulty here is again that the Reynolds stresses need to comply with the structure of the Spalart-Allmaras model (in particular the Boussineq approximation) which is known to be very constraining. Duraisamy et al. [2019] also employed data-assimilation techniques to tune production terms in RANS models to mean-flow data obtained by DNS or experiments. They then used the assimilated eddy-viscosity quantities to design new turbulence models with deep-learning techniques.

In this opening chapter, we extend the time-averaged direct-adjoint optimization approach introduced by Foures et al. [2014], Symon et al. [2017] to higher Reynolds numbers by using as a baseline model RANS equations combined with turbulence models instead of the steady (laminar) Navier-Stokes equations. Contrarily to Li et al. [2017], the tuning parameters are space-dependent functions related to a correction of the modelled Reynolds-stress tensor. In the present work, we choose the Spalart-Allmaras model (SA) Spalart and Allmaras [1994] for its numerical simplicity, even though other models could also be used. Within this model, we explore two correction possibilities.

The first consists in a volume-force acting in the momentum equations. This force is supposed to stand for a correction of the viscous force induced by the eddy-viscosity produced by the SA model. This therefore allows the turbulent stresses not to be constrained by the Boussinesq assumption, which is known to be only well adapted in shear dominated regions. Since this correction is not attached to the turbulence model itself, this approach may be considered as an extension of Foures et al. [2014] and Symon et al. [2017], the difference being the consideration of a background eddy-viscosity that evolves with the data-assimilation process according to a turbulence model (in this work, SA). This modeling is better suited for turbulent flows. For example, the uncorrected solution (for which the correction term is null), which is needed to initialize the optimization algorithm, now corresponds to the RANS-SA solution. Such an initial solution may easily be obtained with standard algorithms and is already much closer to the targeted mean-flow than an initial condition obtained with a laminar model. The assimilation procedure therefore only needs to compensate the small imperfections of the initial turbulence model, while using a laminar model for turbulent flows implies compensating for the whole discrepancy between the laminar steady solution and the targeted mean-flow.

The second correction possibility is a source term acting on the SA equation that governs the eddy-viscosity. This time the control parameter acts within the turbulence model, correcting the production and dissipation terms of the used turbulence model. It is worth-mentioning that a similar technique was employed by Singh and Duraisamy [2016] and Parish and Duraisamy [2017] to identify correction terms in turbulence models: they tuned a space-dependent function modulating the strength of the turbulent-viscosity production term. More recently, this very same procedure was used by He et al. [2019b] to assimilate some PIV measurements. We believe however that an additive source term in the equation governing the eddy viscosity is more general and robust.

This chapter is organized as follows. First (§2), we will describe the physical configuration of interest, by showing the reference solution (obtained by DNS) and the corresponding (uncorrected) RANS-SA solution, pointing out the differences between them, motivating the need for data-assimilation. This configuration consists of a rounded backward-facing step (Dandois et al. [2007]), for which the SA model provides a bad prediction of the re-circulation length. Then (§2), we will introduce the two models for data-assimilation mentioned above, together with the direct-adjoint optimization procedure to tune the correction fields. Finally (§4), we will present the data-assimilation results on the backward-facing step configuration, by comparing the assimilated flowfields obtained with the two approaches to the true mean-flow. We will also assess the performance of the reconstruction as a function of the sparsity of the used velocity measurements.

2 Flow Configuration and numerical solutions

Before introducing the mathematical models and the formalism for data-assimilation, we briefly introduce the physical configuration on which we will apply this technique, together with the reference (DNS) and *a priori* (RANS-SA) solutions. We will in particular highlight and discuss the errors of the (RANS-SA) solution, motivating the data-assimilation procedure.

The chosen configuration, is a rounded Backward-Facing Step (BFS) at Reynolds number $Re = 28275$, based on the height of the step and on the inflow velocity. Those quantities will also be used to make all variables dimensionless in the following. This configuration was considered in Dandois et al. [2007] and consists of a smooth expansion, generating a re-circulation region that detaches at $x_d = 0.53$ and reattaches at $x_r = 3.93$, those positions being measured after the start of the expansion. The geometry of the rounded step is given by the equation:

$$y(x) = [\sin(a\pi x) - a\pi x]/(2\pi), \quad 0 \leq x \leq 2/a \tag{2.1}$$

where $a = 0.703$. As detailed in Dandois et al. [2007], the Direct Numerical Simulation (DNS) was fed at the inlet with fluctuations generated from a time-dependent zero-pressure-gradient turbulence simulation (see Lund [1998]). This allows to describe a situation where the incoming boundary layer is fully turbulent. The resulting mean-flow can be seen in Figure 2.1. We now turn our attention to the solution given by the RANS equations coupled with the Spalart-Allmaras model. Here we consider the incompressible Navier-Stokes equations, describing the velocity $\tilde{\mathbf{u}}$ and pressure $\tilde{p}$ (the $\tilde{\cdot}$ notation referring to quantities obtained by a model), together with an extra equation governing an

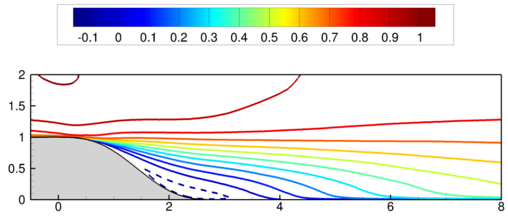

Figure 2.1: DNS solution: streamwise mean-velocity field. Dashed lines represent negative values

eddy-viscosity-related quantity $\tilde{\nu}$:

$$\begin{aligned}
\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} + \nabla \tilde{p} - \nabla \cdot ((\nu + \nu_t(\tilde{\nu}))\nabla_s \tilde{\mathbf{u}}) &= \mathbf{0} && (2.2)\\
\nabla \cdot \tilde{\mathbf{u}} &= 0 && (2.3)\\
\tilde{\mathbf{u}} \cdot \nabla \tilde{\nu} - \nabla \cdot (\eta(\tilde{\nu})\nabla \tilde{\nu}) &= P(\tilde{\nu}, \nabla \tilde{\mathbf{u}}) + D(\tilde{\nu}, \nabla \tilde{\mathbf{u}}) + C(\nabla \tilde{\nu}) && (2.4)\\
&= s(\tilde{\nu}, \nabla \tilde{\nu}, \nabla \tilde{\mathbf{u}}),
\end{aligned}$$

where $\eta(\tilde{\nu})$ is the diffusivity of $\tilde{\nu}$, $P(\tilde{\nu}, \nabla \tilde{\mathbf{u}})$ the production term, $D(\tilde{\nu}, \nabla \tilde{\mathbf{u}})$ the destruction term and $C(\nabla \tilde{\nu})$ the cross diffusion term, all of them being lumped into a single source term $s(\tilde{\nu}, \nabla \tilde{\nu}, \nabla \tilde{\mathbf{u}})$. More information about the original baseline model (especially the functions $\nu_t(\tilde{\nu})$, $P(\tilde{\nu}, \nabla \tilde{\mathbf{u}})$, $D(\tilde{\nu}, \nabla \tilde{\mathbf{u}})$ and $C(\tilde{\nu}, \nabla \tilde{\mathbf{u}})$) and the chosen implementation can be found in Crivellini et al. [2013] and Oliver [2008], as well as in Appendix A. To properly solve this set of equations, we need to specify boundary conditions, especially the inflow velocity profile (a sketch of the flow domain, with the remaining boundary conditions is provided in figure 2.2). We will consider here that the inflow boundary layer is known and mimics the inflow conditions of the DNS (this constraint can in principle be removed by tuning or assimilating this inflow boundary condition with respect to some measurements more downstream; yet, this is not the focus of this work and has not been studied). One of the issues here concerns the value of $\tilde{\nu}$, which should reflect the value of the shear-stress in the DNS. For this variable, we will keep it simple and impose a constant value $(\tilde{\nu}/\nu)_\infty$ with a fast-decay to zero very close to the wall. We have made a parametric study for values of this quantity ranging from $(\tilde{\nu}/\nu)_\infty = 3$ (recommended for a fully-turbulent boundary layer, see Allmaras et al. [2012]) to $(\tilde{\nu}/\nu)_\infty = 0$. The stream-wise velocity, the normalized eddy-viscosity ν_t/ν, the wall-pressure and friction coefficients of the RANS-SA solution for $(\tilde{\nu}/\nu)_\infty = 3$ are respectively shown in Figures 6.1 (a,b,c,d). The dependency of the incoming boundary layer (at the step location) as a function of $(\tilde{\nu}/\nu)_\infty$ is shown in Figure 6.1 (e), indicating a small sensitivity with respect to the choice of $(\tilde{\nu}/\nu)_\infty$. In table 2.1, we can see that the re-circulation length is strongly overestimated by the RANS-SA model and that the choice of $(\tilde{\nu}/\nu)_\infty$ strongly affects these overestimated results. This overestimation may be due to the fact that the Spalart-Allmaras model has been initially tuned for higher Reynolds numbers and the gradients are not sufficiently strong in the present shear-layer to strengthen the eddy-viscosity (and reduce the re-circulation length). For this reason, throughout this work, we fix that value to be $(\tilde{\nu}/\nu)_\infty = 3$, which is consistent with the fact that the DNS itself was performed with a fully-turbulent incoming boundary-layer. As mentioned above, the RANS-SA solution over-predicts the reattachment length by 70%. This also reflects in a strong mismatch in the pressure and friction coefficients (Figure 2.3 (c,d)). These discrepancies motivate the use of data-

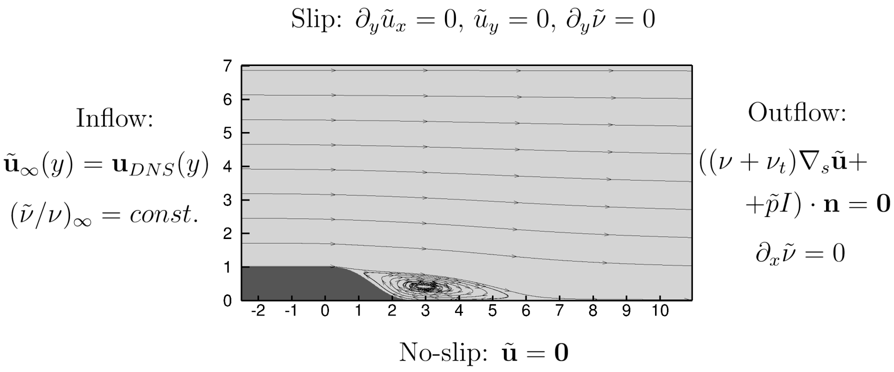

Figure 2.2: Sketch of computational domain, together with its boundary donditions.

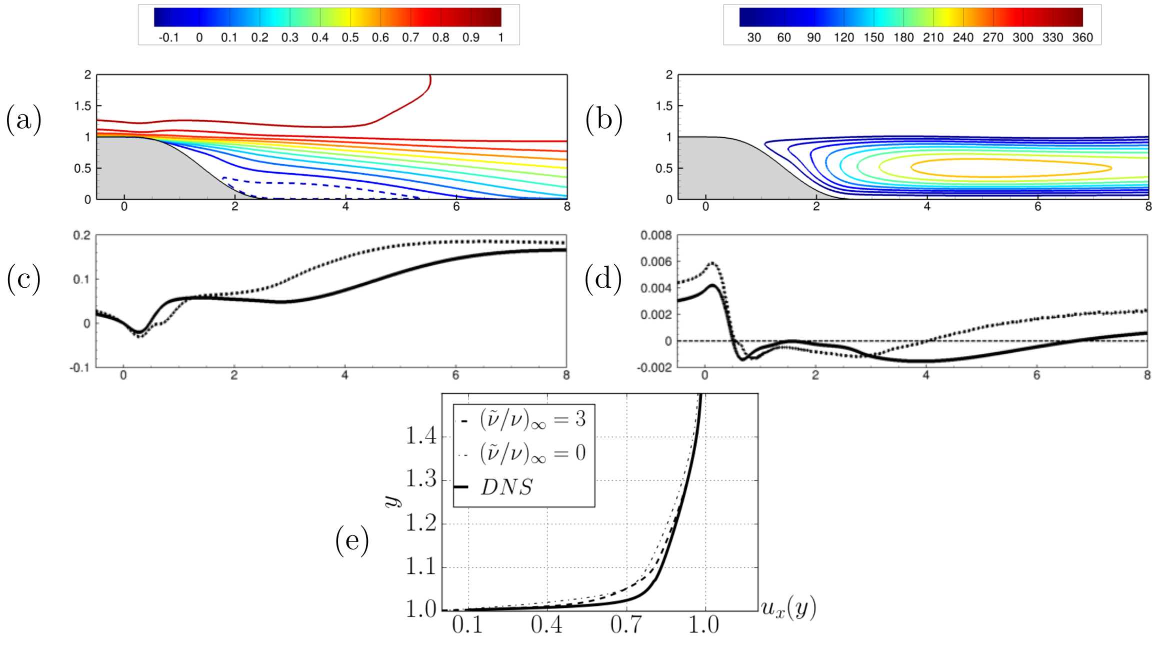

Figure 2.3: RANS-SA solution: (a): streamwise velocity field for solution with $(\tilde{\nu}/\nu)_\infty = 3$, (b): corresponding normalized eddy viscosity ν_t/ν, (c): pressure coefficient, (d): friction coefficient, (e): streamwise velocity at $x = 0$ for RANS-SA solution with different values of $(\tilde{\nu}/\nu)_\infty$ and for time-averaged DNS-solution. In figures (c,d), the dotted line corresponds to the time-averaged DNS solution, and the solid-line to the RANS-SA solution.

assimilation techniques to improve the recovery of the right flow features. In this study, we will consider the time-averaged DNS as the reference solution from which measurements will be extracted to feed the assimilation process.

Table 2.1: Comparison of separation and reattachment points for RANS, LES, and DNS simulations for the smooth backward-facing step, $Re = 28275$

	$(\tilde{\nu}/\nu)_\infty$	x_s	x_r
	3	0.50	6.67
RANS - SA	$3 \cdot 10^{-6}$	0.46	6.87
(present studies)	0	0.40	7.35
DNS Dandois et al. [2007]	–	0.53	3.93
LES Dandois et al. [2007]	–	0.51	3.96

3 Models and Data assimilation procedure

In this section, we will discuss the models we consider for the data-assimilation process. We recall that the time-averaged $\overline{\mathbf{u}}$ and pressure $\overline{p}$ quantities are governed by the RANS equations:

$$\overline{\mathbf{u}} \cdot \nabla \overline{\mathbf{u}} + \nabla \overline{p} - \nabla \cdot (\nu \nabla_s \overline{\mathbf{u}}) = -\nabla \cdot \left(\overline{\mathbf{u}' \otimes \mathbf{u}'} \right) = -\nabla \cdot \underline{\tau} = \overline{\mathbf{f}}, \tag{2.5}$$

where $\underline{\tau}$ is the Reynolds-stress tensor, and the notation $\mathbf{u} \otimes \mathbf{v}$ stands for the tensor $[u_i v_j]_{i,j=1,\dots}$. This tensor contains all the nonlinear interactions of the non-resolved fluctuations $\mathbf{u}'$. For this reason, those equations are not closed and several hypotheses / approximations / models exist in the literature to close the problem, that is provide a closure $\underline{\tau} = \underline{\tau}(\overline{\mathbf{u}}, \overline{p},$ + supplementary variables). It is thus clear that quantities related to this closure can legitimately be used as tuning functions for our data-assimilation procedure. In the following, we will describe two possible tuning functions, both of them relying on the SA-turbulence model. The first one corresponds to adding a volume-force in the momentum equations and the second one to adding a volume source in the $\tilde{\nu}$ equation.

3.1 Model 1: source term in momentum equation

The first model considered here is given by a volume force on the momentum equations. However, using the Navier-Stokes equations with solely a molecular viscosity as a base-model and compensating the measurement mismatch through only a volume-force (as done in Foures et al. [2014]) may be prone of difficulties at high Reynolds numbers for several reasons: Symon et al. [2017] had to perform several expensive optimization steps at increasing Reynolds numbers to deal with a case at $Re = O(10^4)$, the correction term may be very large since the uncorrected laminar steady state is usually very far from the turbulent mean, and in some cases, this laminar state may not even exist or there might be lots of such steady states Sipp et al. [2010]. All these reasons render the assimilation process difficult and expensive. In any case, it seems a good idea to have an assimilation process with an initial guess that is easy to obtain and not too far from the targeted solution so that only a small discrepancy needs to be compensated by the optimization process.

In order to keep the physical idea of a volume-force in the momentum equations and to improve the robustness of the assimilation process, we choose to use existing turbulence models in connection with a momentum force in the RANS equations to serve as a model for the optimization. This means that the Reynolds stress tensor will be divided into a modeled and a tuned part:

$$\overline{\mathbf{f}} = -\nabla \cdot \underline{\tau} \approx -\nabla \cdot (2/3\kappa I - \nu_t \nabla_s \tilde{\mathbf{u}}) + \tilde{\mathbf{f}}_{\mathbf{x}}. \tag{2.6}$$

The first term on the right-hand-side of equation (2.6) corresponds to the Boussinesq term, stating that part of the Reynolds-stresses can be modeled by an eddy-viscosity ν_t (the term involving the turbulent kinetic energy $\kappa = \overline{\mathbf{u}' \cdot \mathbf{u}'}/2$ being incorporated in the pressure). The second term is the correction term $\tilde{\mathbf{f}}_{\mathbf{x}}$, which can in principle account for non-Boussinesq effects or other modelling errors. With this approximation, instead of (6.1), we consider the following forced momentum equations:

$$\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} + \nabla \underbrace{(\tilde{q} + 2/3\kappa)}_{\tilde{p}} - \nabla \cdot ((\nu + \nu_t) \nabla_s \tilde{\mathbf{u}}) = \tilde{\mathbf{f}}_{\mathbf{x}}, \tag{2.7}$$

where the turbulent kinetic energy $\tilde{p} = \tilde{q} + 2\kappa/3$ has been lumped intro the pressure term. The eddy viscosity ν_t present in these equations is given by a turbulence model, here the Spalart-Allmaras model $\nu_t(\tilde{\nu})$ where $\tilde{\nu}$ is governed by (2.4).

3.2 Model 2: source term in $\tilde{\nu}$ equation

The second tuning function considered corresponds to a source term in the equation governing $\tilde{\nu}$. Equation (2.4) is replaced by

$$\tilde{\mathbf{u}} \cdot \nabla \tilde{\nu} - \nabla \cdot (\eta(\tilde{\nu}) \nabla \tilde{\nu}) = s(\tilde{\nu}, \nabla \tilde{\nu}, \nabla \tilde{\mathbf{u}}) + \tilde{f}_{\tilde{\nu}}. \tag{2.8}$$

The correction $\tilde{f}_{\tilde{\nu}}$ only has an indirect effect on the velocity $\tilde{\mathbf{u}}$ and pressure $\tilde{p}$ fields, since $\tilde{f}_{\tilde{\nu}}$ only modifies the source-production-diffusion term s acting on $\tilde{\nu}$, which then modifies the eddy-viscosity ν_t that appears in the momentum equations (6.1). The Reynolds stress tensor therefore remains constraint by the Boussinesq assumption, which is nevertheless a valid assumption in a great variety of flows, especially in shear-dominated flows. This model is sufficiently flexible to allow for example correction of a wrong free shear-layer expansion ratio, as is the case in figure 2.3 (a). Model 2 is nevertheless far more constrained than model 1, since with model 1 we are able to encompass all corrections obtained by model 2, while the reverse is not true. Yet, this should not just be seen as drawback for model 2, since the associated solution will be forced to remain in a subspace that might in some cases be more physical than the solution obtained with a much more free model, such as model 1. One has to keep in mind that depending on the number of measurements that are available, the inverse problem that aims at determining the tuning function that yields the best measurement match is more or less severely under-determined: for example, in presence of sparse (resp. many) measurements, it might be better to pick model 2 (resp. model 1) that is more constrained (resp. freer).

3.3 Data Assimilation procedure

In this section, we describe the data-assimilation procedure employed. Let $\overline{\mathbf{m}}$ be a set of higher-fidelity or experimental measurements that correspond to information extracted from the flow and $\mathcal{M}(\cdot)$ the measurement operator that allows to extract the corresponding measure from our simulation result $\tilde{\mathbf{u}}$. In this chapter, since we are dealing only with velocity measurements, this operator will act on the velocity field, $\tilde{\mathbf{u}}$, yielding $\tilde{\mathbf{m}} = \mathcal{M}(\tilde{\mathbf{u}}) \in M$ (other measure operators will be considered in the next chapter). M is the measurement space, whose norm is given, generically, by $||\cdot||_M$. The data-assimilation problem can now be cast into an optimization one, for which we want to tune the forcing terms (either $\tilde{\mathbf{f}}_{\mathbf{x}}$ or $\tilde{f}_{\tilde{\nu}}$) such that the cost functional:

$$\mathcal{J}(\tilde{\mathbf{q}}) = \frac{1}{2}||\mathcal{M}(\tilde{\mathbf{q}}) - \overline{\mathbf{m}}||^2_M \tag{2.9}$$

is minimal, with the following constraints:

$$\begin{aligned}
\tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}} + \nabla\tilde{p} - \nabla\cdot((\nu+\nu_t(\tilde{\nu}))\nabla_s\tilde{\mathbf{u}}) &= \tilde{\mathbf{f}}_{\mathbf{x}} && (2.10)\\
\nabla\cdot\tilde{\mathbf{u}} &= 0 && (2.11)\\
\tilde{\mathbf{u}}\cdot\nabla\tilde{\nu} - \nabla\cdot(\eta(\tilde{\nu})\nabla\tilde{\nu}) - s(\tilde{\nu},\nabla\tilde{\nu},\nabla\tilde{\mathbf{u}}) &= \tilde{f}_{\tilde{\nu}}, && (2.12)
\end{aligned}$$

where here, for conciseness, we consider a general framework in which both $\tilde{\mathbf{f}}_{\mathbf{x}}$ and $\tilde{f}_{\tilde{\nu}}$ could simultaneously be considered in the optimization process. To solve this optimization problem, following Foures et al. [2014], Parish and Duraisamy [2017], Singh and Duraisamy [2016] or Mons et al. [2014], we choose a gradient-based algorithm. For the descent algorithm, we pick the low memory version of the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm, which is far more efficient than a simple steepest gradient-descent one. It requires in particular the computation of the cost functional gradient with respect to the forcing parameter, $\nabla_{\tilde{\mathbf{f}}_{\mathbf{x}}}\mathcal{J}$ or $\nabla_{\tilde{f}_{\tilde{\nu}}}\mathcal{J}$. For this, we resort to a Lagrangian formalism, in which the state is augmented with a set of Lagrange multiplier (or adjoint) variables $(\tilde{\mathbf{u}}^\dagger, \tilde{p}^\dagger, \tilde{\nu}^\dagger)$ and we look for critical points of the Lagrangian functional:

$$\begin{aligned}
\mathcal{L}([\tilde{\mathbf{u}},\tilde{p},\tilde{\nu}],[\tilde{\mathbf{u}}^\dagger,\tilde{p}^\dagger,\tilde{\nu}^\dagger],[\tilde{\mathbf{f}}_{\mathbf{x}},\tilde{f}_{\tilde{\nu}}]) =& \mathcal{J}(\tilde{\mathbf{u}})\\
&+\left(\tilde{\mathbf{u}}^\dagger, \tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}} + \nabla p - \nabla\cdot((\nu+\nu_t(\tilde{\nu}))\nabla_s\tilde{\mathbf{u}}) - \tilde{\mathbf{f}}_{\mathbf{x}}\right)_\Omega\\
&+\left(\tilde{p}^\dagger, \nabla\cdot\tilde{\mathbf{u}}\right)_\Omega\\
&+\left(\tilde{\nu}^\dagger, \tilde{\mathbf{u}}\cdot\nabla\tilde{\nu} - \nabla\cdot(\eta(\tilde{\nu})\nabla\tilde{\nu}) - s(\tilde{\nu},\nabla\tilde{\nu},\nabla\tilde{\mathbf{u}}) - \tilde{f}_{\tilde{\nu}}\right)_\Omega,
\end{aligned} \tag{2.13}$$

where $(\mathbf{v}_1,\mathbf{v}_2)_\Omega = \int_\Omega \mathbf{v}_1\cdot\mathbf{v}_2 d\Omega$ represents the inner product related to the classical (real valued) energy-norm. Setting the variation of this Lagrangian with respect to the adjoint variables to zero, yields the governing equations (2.10). Setting its variation with respect to the direct variables, we obtain the adjoint equations:

$$\begin{aligned}
\tilde{\mathbf{u}}^\dagger\cdot(\nabla\tilde{\mathbf{u}})^T - \tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}}^\dagger - \nabla\cdot((\nu+\nu_t)\nabla_s\tilde{\mathbf{u}}^\dagger) - \nabla\tilde{p}^\dagger & && (2.14)\\
+\tilde{\nu}^\dagger\nabla\tilde{\nu} + \nabla\cdot(\tilde{\nu}^\dagger\partial_{\nabla\tilde{\mathbf{u}}}s) &= -\frac{\partial\mathcal{M}}{\partial\tilde{\mathbf{u}}}(\mathcal{M}(\tilde{\mathbf{u}}) - \overline{\mathbf{m}}) &&\\
\nabla\cdot\tilde{\mathbf{u}}^\dagger &= 0 && (2.15)\\
-\tilde{\mathbf{u}}\cdot\nabla\tilde{\nu}^\dagger - \nabla\cdot(\eta\nabla\tilde{\nu}^\dagger) + (\partial_{\tilde{\nu}}\eta)\nabla\tilde{\nu}^\dagger\cdot\nabla\tilde{\nu} + (\partial_{\tilde{\nu}}\nu_t)\nabla\tilde{\mathbf{u}}^\dagger : \nabla_s\tilde{\mathbf{u}} & &&\\
-(\partial_{\tilde{\nu}}s)\tilde{\nu}^\dagger + \nabla\cdot(\tilde{\nu}^\dagger\partial_{\nabla\tilde{\nu}}s) &= 0, && (2.16)
\end{aligned}$$

where the boundary conditions for this adjoint variables are given by:

$$No-slip/Inflow: \qquad \tilde{\mathbf{u}}^{\dagger} = \mathbf{0}, \quad \tilde{\nu}^{\dagger} = 0 \tag{2.17}$$

$$Outflow: \quad \mathbf{n}\cdot\left(\tilde{\mathbf{u}}^{\dagger}\otimes\tilde{\mathbf{u}} + (\nu+\nu_t)\nabla_s\tilde{\mathbf{u}}^{\dagger} + I\tilde{p}^{\dagger} + (\partial_{\nabla\tilde{\mathbf{u}}}s)\tilde{\nu}^{\dagger}\right) = \mathbf{0} \tag{2.18}$$

$$\mathbf{n}\cdot\left(\tilde{\mathbf{u}}\tilde{\nu}^{\dagger} + \eta\nabla\tilde{\nu}^{\dagger} - \tilde{\nu}^{\dagger}(\partial_{\tilde{\nu}}\eta)\nabla\tilde{\nu} - (\partial_{\nabla\tilde{\nu}}s)\tilde{\nu}^{\dagger}\right) = 0 \tag{2.19}$$

$$Slip: \qquad \partial_y\tilde{\mathbf{u}}^{\dagger}_x = 0, \quad \tilde{\mathbf{u}}^{\dagger}_y = 0, \quad \eta\partial_y\tilde{\nu}^{\dagger} = (\partial_{\partial_y\tilde{\nu}}s)\tilde{\nu}^{\dagger}. \tag{2.20}$$

Although here we derive the strong equations for the adjoint problem, together with its boundary conditions, in practice we use the discrete adjoint, meaning the transpose of the Jacobian matrix. This procedure is fully justified in a Finite-Element Framework, since it provides an actual FEM discretization of the adjoint equations. For further details on this issue, see Appendix B 4. Taking now the variation of the Lagrangian with respect to the forcing terms (either $\tilde{\mathbf{f}}_{\mathbf{x}}$ or $\tilde{f}_{\tilde{\nu}}$), we have:

$$\left(\frac{\partial\mathcal{L}}{\partial\tilde{\mathbf{f}}_{\mathbf{x}}}, \delta\tilde{\mathbf{f}}_{\mathbf{x}}\right)_{\Omega} = -\left(\tilde{\mathbf{u}}^{\dagger}, \delta\tilde{\mathbf{f}}_{\mathbf{x}}\right)_{\Omega} \;\rightarrow\; \nabla_{\tilde{\mathbf{f}}_{\mathbf{x}}}\mathcal{J} = -\tilde{\mathbf{u}}^{\dagger} \tag{2.21}$$

$$\left(\frac{\partial\mathcal{L}}{\partial\tilde{f}_{\tilde{\nu}}}, \delta\tilde{f}_{\tilde{\nu}}\right)_{\Omega} = -\left(\tilde{\nu}^{\dagger}, \delta\tilde{f}_{\tilde{\nu}}\right)_{\Omega} \;\rightarrow\; \nabla_{\tilde{f}_{\tilde{\nu}}}\mathcal{J} = -\tilde{\nu}^{\dagger}. \tag{2.22}$$

At this point we have all the necessary ingredients for the BFGS optimization algorithm. A brief description of the iterative procedure is depicted in figure 2.4. Roughly speaking, this algorithm is more elaborate than usual gradient-descent methods, since the descent direction is obtained by inverting an approximation of the Hessian matrix, constructed from previous gradient evaluations. This procedure allows one to capture the second-derivative behaviour of the cost functional $\mathcal{J}$, which, formally provides a second-order convergence of the optimization (further details on its implementation with a finite-element solver in Appendix B). It is interesting to notice that, in the case where we use the $\tilde{\mathbf{f}}_{\mathbf{x}}$ tuning parameter, even if the original gradient is modified by the Hessian matrix, the finally chosen descent direction is still divergence-free, just as in Foures et al. [2014] for the adjoint velocity field $\tilde{\mathbf{u}}^{\dagger}$. However, since we now have a background eddy-viscosity, the total force in 2.6 that acts on the momentum equations is not anymore divergence-free. We remark as well that, at any iteration of the optimization procedure, the forcing term $\tilde{\mathbf{f}}_{\mathbf{x}}$ may have a non zero projection onto the space of all forces compatible with a Boussinesq hypothesis. This means that the eddy-viscosity $\nu_t = \nu_t(\tilde{\nu})$ does not hold anymore the 'total amount' of eddy-viscosity, a part of which can as well be contained in $\tilde{\mathbf{f}}_{\mathbf{x}}$. This slightly hides the physical meaning of the sole term ν_t.

4 Results

We will now analyze the reconstruction results of the backward-facing step flow for two different measurement operators. Firstly, we will consider the complete mean velocity field as known in order to evaluate the ability of the two models to accurately reconstruct the velocity, pressure and Reynolds stress forcing. This is interesting because the nature of the forcing terms in the two models is different, which induces different constraints on the reconstructed fields, so that the ensemble of reachable velocity-pressure fields is different between the two models. For example, the mean-flow generated by the DNS /

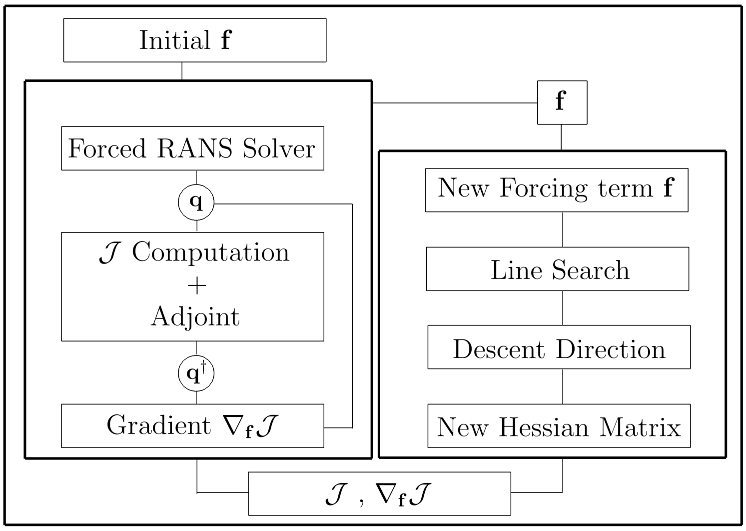

Figure 2.4: Sketch of the BFGS algorithm for a general problem, having $\mathbf{f}$ as a control parameter and $\mathbf{q}$ as state. The adjoint variable is denoted by $\mathbf{q}^\dagger$.

LES solutions might in fact not be fully compatible with the Boussinesq approximation of model 2. Secondly, we will use only few point-wise velocity measurements such as those given by a Pitot probe or a hot-wire measurement that is moved along a line to extract cross-stream velocity profiles. Such a situation is actually much more realistic when considering industrial wind-tunnels, since optical measurements such as Particle-Image-Velocimetry (PIV) measurements remain difficult to implement in such facilities. We actually expect different behaviors between model 1 and model 2 when only few measurements are available since the reconstruction constraints are different between the two models. Foures et al. [2014] and Symon et al. [2017] have already explored point-wise measurements in the case of a laminar model at lower Reynolds numbers: in particular, Symon et al. [2017] showed that a measurement operator involving a spatial averaging of the measure over a finite region improved the quality of the reconstruction by smoothing the gradients of the reconstructed velocity field in the vicinity of the measurements. We will apply here a similar approach.

4.1 Complete velocity field measurements

We consider the measure as being the whole exact (mean) velocity field $\overline{\mathbf{m}} = \overline{\mathbf{u}}$. This field was extracted from the DNS results of Dandois et al. [2007] and then interpolated on our mesh, which is coarser than the one from the DNS. The measure is thus defined on the same mesh as the solution itself. For this reason, the measurement space M ends up being the velocity space itself. Its norm also simplifies into $||\cdot||_M = ||\cdot||_\Omega$, the norm induced by the kinetic energy inner product (without the pressure component). We then have:

$$-\left(\frac{\partial \mathcal{M}}{\partial \tilde{\mathbf{u}}}\right)^\dagger (\mathcal{M}(\tilde{\mathbf{u}}) - \overline{\mathbf{m}}) = -(\tilde{\mathbf{u}} - \overline{\mathbf{u}}) \tag{2.23}$$

Under those conditions, the optimization of the tuning fields $\tilde{\mathbf{f}}_\mathbf{x}$ and $\tilde{f}_{\tilde{\nu}}$ proceeds as shown in figure 2.5. Model 1 with the optimization of the volume-force $\tilde{\mathbf{f}}_\mathbf{x}$ manages to reconstruct the time-averaged DNS solution extremely accurately due to the large flexibility of this

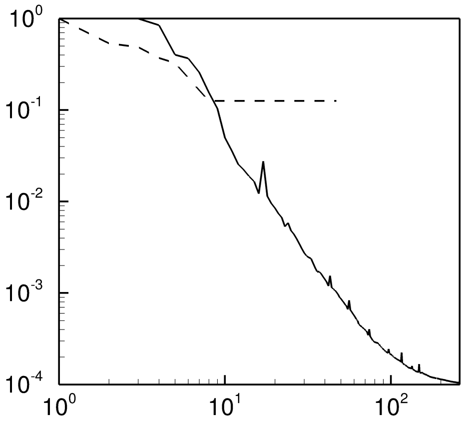

Figure 2.5: Convergence history $\mathcal{J}_n/\mathcal{J}_0$ as a function of the iteration count n when optimizing $\tilde{\mathbf{f}}_\mathbf{x}$ (solid line) or $\tilde{f}_{\tilde{\nu}}$ (dashed line).

model: the decrease in the cost functional $\mathcal{J}_n/\mathcal{J}_0$ is around 10^{-4}, which is very good. Note also that this decrease is achieved in less than 200 iterations, which is much quicker than the 10^3 iterations required by Foures et al. [2014] to achieve similar results. This gain is due to the choice of far-superior l-BFGS optimization algorithm, which provides an approximation of the Hessian matrix based on previous gradient evaluations, improving the descent direction estimation. When using model 2, the optimization procedure reaches a plateau after $n \approx 10$ iterations. In this case, the reduction of the cost functional is of the order of $\mathcal{J}_n/\mathcal{J}_0 \approx 0.13$. This is less accurate than with model 1. Yet, it is coherent with the fact that the control parameter $\tilde{f}_{\tilde{\nu}}$ only indirectly affects the velocity field $\tilde{\mathbf{u}}$ (through ν_t) , leading to a far more constrained reconstructed solution. As for the reconstructed velocity fields, we can clearly see that model 1 (figure 2.6) manages to almost exactly match the streamlines of the reference solution (in dashed lines), whereas the model 2 only approximately achieves this goal. For example, the resulting reattachement point is around $x_r = 4.64$ for model 2, the reference solution of the DNS being $x_r = 3.93$ while the uncorrected RANS-SA gave $x_r = 6.67$. Yet, the overall picture provided by model 2 is still acceptable and definitely improves the base-line results provided by RANS-SA model.

It is interesting to note that for model 1, which involves the tuning of the momentum forcing $\tilde{\mathbf{f}}_\mathbf{x}$, the final eddy-viscosity values (figure 2.7 (a)) have decreased with respect to the baseline values (the maximum is now around $\nu_t/\nu \approx 140$, compared to 240 for the uncorrected RANS-SA solution). This indicates that the baseline SA turbulence model induces errors that may be compensated most efficiently by replacing part of the Reynolds-stresses induced by the SA model by a general unconstrained forcing. In model 2, where the source term $\tilde{f}_{\tilde{\nu}}$ is optimized for in the $\tilde{\nu}$ equation, the final eddy-vicosity ν_t values have strongly increased for the bubble to become shorter (figure 2.7 (b)), its maximum value being here $\nu_t/\nu \approx 330$. We can also see that the Reynolds-stress induced forcing produced in the momentum equations (either stemming from the viscous contribution alone for model 2 or combined with the volume-force $\tilde{\mathbf{f}}_\mathbf{x}$ for model 1) are similar and located in the vicinity of the separation point (figures 2.7 (c,d)) and agree with the overall characteristics of the actual forcing term $\overline{\mathbf{f}}$ (figure 2.7 (f)). We however remark that, for model 1, those two forcing terms are not exactly the same (compare figure (c) and (e)), even though their corresponding velocity fields nearly exactly match. The same

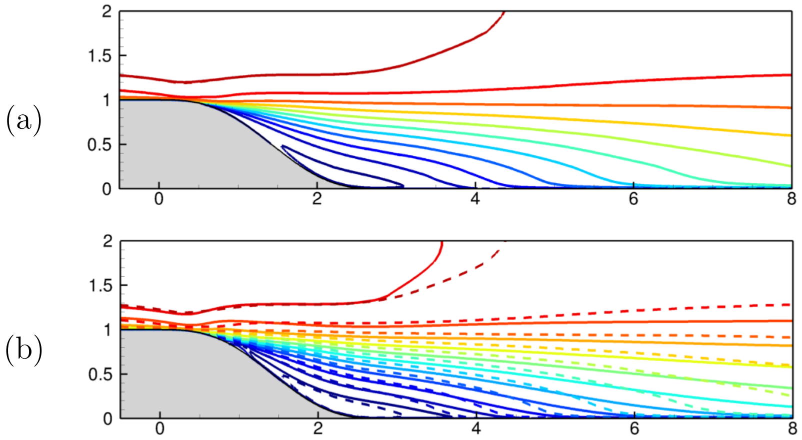

Figure 2.6: Complete velocity field assimilation: optimal velocity field when optimizing (a) $\tilde{\mathbf{f}}_{\mathbf{x}}$ (model 1) and (b) $\tilde{f}_{\tilde{\nu}}$ (model 2).

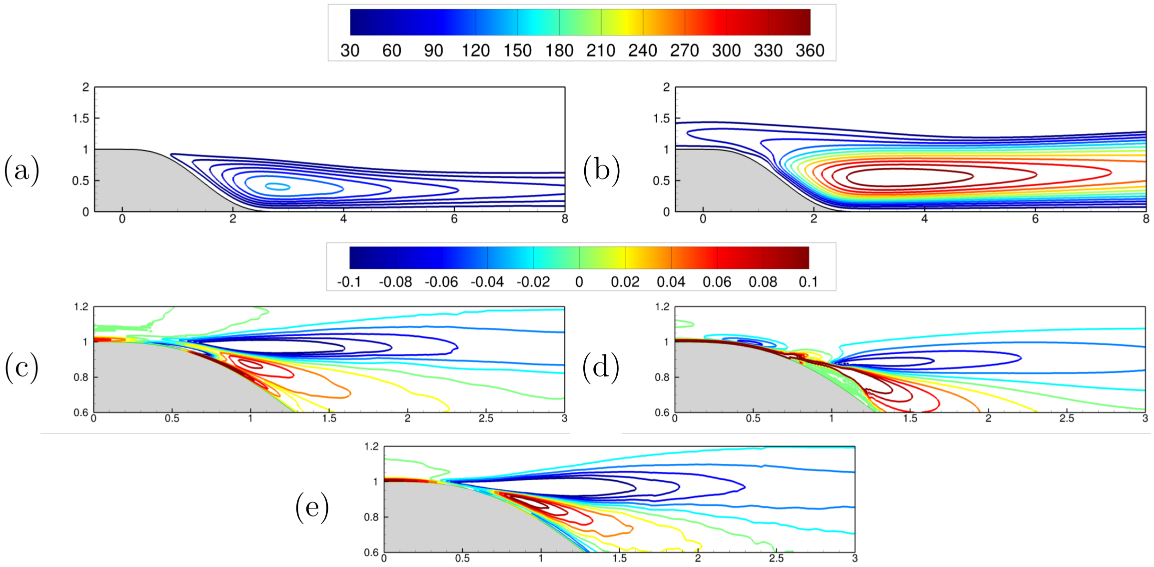

Figure 2.7: Complete velocity field assimilation: ν_t/ν for model 1 (optimization of $\tilde{\mathbf{f}}_{\mathbf{x}}$) (a) and model 2 (optimization of $\tilde{f}_{\tilde{\nu}}$) (b). Total stream-wise momentum forcing for (c) volume-force correction ($\nabla \cdot (\nu_t \nabla_s \tilde{\mathbf{u}}) + \tilde{\mathbf{f}}_{\mathbf{x}}$), (d) SA-correction ($\nabla \cdot (\nu_t \nabla_s \tilde{\mathbf{u}})$) and (e) exact one $\overline{\mathbf{f}}$.

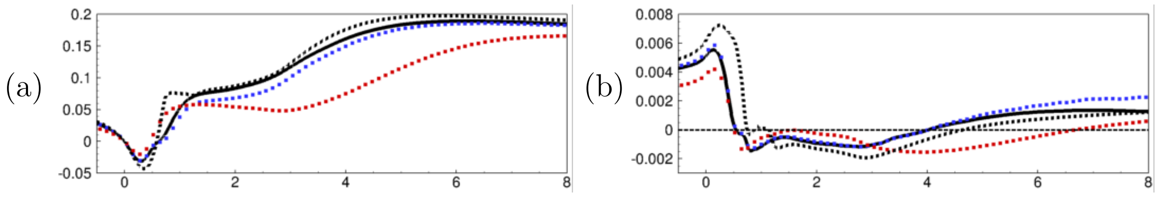

Figure 2.8: Complete velocity field assimilation: (a) wall-pressure and (b) wall-friction. Red dots correspond to the RANS-SA baseline solution, blue dots to the time-averaged DNS, continuous black line to model 1 and black dashed line to model 2.

issue was reported by Foures et al. [2014], who explained this discrepancy by the fact that the adjoint fields (and thus the forcing term $\tilde{\mathbf{f}}_\mathbf{x}$) are divergence-free by construction if no pressure measurements are considered in the cost functional. This constraint ends up preventing the pressure to be fully reconstructed, since we have no information on the potential part of the forcing term. It is interesting to notice however that, despite this fact, the wall-pressure (figure 2.8 (a)) is well represented. To explain this, we turn our attention to the exact RANS equations (3.1) and the approximated ones (2.7). We can see that subtracting the two equations yields:

$$\nabla(\tilde{p}-\overline{p})+(\tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}}-\overline{\mathbf{u}}\cdot\nabla\overline{\mathbf{u}})-\nabla\cdot(\nu\nabla_s(\tilde{\mathbf{u}}-\overline{\mathbf{u}}))=\tilde{\mathbf{f}}_\mathbf{x}+\nabla\cdot(\nu_t\nabla_s\tilde{\mathbf{u}})-\overline{\mathbf{f}}. \tag{2.24}$$

To evaluate the difference between the approximated wall-pressure $\tilde{p}$ and the exact one $\overline{p}$, we evaluate this equation on the tangential direction at the solid wall. From the no-slip condition (satisfied both by the approximation and by the DNS), we have $\tilde{\mathbf{u}}=\overline{\mathbf{u}}=\overline{\mathbf{f}}=\mathbf{0}$. We also have that the viscous term $\nabla\cdot(\nu_t\nabla_s\tilde{\mathbf{u}})$ has to vanish at the wall, since $\nabla\cdot(\nu_t\nabla_s\tilde{\mathbf{u}})=\nu_t\nabla\cdot\nabla_s\tilde{\mathbf{u}}+\nabla\nu_t\cdot\nabla_s\tilde{\mathbf{u}}$, the first term being zero since $\nu_t=0$ at the walls and the second being also zero since, with the Spalart-Allmaras model $\nabla\nu_t=\partial_{\tilde{\nu}}\nu_t\nabla\tilde{\nu}$, which is zero because $\partial_{\tilde{\nu}}\nu_t|_{\tilde{\nu}=0}=0$. Moreover, by imposing the divergence-free property of both the reference velocity field and the approximated one and projecting the result onto the wall-tangential direction $\mathbf{t}$ we have:

$$\partial_\mathbf{t}(\tilde{p}-\overline{p})-\nu\Delta(\tilde{\mathbf{u}}_\mathbf{t}-\overline{\mathbf{u}}_\mathbf{t})=\partial_\mathbf{t}(\tilde{p}-\overline{p})-\nu\partial_{\mathbf{nn}}(\tilde{\mathbf{u}}_\mathbf{t}-\overline{\mathbf{u}}_\mathbf{t})=\tilde{\mathbf{f}}_\mathbf{x}\cdot\mathbf{t}. \tag{2.25}$$

Since, at any iteration of the optimization procedure, the forcing term $\tilde{\mathbf{f}}_\mathbf{x}$ is a combination of adjoint variables that satisfy the no-slip boundary condition, we have $\tilde{\mathbf{f}}_\mathbf{x}=\mathbf{0}$ at the wall. If we apply the approximation $\tilde{\mathbf{u}}\approx\overline{\mathbf{u}}$ (true at the end of successful optimization process), we have that the variation of the pressure difference along the wall is approximately null $\partial_\mathbf{t}(\tilde{p}-\overline{p})\approx 0$, justifying the good estimation of the wall-pressure. We remark that this approximation is valid only if, at the wall, the reconstructed tangential velocity field exhibits normal-derivatives that are close to the reference ones $\nu\partial_{\mathbf{nn}}\tilde{\mathbf{u}}_\mathbf{t}\approx\nu\partial_{\mathbf{nn}}\overline{\mathbf{u}}_\mathbf{t}$. This condition is therefore only met if the quality of the reconstruction is very good near the wall, as is the case here with complete velocity measurements. Also, as mentioned before, this approximation for the pressure obtained at the wall does not hold anymore in the bulk of the flow (see figure 2.9). This is due to the fact that the evolution of $\tilde{\mathbf{f}}_\mathbf{x}$ is restricted to a solenoidal vector space. In order to study this error in the pressure field, following Foures et al. [2014], (2.24) shows that the pressure difference is related to the optimal forcing $\tilde{\mathbf{f}}_\mathbf{x}$, the actual DNS forcing $\overline{\mathbf{f}}$ and the eddy-viscosity term through:

$$\nabla(\tilde{p}-\overline{p})\approx\tilde{\mathbf{f}}_\mathbf{x}+\nabla\cdot(\nu_t\nabla_s\tilde{\mathbf{u}})-\overline{\mathbf{f}}. \tag{2.26}$$

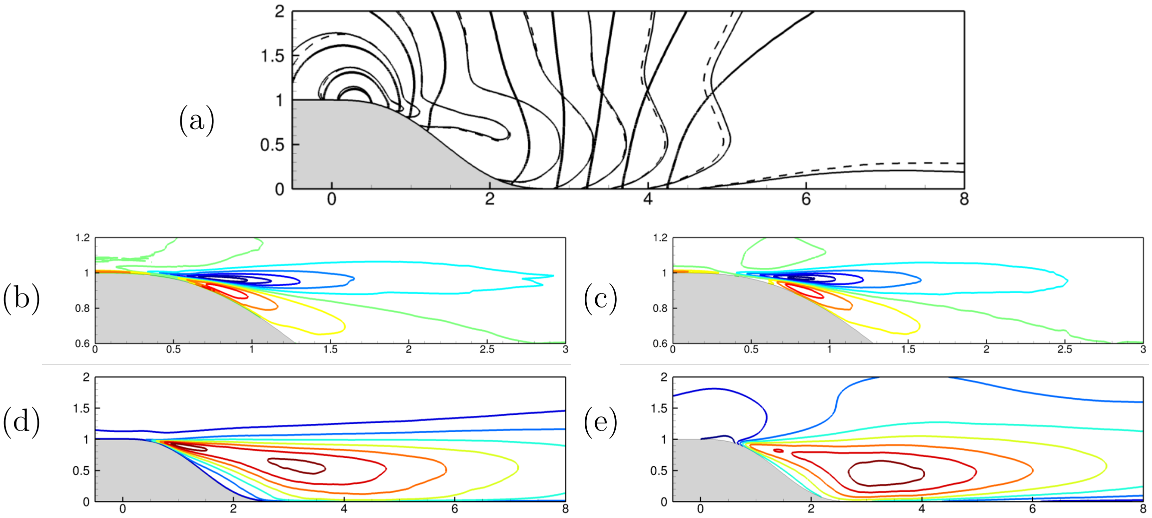

Figure 2.9: Assimilated pressure (a) $\tilde{p}$ (thick solid lines), reconstructed one $\tilde{p} - \tilde{\phi} + \overline{\phi}$ (thin solid lines) and DNS one $\overline{p}$ (dashed lines). The divergence-free forcing terms $\tilde{\mathbf{f}}_{\mathbf{x}} + \tilde{\mathbf{f}}_s$ coming from the model is also compared with the one coming from the DNS $\overline{\mathbf{f}}_s$ (c). The turbulent kinetic energy $2/3\kappa$ (computed from DNS statistics) (b) and the combination of the potentials $-\tilde{\phi} + \overline{\phi}$ (c).

The forcing term $\tilde{\mathbf{f}}_{\mathbf{x}}$ is solenoidal and the pressure error $\tilde{p} - \overline{p}$ appears as a scalar potential. In order to identify in this equation solenoidal and scalar-potential parts, we perform a Helmholtz decomposition of the DNS forcing term $\overline{\mathbf{f}}$ and of the eddy-viscosity term related to turbulence model:

$$\overline{\mathbf{f}} = \overline{\mathbf{f}}_s + \nabla\overline{\phi} \tag{2.27}$$

$$\nabla \cdot (\nu_t \nabla_s \tilde{\mathbf{u}}) = \tilde{\mathbf{f}}_s + \nabla\tilde{\phi} \tag{2.28}$$

where $\overline{\mathbf{f}}_s$ and $\tilde{\mathbf{f}}_s$ are the solenoidal functions and $\overline{\phi}$ and $\tilde{\phi}$ are the scalar potentials. We therefore have:

$$\overline{p} - \overline{\phi} \approx \tilde{p} - \tilde{\phi} \tag{2.29}$$

$$\overline{\mathbf{f}}_s \approx \tilde{\mathbf{f}}_{\mathbf{x}} + \tilde{\mathbf{f}}_s, \tag{2.30}$$

where on the left-hand-side we have only terms stemming from the reference itself (thus the notation $\bar{\cdot}$) while those on the right-hand-side are linked to the optimal reconstruction (thus the notation $\tilde{\cdot}$). If we plot the reconstructed pressure $\tilde{p} + \overline{\phi} - \tilde{\phi}$, we can see that it matches very well with the actual pressure $\overline{p}$ (figure 2.9). The "divergence-free" component of the DNS forcing term $\overline{\mathbf{f}}_s$ compares now much better with the optimal one $\tilde{\mathbf{f}}_{\mathbf{x}} + \tilde{\mathbf{f}}_s$ (figure 2.9(b,c)). Also, in (2.7), we have seen that $\tilde{p} = \tilde{q} + 2/3\kappa$, so that $\tilde{\phi} - \overline{\phi} \approx 2/3\kappa$ if we assume that $\tilde{q} = \overline{p}$, which is also verified (see figure 2.9 (d,e)).

4.2 Sparse velocity measurements

4.2.1 Sparse velocity measurement operator

Now that we have performed and analysed the complete-velocity field measure, we turn our attention to a more realistic case where fewer (but still) velocity measurements are considered in the assimilation process. Instead of a 'spatial average' of the velocity field around the measurement location, we will consider a measure operator under the form:

$$\mathcal{M}(\tilde{\mathbf{u}}) = \{\tilde{\mathbf{u}}(\mathbf{x}_m^i)\}_{i=1,\cdots,N}. \tag{2.31}$$

Here $\mathbf{x}_m^i$ are N points where the velocity measurements are performed. Typically, these N measurements may correspond to cross-stream profiles of N_y points at N_x stream-wise stations, such that $N = N_x \times N_y$. Experimentally, this would correspond a Pitot probe that is moved in the cross-stream direction at N_x different stations, a rather common measure in experimental fluid dynamics. This way, the cost functional reads:

$$\mathcal{J} = \frac{1}{2}\sum_{i=1}^{N}(\tilde{\mathbf{u}}(\mathbf{x}_m^i) - \overline{\mathbf{m}}_i)^2, \tag{2.32}$$

where $\overline{\mathbf{m}}_i = \overline{\mathbf{u}}(\mathbf{x}_m^i)$. With this definition of the measurement operator, the right-hand-side of the adjoint equations becomes:

$$-\left(\frac{\partial \mathcal{M}}{\partial \tilde{\mathbf{u}}}\right)^{\dagger}(\mathcal{M}(\tilde{\mathbf{u}}) - \overline{\mathbf{m}}) = -\sum_{i=1}^{N}(\tilde{\mathbf{u}}(\mathbf{x}_m^i) - \overline{\mathbf{m}}_i)\delta_{\mathbf{x}_m^i} \tag{2.33}$$

where $\delta_{\mathbf{x}_m}$ is the Dirac-mass centered at $\mathbf{x}_m$. With this right-hand-side defined, we can perform the data-assimilation for both models.

4.2.2 Reconstruction results using model 1 and model 2

Results of the reconstruction using model 1 and model 2 are presented in figure 2.10 for $N_x = 3$ and $N_y = 10$. From this figure, we can see that the volume-force model produces a solution which, although it respects the measurements ($\mathcal{J}_n/\mathcal{J}_0 \approx 10^{-6}$), does not correspond well to the reference $\overline{\mathbf{u}}$. We can see from figure 2.10 (a) that the iso-contours of the assimilated velocity are not smooth and non physical. For example, the wall-friction shown in figure 2.10 (e) shows a (rather large) secondary re-circulation bubbles, a feature that does not exist neither in the RANS-SA solution nor in the DNS. We can explain these observations by noticing that the optimal forcing term $\tilde{\mathbf{f}}_{\mathbf{x}}$ is given by a linear combination of adjoint fields with right-hand-sides given by (2.33). Since those equations are excited on the momentum equation by a sum of Dirac masses, the velocity-adjoint variables are 'peaked' around the measurements, which contaminates the solution, leading to those undesired oscillations. On the other hand, this does not occur in the $\tilde{\mathbf{f}}_{\mathbf{x}}$-data-assimilation, since the $\tilde{\nu}$-adjoint is not forced directly by the Dirac masses, leaving the gradients smooth. Furthermore, one can see that the result for model 2 is very similar to the complete velocity field assimilation. If we look at the wall-pressure or wall-friction plot (figure 2.10 (d,e)), we can see that the results in the sparse case present the same general features as in the complete case, including the (rather small) re-circulation bubble presented before. Model 2 therefore appears to be more robust than model 1 in the sense that, seemingly, for any type of velocity measurements (dense, sparse), the results of the reconstruction will be almost the same. This is confirmed when looking at the global errors:

$$e_{\Omega} = \sqrt{\frac{1}{V_{\Omega}}\int_{\Omega}(\tilde{\mathbf{u}} - \overline{\mathbf{u}})^2 d\Omega}, \tag{2.34}$$

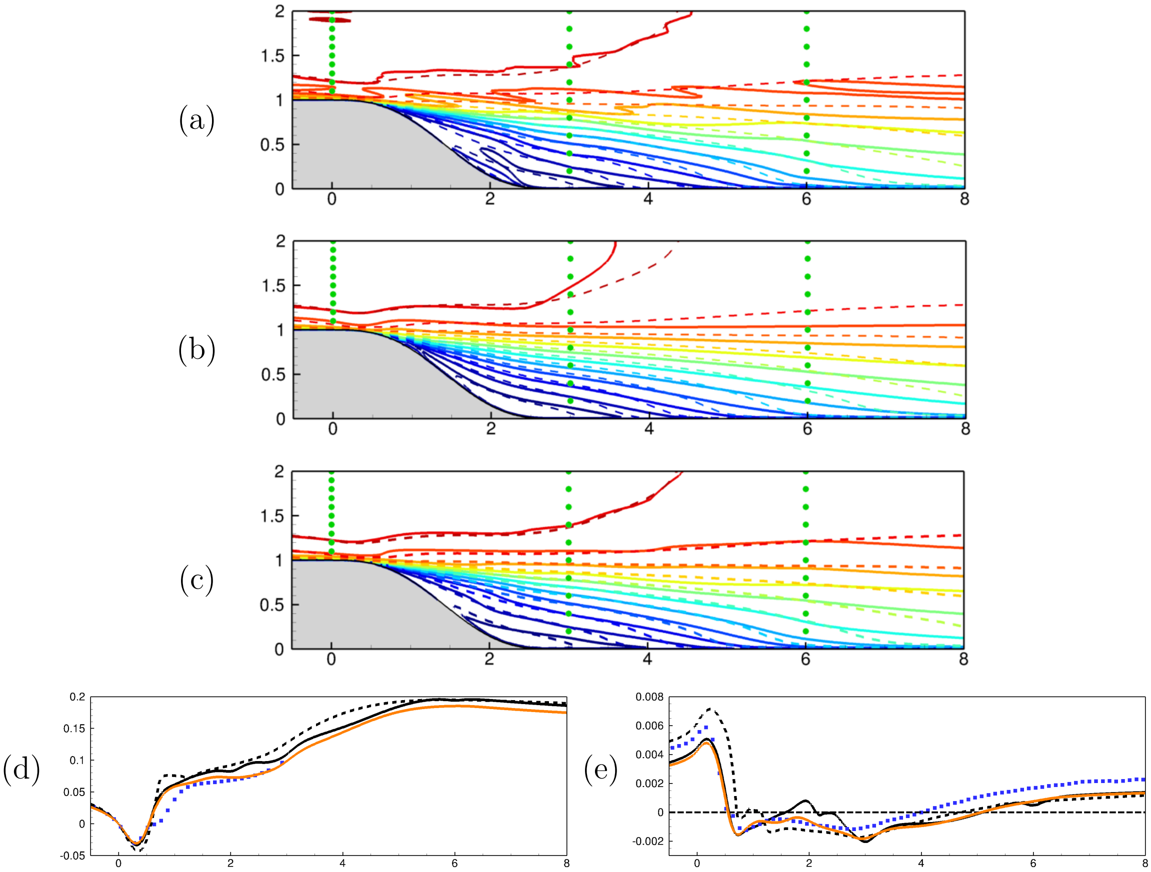

Figure 2.10: Sparse-velocity measurement assimilation: assimilated velocity field $\tilde{\mathbf{u}}$ (solid lines) together with reference solution $\overline{\mathbf{u}}$ (dashed lines) for (a) $\tilde{\mathbf{f}}_{\mathbf{x}}$, (b) $\tilde{f}_{\tilde{\nu}}$ and (c) $\tilde{\mathbf{f}}_{\mathbf{x}}$ with penalized gradients. Green dots correspond to measurement locations $\mathbf{x}_m^i$. Wall-pressure (d) and wall-friction (e) are plotted as well. Blue dots correspond to the DNS solution, continuous black line to model 1, black dashed line to model 2 and orange line to model 1 with γ-penalisation.

Table 2.2: Global results for partial-velocity data-assimilations

		e_Ω			$\mathcal{J}_n/\mathcal{J}_0$		
		$\mathbf{f_x}$	$\mathbf{f_x}$, $\gamma^2 = 1$	$f_{\tilde{\nu}}$	$\mathbf{f_x}$	$\mathbf{f_x}$, $\gamma^2 = 1$	$f_{\tilde{\nu}}$
Total	Measure	$\sim 10^{-3}$	–	0.035	$\sim 10^{-4}$	–	12.3 %
$N_x = 3$	$N_y = 5$	0.051	0.042	0.036	$\sim 10^{-6}$	0.1 %	8.7 %
	$N_y = 10$	0.037	0.031	0.037	$\sim 10^{-6}$	0.3 %	8.7 %
	$N_y = 20$	0.030	0.028	0.037	$\sim 10^{-6}$	0.4 %	9.0 %
$N_x = 6$	$N_y = 5$	0.042	0.035	0.037	$\sim 10^{-6}$	0.1 %	11.9 %
	$N_y = 10$	0.032	0.030	0.037	$\sim 10^{-6}$	0.5 %	9.4 %
	$N_y = 20$	0.027	0.027	0.038	$\sim 10^{-6}$	1.1 %	13.4 %
RANS-SA			0.094			–	

which are provided in table 2.2. Indeed, for several values of N_x, N_y, the optimization based on model 2 leads to very similar values for e_Ω. whereas for model 1, those values become smaller if more measurements are considered. From table 2.2 we can also see that, in the cases where few measurements are provided ($N_y = 5$, for example), the error made with model 2 is smaller than with model 1. This means that, whenever we have only few measurements, it is preferable to use the more robust model 2 than the more flexible model 1 as a means to compensate the lack of knowledge of the actual reference flow.

4.2.3 Optimization with gradient penalization of model 1

However, in the case of denser velocity measurements, we may still prefer to use the more flexible model 1, since we have shown a more drastic reduction in e_Ω for model 1 than for model 2. Yet, we need to improve the reconstruction associated to model 1 by trying to damp the oscillations seen in the solution (see figure 2.10 (a)). For this, we suggest to penalize the forcing term by modifying the cost functional $\tilde{\mathcal{J}}$ following:

$$\tilde{\mathcal{J}} = \frac{\mathcal{J}}{\mathcal{J}_0} + \frac{\gamma^2}{2} \int_\Omega |\nabla \tilde{\mathbf{f}}|^2 \, d\Omega \tag{2.35}$$

We remark that, by dividing the original cost functional by its value at the first iteration, we are normalizing the penalization with a unity value, making the penalization term independent off the measurement and only dependent on the parameter γ^2. This parameter should not taken too small (the penalization needs to be effective) and not too big (the optimization should not get stuck in the vicinity of the RANS-SA solution). It tells the algorithm to favor a solution that is smoother, since the tuning term $\tilde{\mathbf{f}}$ in the momentum-equations will exhibit smaller gradient values. This is a physically sound constraint since we know that the Reynolds-stress forcing is very smooth. The results for this modified data-assimilation procedure with model 1 are shown in figure 2.10 (c) for $\gamma^2 = 1$. We can see that the resulting solution is smoother and still matches well with the measurements ($\mathcal{J}_n/\mathcal{J}_0 \lessapprox 1\%$), suggesting that the kind of penalisation employed (here derivatives of $\tilde{\mathbf{f}}_\mathbf{x}$) still allows enough freedom for the tuning field to match the measurements, while constraining the solution in a smooth subspace. Furthermore, except for the very sparse setup ($N_x = 3$, $N_y = 5$), the penalized algorithm even manages to give a better reconstruction of the flow away from the measurement points (see table 2.2). Indeed, for all such test-cases, the penalized model 1 solution always manages to outperform both the non-penalized model 1 reconstruction and the model 2 reconstruction.

4.3 Understanding the 'rigidity' of model 2 through observability Gramian analysis

This section is devoted to the understanding of the rigidity of model 2, pointed out in the last section. One possible way to understand this observation is to examine the linearized optimization problem, say, around the RANS-SA solution. In this linear framework, a variation of the control parameter $\delta\tilde{\mathbf{f}}$ (here, either $\tilde{\mathbf{f}}_\mathbf{x}$ or $\tilde{f}_{\tilde{\nu}}$) induces a variation of the state variable $\tilde{\mathbf{q}} = (\tilde{\mathbf{u}}, \tilde{p}, \tilde{\nu})$:

$$\delta\tilde{\mathbf{q}} = \left(\frac{\partial \mathcal{R}}{\partial \tilde{\mathbf{q}}}\right)^{-1} P\delta\tilde{\mathbf{f}} = (\partial_{\tilde{\mathbf{q}}}\mathcal{R})^{-1} P\delta\tilde{\mathbf{f}} \tag{2.36}$$

where $\mathcal{R}(\tilde{\mathbf{q}})$ is the nonlinear residual composing the RANS-SA equations and P being an operator that transforms the variation of the control parameter into the actual forcing term of the Jacobian. This variation in the state induces a variation on the observable data according to:

$$\delta\tilde{\mathbf{m}} = \frac{\partial\mathcal{M}}{\partial\tilde{\mathbf{q}}}\delta\tilde{\mathbf{q}} = (\partial_{\tilde{\mathbf{q}}}\mathcal{M})\delta\tilde{\mathbf{q}} = (\partial_{\tilde{\mathbf{q}}}\mathcal{M})(\partial_{\tilde{\mathbf{q}}}\mathcal{R})^{-1}P\delta\tilde{\mathbf{f}} = \mathcal{A}\delta\tilde{\mathbf{f}} \tag{2.37}$$

To be able to identify which forcing terms induce the most energetic variation on the measure we may optimize the gain:

$$G(\delta\tilde{\mathbf{f}}, \delta\tilde{\mathbf{m}}) = \frac{||\delta\tilde{\mathbf{m}}||_M^2}{||\delta\tilde{\mathbf{f}}||^2} = \frac{\langle\delta\tilde{\mathbf{m}}, \delta\tilde{\mathbf{m}}\rangle_M}{(\delta\tilde{\mathbf{f}}, \delta\tilde{\mathbf{f}})_\Omega}, \tag{2.38}$$

which is equivalent to solving the Singular-Value decomposition of the operator $\mathcal{A} = (\partial_{\tilde{\mathbf{q}}}\mathcal{M})(\partial_{\tilde{\mathbf{q}}}\mathcal{R})^{-1}P$:

$$\mathcal{A}^\dagger\mathcal{A}\delta\tilde{\mathbf{f}}_i = \lambda_i^2\delta\tilde{\mathbf{f}}_i, \tag{2.39}$$

where λ_i^2 are the ranked squared singular-values ($\lambda_i^2 \geq \lambda_{i+1}^2$). The values λ_i represent the measurement variations along the unit optimal measurement directions $\delta\tilde{\mathbf{m}}_i$ ($=\lambda_i^{-1}\mathcal{A}\delta\tilde{\mathbf{f}}_i$) associated to the unit optimal forcing directions $\delta\tilde{\mathbf{f}}_i$. If all those singular-values are of the same order, we can say that any measurement $\delta\tilde{\mathbf{m}}$ is equally reachable (or, more suitably, 'observable'). If we have a strong separation of singular-values, that is $\lambda_0^2 \gg \lambda_i^2$ for some index i, we have measurement states that cannot be reached ('observed') with the chosen model. This can be seen by considering that, for a given $\delta\tilde{\mathbf{f}}$, the resulting perturbation on the measure is given by:

$$\delta\tilde{\mathbf{m}} = \mathcal{A}\delta\tilde{\mathbf{f}} = \mathcal{A}\sum_i \alpha_i\delta\tilde{\mathbf{f}}_i = \sum_i \alpha_i\lambda_i\delta\tilde{\mathbf{m}}_i = \lambda_0\sum_i \alpha_i\frac{\lambda_i}{\lambda_0}\delta\tilde{\mathbf{m}}_i. \tag{2.40}$$

Since $\delta\tilde{\mathbf{m}}_i$ is unit norm, this shows that the amplitude α_i required for the forcing term to reach measurement $\delta\tilde{\mathbf{m}}_i$ scales as λ_0/λ_i, which can be very large, and therefore not achievable. In figure 2.11, we have plot the quantities $(\lambda_i/\lambda_0)^2$ for model 1 and model 2, taking as measurement operator the full velocity field $\mathcal{M}(\tilde{\mathbf{q}}) = \tilde{\mathbf{u}}$. We can see that the $\tilde{\mathbf{f}}_\mathbf{x}$ model 1 is much more flexible than the $\tilde{f}_{\tilde{\nu}}$ model 2 so that more measurement states $\delta\tilde{\mathbf{m}}$ can be reached. As for the $\tilde{f}_{\tilde{\nu}}$ model 2, we can see that the separation is very strong and increases very rapidly for $i > 1$.

5 Conclusion

In this chapter, we have introduced two models for the reconstruction of mean-flow features from mean-velocity measurements. The first one consists in the perturbation of the momentum equations by a volume-force in a similar fashion as in Foures et al. [2014], the difference being that we have a background eddy-viscosity turbulence model, making the numerical procedure more robust and well-posed for high Reynolds number flows. The second one consists in the correction of the equation governing the eddy-viscosity (here the Spalart-Allmaras model). The physical configuration on which those models are tested is a smooth backward-facing step. We can show that, if the whole velocity field is considered as measured (like in a PIV setup), the $\tilde{\mathbf{f}}_\mathbf{x}$ model 1 produces a solution that matches exactly

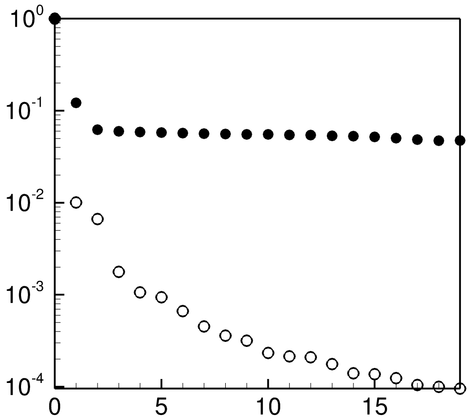

Figure 2.11: Separation of Singular-Values $(\lambda_i/\lambda_0)^2$ for the $\tilde{\mathbf{f}}_{\mathbf{x}}$ model 1 (continuous lines) and the $\tilde{f}_{\tilde{\nu}}$ model 2 (dashed lines), computed around the RANS-SA solution.

the reference, showing the high flexibility of thi model. Moreover, we could also show that the optimal forcing parameter matches almost perfectly the projection of the real forcing term onto the space of solenoidal functions, in a similar way as Foures et al. [2014]. As for the $\tilde{f}_{\tilde{\nu}}$ model 2, we showed that, if a complete velocity field measurement is available, it is not capable to reproduce exactly the reference state. The model is too 'rigid' and lots of measurements are not accessible with this model. If now only sparse point-wise velocity measurements are available, we show that the first model, despite its high flexibility to exactly recover the measurements, may lead to noisy / unphysical state reconstructions. For this reason, an additional penalisation on the gradients of the tuning field needed to be considered to obtain smooth solutions. As for the second model, it turned out to be more rigid in the sense that, independently of the quantity of available measurements, the optimal state was approximately the same. This behaviour could be analysed by performing a singular value decomposition of the linear operator between the forcing space and the measurement space. We showed that the controlability/observability of the second model was actually restricted to a rather small subspace. However, such a rigidity may be helpful whenever the number of measurements is very low since its optimal solution is almost independent of the number of measurements. Hence, model 2 can provide in such cases a better reconstruction of the flow field than the much more flexible model 1.

Chapter 3

Wall-measure assimilation for RANS equations

1 Introduction

In the second chapter, we dealt with the case where the information or data to be assimilated was considered to be sparse velocity-field information. The measure of this kind of quantity from experiments can be easy to do if not many probes are used, employing, for example, hot-wire or Pitot probes. However if one needs a great number of velocity measurements, more refined experimental techniques may be needed (for example, Particle Image Velocimetry, Raffel et al. [2018]), the drawback in those cases being the requirement of more sophisticated experimental apparatus. Instead, in this chapter, we evaluate other quantities that are more easily obtained experimentally, for example, the wall-pressure and friction. We remark that those measures are more indirect, regarding the velocity field itself, and their assimilation's performance needs to be investigated. It is interesting to note that, from a practical point of view, the assimilation of, say, the wall-friction, can be useful even in cases where sufficiently bulk velocity field information is available since it provides an easy way to express the localized behavior of the flow near the wall, making the optimization procedure to impose its correct behavior there.

From a mathematical point of view, in an incompressible framework, the derivation of the gradient for this cost functional is less straightforward than the velocity measure one. The reasons for that are related with a seemingly over-constrained set of boundary conditions for the adjoint problem. Since this issue is not so much discussed in the literature (even if similar physical problems have already been addressed, with the correct adjoint equations, see Bewley and Protas [2004]), we will dedicate the next section to the exposure of the mathematical problem and to the proper derivation of the adjoint equations on a laminar case, together with a few numerical results on the transitional flow around a circular cylinder (as in Foures et al. [2014]), where the control parameter will be a volume-force acting on momentum equations. Then, we will apply this procedure in the case of the turbulent Backward-Facing Step (BFS), already introduced in the previous paragraph, using for this purpose the two models discussed there. As we will see, this case is much more challenging since, depending on the model used and on the quantity we measure, the model will provide a solution that corresponds more or less to the reference. For instance, if the wall-friction assimilation using model 2, the optimal solution will produce a re-circulation bubble that is much thinner than the actual one, over-estimating the pressure. This chapter can thus be viewed as an extension of the first, since the overall goals and methods will essentially be the same.

2 Wall-measure for laminar RANS equations

2.1 Theoretical developments on laminar equations

In this section, we introduce the theoretical aspects of the adjoint derivation for the wall-information measure operator. For the sake of clarity, we choose to place ourselves first in the case where the flow is governed by the laminar RANS equations. We remark that all technicalities related to the adjoint boundary conditions, that will allow us to assimilate wall-information are present in the same manner if we consider a turbulence model. The exact RANS equations, starting point of the analysis, are given by:

$$\overline{\mathbf{u}} \cdot \nabla \overline{\mathbf{u}} + \nabla \overline{p} - \nu \Delta \overline{\mathbf{u}} = -\nabla \cdot \left(\overline{\mathbf{u}' \otimes \mathbf{u}'} \right) = \overline{\mathbf{f}}, \quad \nabla \cdot \overline{\mathbf{u}} = 0, \tag{3.1}$$

which differ from the steady solution equation by the presence of the right-hand side term $\overline{\mathbf{f}}$, coming from the nonlinear interactions of the fluctuation $\mathbf{u}'$. Following Foures et al. [2014], Symon et al. [2017], we will use this right-hand side term as tuning parameter for our assimilation procedure. Its goal will be then to find the forcing term $\tilde{\mathbf{f}}$, satisfying:

$$\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} + \nabla \tilde{p} - \nu \Delta \tilde{\mathbf{u}} = \tilde{\mathbf{f}}, \quad \nabla \cdot \tilde{\mathbf{u}} = 0, \tag{3.2}$$

such that the following cost functional:

$$\mathcal{J}_\Gamma(\tilde{\mathbf{u}}, \tilde{p}) = \frac{1}{2} || \mathcal{M}_\Gamma(\nabla \tilde{\mathbf{u}}, \tilde{p}) - \overline{\mathbf{m}}_\Gamma ||^2_{M_\Gamma} \tag{3.3}$$

is minimal. $\mathcal{M}_\Gamma$ is now a (rather general) measure operator that takes as input the information of the gradient of velocity field and pressure at (a portion of) the wall Γ and gives the measure output on the measure space M_Γ:

$$\mathcal{M}_\Gamma : (\nabla \tilde{\mathbf{u}}, \tilde{p})|_\Gamma \mapsto \mathcal{M}_\Gamma(\nabla \tilde{\mathbf{u}}, \tilde{p}) \in M_\Gamma \tag{3.4}$$

and $\overline{\mathbf{m}}_\Gamma = \mathcal{M}_\Gamma(\nabla \overline{\mathbf{u}}, \overline{p})$ represents the measure applied on the actual reference solution. We note that, since we suppose the measure operator acts on a solid wall Γ, on which the no-slip boundary conditions are assumed, it does not make sense to measure the velocity field. However, in order to measure the wall-friction, its derivatives are needed. The measure space is equipped with a norm $|| \cdot ||_{M_\Gamma}$ that will quantify the distance from the measure of our model and the actual target data to be assimilated $\mathbf{m}_\Gamma$. Just as before, in order to derive the gradient, we consider the Lagrangian:

$$\mathcal{L} = \mathcal{J}_\Gamma + (\tilde{\mathbf{u}}^\dagger, \tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} + \nabla \tilde{p} - \nu \Delta \tilde{\mathbf{u}} - \tilde{\mathbf{f}})_\Omega + (\tilde{p}^\dagger, \nabla \cdot \tilde{\mathbf{u}})_\Omega \tag{3.5}$$

The variations of this Lagrangian with respect to the adjoint variables will, again, lead to the set of direct equations (3.2). Its variations with respect to the state $(\tilde{\mathbf{u}}, \tilde{p})$ will lead to the adjoint equations, which will be slightly different from those found in the previous chapter. The reason for this, will be the nature of the measure operator, that takes as input boundary terms. As we will see, this fact will change the forcing term of the adjoint equations from a volume force to a non-homogeneous boundary condition. To properly derive them, we will proceed with the formal derivation of the Lagrangian:

$$\begin{aligned}\left(\frac{\partial\mathcal{L}}{\partial(\tilde{\mathbf{u}},\tilde{p})},\delta(\tilde{\mathbf{u}},\tilde{p})\right)_\Omega &= \left(\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p})-\mathbf{m}_\Gamma,\frac{\partial\mathcal{M}_\Gamma}{\partial(\nabla\tilde{\mathbf{u}},\tilde{p})}\cdot(\nabla\delta\tilde{\mathbf{u}},\delta\tilde{p})\right)_{M_\Gamma} \\ &+ \left(\tilde{\mathbf{u}}^\dagger,\tilde{\mathbf{u}}\cdot\nabla\delta\tilde{\mathbf{u}}+\delta\tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}}+\nabla\delta\tilde{p}-\nu\Delta\delta\tilde{\mathbf{u}}\right)_\Omega+\left(\tilde{p}^\dagger,\nabla\cdot\delta\tilde{\mathbf{u}}\right)_\Omega=0\end{aligned} \tag{3.6}$$

We remark that, for any linear operator that acts on the boundary such as:

$$A_\Gamma : u_\Gamma \mapsto A_\Gamma u \in M_\Gamma \tag{3.7}$$

we can define its adjoint as:

$$(\mathbf{m},A_\Gamma u)_{M_\Gamma}=\left(A_\Gamma^\dagger\mathbf{m},u\right)_\Gamma \tag{3.8}$$

where $(\mathbf{v}_1,\mathbf{v}_2)_\Gamma=\int_\Gamma\mathbf{v}_1\cdot\mathbf{v}_2 d\Gamma$. With this in mind, and performing the appropriated integration by parts, we can derive the adjoint equations to be:

$$\tilde{\mathbf{u}}^\dagger\cdot(\nabla\tilde{\mathbf{u}})^T-\tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}}^\dagger-\nabla\tilde{p}^\dagger-\nu\Delta\tilde{\mathbf{u}}^\dagger=0,\quad \nabla\cdot\tilde{\mathbf{u}}^\dagger=0 \tag{3.9}$$

We remark that, since the measure operator acts now on the boundary of the domain, the adjoint equations are no longer forced by a volume term. Those equations must however be completed with a set of boundary conditions. For the inflow Γ_I, outflow Γ_O and slip Γ_S boundaries the boundary conditions are exactly the ones obtained by Foures et al. [2014]. However, since now we are performing some measurements on the solid wall Γ, the classical no-slip boundary condition does not hold anymore. To re-establish those conditions there, we have to impose that all boundary terms, coming from integration by parts, vanish (having in mind that $\delta\tilde{\mathbf{u}}=\mathbf{0}$ at Γ). Those terms can be written as:

$$\begin{aligned}&\left(-\nu\mathbf{n}\otimes\tilde{\mathbf{u}}^\dagger+\left[\frac{\partial\mathcal{M}_\Gamma}{\partial\nabla\tilde{\mathbf{u}}}\right]^\dagger(\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p})-\mathbf{m}_\Gamma),\nabla\delta\tilde{\mathbf{u}}\right)_\Gamma \\ &+\left(\mathbf{n}\cdot\tilde{\mathbf{u}}^\dagger+\left[\frac{\partial\mathcal{M}_\Gamma}{\partial\tilde{p}}\right]^\dagger(\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p})-\mathbf{m}_\Gamma),\delta p\right)_\Gamma=0,\quad\forall(\delta\tilde{\mathbf{u}},\delta\tilde{p})\end{aligned} \tag{3.10}$$

At a first glance, this seems impossible since, if we take $\nabla\delta\tilde{\mathbf{u}}=0$, we see that we need to impose that the normal adjoint velocity has to be driven by the measure operator on the pressure:

$$\mathbf{n}\cdot\tilde{\mathbf{u}}^\dagger=-\left[\frac{\partial\mathcal{M}_\Gamma}{\partial\tilde{p}}\right]^\dagger(\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p})-\mathbf{m}_\Gamma) \tag{3.11}$$

Inversely, if we take $\delta\tilde{p}=0$, we have that:

$$\nu\mathbf{n}\otimes\tilde{\mathbf{u}}^\dagger=\left[\frac{\partial\mathcal{M}_\Gamma}{\partial\nabla\tilde{\mathbf{u}}}\right]^\dagger(\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p})-\mathbf{m}_\Gamma) \tag{3.12}$$

which not only demands the imposition of all the adjoint velocity on the wall but also imposes some restrictions on the measure operator itself, since the above equation is a tensor-equation. We can see then that, instead of having a tensor-like equation, we should have a scalar one, such that, together with 3.11, a complete set of boundary conditions for $\tilde{\mathbf{u}}^\dagger$ is given. To try to achieve that, we will simplify the first term of equation 3.10 using the fact that, along the wall, the test function must not change (in fact it is identically zero). This condition can be written as $\mathbf{t} \cdot \nabla \delta \tilde{\mathbf{u}} = \mathbf{0}$, where $\mathbf{t}$ is any tangent vector ($\mathbf{t} \cdot \mathbf{n} = 0$). In what follows, the choice for $\mathbf{t}$ will be fixed to be $\mathbf{t} = (-n_y, n_x)$ (in a 2D configuration). With this arbitrary choice for the tangent direction, we can rewrite this condition as:

$$\mathbf{t} \cdot \nabla \delta \tilde{u}_x = t_x \partial_x \delta \tilde{u}_x + t_y \partial_y \delta \tilde{u}_x = -n_y \partial_x \delta \tilde{u}_x + n_x \partial_y \delta \tilde{u}_x = 0$$

$$\mathbf{t} \cdot \nabla \delta \tilde{u}_y = t_x \partial_x \delta \tilde{u}_y + t_y \partial_y \delta \tilde{u}_y = -n_y \partial_x \delta \tilde{u}_y + n_x \partial_y \delta \tilde{u}_y = 0$$

For a general point on the solid wall (for which $n_x n_y \neq 0$), we can solve this set of equations for $\partial_x \delta \tilde{u}_y$ and $\partial_y \delta \tilde{u}_x$. Those extra scalar equations establish two links between the components of the tensor $\nabla \delta \tilde{\mathbf{u}}$, reducing the possible variations of it:

$$\nabla \delta \tilde{\mathbf{u}} = \begin{bmatrix} \partial_x \delta \tilde{u}_x & \partial_x \delta \tilde{u}_y \\ \partial_y \delta \tilde{u}_x & \partial_y \delta \tilde{u}_y \end{bmatrix} = \begin{bmatrix} \partial_x \delta \tilde{u}_x & n_x / n_y \partial_y \delta \tilde{u}_y \\ n_y / n_x \partial_x \delta \tilde{u}_x & \partial_y \delta \tilde{u}_y \end{bmatrix} \tag{3.13}$$

Using those conditions on equation (3.10), we now can re-derive a new equation for the boundary conditions for $\tilde{\mathbf{u}}^\dagger$, taking again $\delta \tilde{p} = 0$. Since we use now a supplementary vector-like equation ($\mathbf{t} \cdot \nabla \delta \tilde{\mathbf{u}} = \mathbf{0}$), we manage to reduce the previous equation from a tensor-like one to a vector-like one. However, we still observe contradiction with respect to the boundary condition (3.11). For this reason, at this point, we also suppose that $\nabla \cdot \delta \tilde{\mathbf{u}} = 0$ at Γ. This hypothesis may seem arbitrary, since the test function $\delta \tilde{\mathbf{u}}$ does not need, in general, to satisfy one of the volume-differential equations (continuity) of our model, specially at the boundary Γ. However, one may make this argument rigorous by changing the space of our test functions such that it verifies $\nabla \cdot \delta \tilde{\mathbf{u}} = 0$ at Γ. With this extra condition, the quantity $\nabla \delta \tilde{\mathbf{u}}$ becomes:

$$\nabla \delta \tilde{\mathbf{u}} = \begin{bmatrix} 1 & -n_x / n_y \\ n_y / n_x & -1 \end{bmatrix} \partial_x \delta \tilde{u}_x = \mathbb{D} \partial_x \delta \tilde{u}_x \tag{3.14}$$

Replacing this assumption on the first term of equation 3.10, we have:

$$\left(-\nu \mathbf{n} \otimes \tilde{\mathbf{u}}^\dagger : \mathbb{D} + \left[\frac{\partial \mathcal{M}_\Gamma}{\partial \nabla \tilde{\mathbf{u}}} : \mathbb{D} \right]^\dagger (\mathcal{M}_\Gamma(\nabla \tilde{\mathbf{u}}, \tilde{p}) - \mathbf{m}_\Gamma), \partial_x \delta \tilde{u}_x \right)_\Gamma \tag{3.15}$$

Where the notation $\mathbb{A} : \mathbb{B}$ stands for $\sum_{i,j} \mathbb{A}_{i,j} \mathbb{B}_{i,j}$. We can see that we now have a scalar equation that will lead to only one extra boundary condition for the adjoint, to be imposed with (3.11). If we simplify the term:

$$-\nu \mathbf{n} \otimes \tilde{\mathbf{u}}^\dagger : \mathbb{D} = -\nu \partial_x \delta u_x \left\{ \frac{\tilde{u}_x^\dagger}{n_x} - \frac{\tilde{u}_y^\dagger}{n_y} \right\} = \nu (n_x n_y)^{-1} \partial_x \delta u_x \mathbf{t} \cdot \mathbf{u}^\dagger \tag{3.16}$$

we notice that, for the boundary-integral term to vanish, we need to impose the tangential velocity from the measurements on the viscous friction together with the normal velocity from the pressure at the wall:

$$\begin{aligned} \mathbf{t} \cdot \tilde{\mathbf{u}}^\dagger &= -\frac{2(n_x n_y)}{\nu} \left\{ \frac{\partial \mathcal{M}_\Gamma}{\partial \nabla \tilde{\mathbf{u}}} : \mathbb{D} \right\}^\dagger (\mathcal{M}_\Gamma(\nabla \mathbf{u}, p) - \mathbf{m}_\Gamma) \\ \mathbf{n} \cdot \tilde{\mathbf{u}}^\dagger &= -\left[\frac{\partial \mathcal{M}_\Gamma}{\partial \tilde{p}} \right]^\dagger (\mathcal{M}_\Gamma(\nabla \tilde{\mathbf{u}}, \tilde{p}) - \mathbf{m}_\Gamma) \end{aligned} \tag{3.17}$$

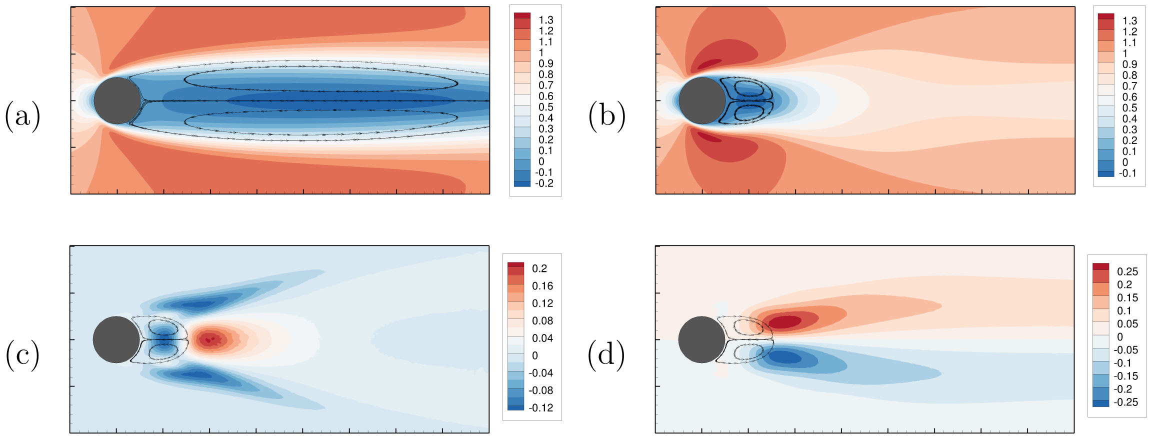

Figure 3.1: (a) Unstable steady-state solution and also the starting point of our optimization procedure and (b) mean-flow and exact divergence Reynolds-stress tensor $\overline{f}_x, \overline{f}_y$ (c,d)

With those boundary conditions, equations (2.1) are will posed. Lastly, by taking the variation of the Lagrangian with respect to the forcing term (in a similar way as before), we obtain the gradient $\nabla_{\tilde{\mathbf{f}}}\mathcal{J}_\Gamma = -\tilde{\mathbf{u}}^\dagger$. With those tools, we are ready to perform wall-measure data-assimilation. We will first apply this technique on a laminar/transitional flow, described in the next paragraph.

2.2 Physical Description of the Considered Laminar Flow

The physical configuration that will be treated here is the same as the one in Foures et al. [2014], the flow around a cylinder at $Re = 150$. The corresponding steady solution (figure 3.1 (a)) is unstable and the time-dependent solution deviates from it, saturating into a nonlinear limit cycle, whose time-average $\overline{u}_x$ can be depicted in figure 3.1 (b), together with the corresponding divergence of the Reynolds-stress tensor $\overline{\mathbf{f}}$ (figure 3.1 (c,d)). The re-circulation bubble's length is much shorter for the mean-flow velocity field ($L_{DNS} = 1.55$, when $L_{Steady} = 9.75$). The goal of the optimizations described on this section will be to get as close as possible to those fields. The reference mean-flow state was obtained through a two-dimensional Direct Numerical Simulation of equations (5.1). This simulation was performed in the software FreeFem++, using $\mathbb{P}^1_b, \mathbb{P}^1_b, \mathbb{P}^1$ discretization (see Hecht [2012]) on a mesh of around 70000 triangles with local refinement on the cylinder's surface and on the wake region. This simulation will also provide the necessary data to perform the assimilation in the following sections.

2.3 Results

In this section we will evaluate how our optimization procedure works for wall-pressure and wall-friction information assimilation.

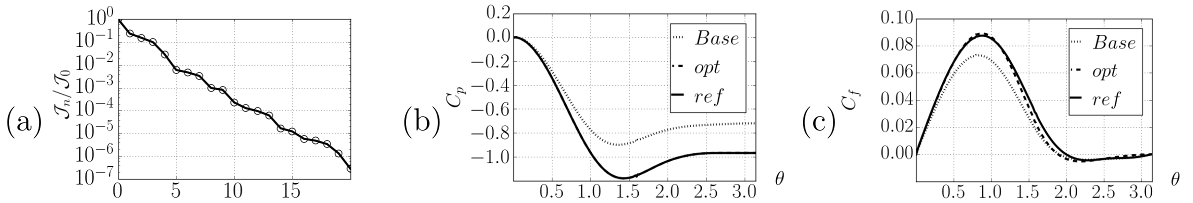

Figure 3.2: Results for pressure on the wall optimization: (a) convergence of the cost functional, (b) wall-pressure and (c) wall-friction in function of the angular coordinate of the cylinder's surface, starting at the stagnation point ($\theta = 0$).

2.3.1 Wall-Pressure Assimilation

To account for the wall-pressure, the generic measure operator introduced before needs to be particularized. Since here we do not perform any measure on the velocity-gradient-related quantities, this operator takes the form:

$$\mathcal{M}_\Gamma(\nabla \tilde{\mathbf{u}}, \tilde{p}) = \tilde{p} \tag{3.18}$$

We remark that, by this choice of measure-operator, we implicitly suppose that the information of wall-pressure of the reference state (DNS presented in the last section) is given on the same computational mesh as the solution. This hypothesis may hold in practice when several pressure probes are available on the surface of the cylinder. However, here, for the laminar case, since our goal is to validate the assimilation procedure, we will not be interested in coarsening the spatial distribution of the measure. For this reason, the measure space M_Γ will be the set of all the functions with support on the boundary. A natural norm for this space is given by $||\cdot||_\Gamma = (\int_\Gamma (\cdot)^2 d\Gamma)^{1/2}$. We can thus infer that:

$$\frac{\partial \mathcal{M}_\Gamma}{\partial \nabla \tilde{\mathbf{u}}} = 0, \frac{\partial \mathcal{M}_\Gamma}{\partial \tilde{p}} = I \Rightarrow \left[\frac{\partial \mathcal{M}_\Gamma}{\partial \tilde{p}}\right]^\dagger = I \tag{3.19}$$

And, consequently, we have that, in a simplified manner:

$$\mathbf{t} \cdot \tilde{\mathbf{u}}^\dagger = 0, \quad \mathbf{n} \cdot \tilde{\mathbf{u}}^\dagger = -(\mathcal{M}_\Gamma(\nabla \tilde{\mathbf{u}}, \tilde{p}) - \overline{\mathbf{m}}_\Gamma) = -(\tilde{p} - \overline{p}) \tag{3.20}$$

With those considerations, the optimization procedure leads to the results presented in figure 3.2 and 3.3.

We can see that the wall-pressure reconstruction manages to exactly capture the actual reference pressure coming from the DNS. Moreover, it does so in a very small number of iterations (Figure 3.2), such that the cost functional $((\mathcal{J}_\Gamma)_n/(\mathcal{J}_\Gamma)_0)$ converges in a near-exponential manner. We can safely say then that the data-assimilation procedure, based on the theory presented in the last section, was successful for the pressure measurement. Furthermore, we can see that, by measuring the pressure, one also reduces the error in the wall-friction (3.2). Looking now at the reconstructed velocity field (3.3 (a)), we can see that, although the re-circulation bubble is now shorter ($L \approx 5.15$), it is still far away from the actual one from the DNS ($L_{DNS} = 1.55$), showing that wall-pressure information may be insufficient for reconstruction of the reference state. Furthermore, we can see that the spatial structure of the forcing term (3.3 (b)) does not compare with the DNS one (Figure 3.1 (c,d)) and is localized around the cylinder due to the inhomogeneity of the adjoint's boundary condition.

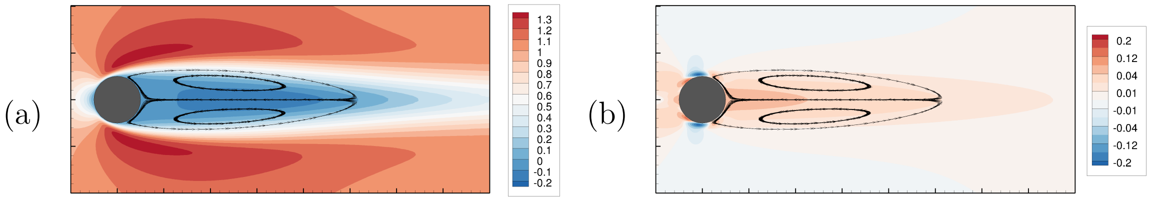

Figure 3.3: Reconstructed velocity field (a) and Reynolds-stress tensor force (b) for wall-pressure assimilation.

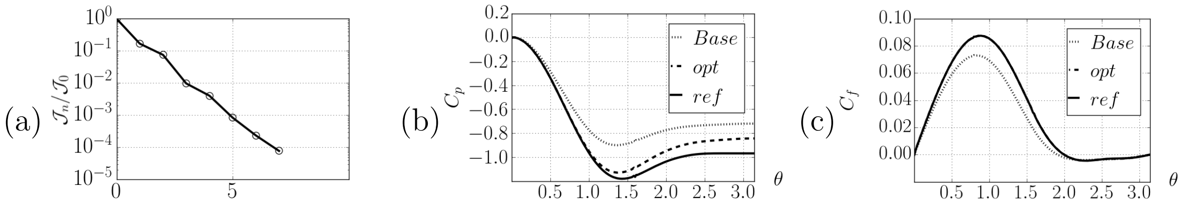

Figure 3.4: Results for wal-friction assimilation: (a) convergence of the cost functional, (b) wall-pressure and (c) wall-friction in function of the angular coordinate of the cylinder's surface, starting at the stagnation point ($\theta = 0$).

2.3.2 Wall-Friction Assimilation

We turn now our attention to the wall-friction measure. In this case, the measure operator will be given by:

$$\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p}) = \nu\mathbf{t}\cdot\nabla_s\tilde{\mathbf{u}}\cdot\mathbf{n} = \nu(\mathbf{n}\otimes\mathbf{t}+\mathbf{t}\otimes\mathbf{n}):\nabla\tilde{\mathbf{u}} \tag{3.21}$$

As before, the norm of the measure space is given by $||\cdot||_\Gamma = \left(\int_\Gamma(\cdot)^2 d\Gamma\right)^{1/2}$, leading to:

$$\frac{\partial\mathcal{M}_\Gamma}{\partial\tilde{p}} = 0, \frac{\partial\mathcal{M}_\Gamma}{\partial\nabla\tilde{\mathbf{u}}} = \nu(\mathbf{n}\otimes\mathbf{t}+\mathbf{t}\otimes\mathbf{n}) \Rightarrow \left\{\frac{\partial\mathcal{M}_\Gamma}{\partial\nabla\tilde{\mathbf{u}}}:\mathbb{D}\right\}^\dagger = \nu(\mathbf{n}\otimes\mathbf{t}+\mathbf{t}\otimes\mathbf{n}):\mathbb{D} \tag{3.22}$$

The boundary conditions for the adjoint in this case are:

$$\mathbf{t}\cdot\tilde{\mathbf{u}}^\dagger = -(\mathcal{M}_\Gamma(\nabla\tilde{\mathbf{u}},\tilde{p}) - \overline{\mathbf{m}}_\Gamma), \quad \mathbf{n}\cdot\tilde{\mathbf{u}}^\dagger = 0 \tag{3.23}$$

The results for this optimization, presented in figures 3.4 and 3.5 are close to the pressure-measurements one. First of all, as before, we can see that the convergence of the cost functional is very fast and the reconstructed wall-friction matches perfectly with the one coming from the DNS. The wall-pressure distribution, although not considered in the cost functional, is also better represented. Furthermore, the reconstructed velocity field presents a re-circulation bubble of $L \approx 5.20$, which is very similar with the one previously shown. Moreover, the forcing term, although slightly different from the one obtained for the pressure measurement, is still concentrated near the wall, also a symptom of the inhomogeneous boundary conditions for the adjoint variable, now on the tangential velocity field.

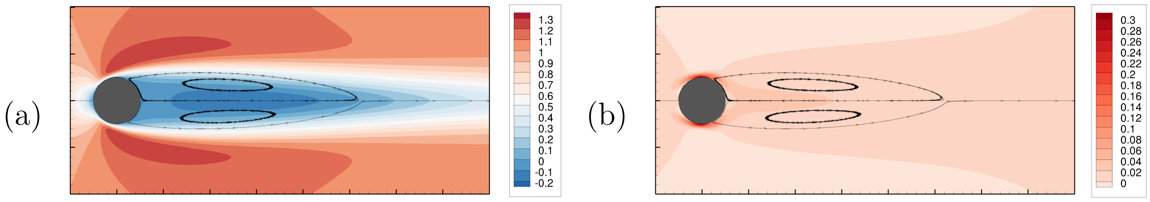

Figure 3.5: Reconstructed velocity field (a) and Reynolds-stress tensor force (b) for wall-friction assimilation.

3 Wall-measure for turbulent RANS

Once the data-assimilation is now validated, we proceed to the turbulent case. Although now the equations are more complex than in the laminar case, due to the consideration of the Spalart-Allmaras model, the volume-adjoint equations remain the same as the ones presented in the previous chapter, with a null volume-forcing term:

$$\begin{aligned}
\tilde{\mathbf{u}}^{\dagger}\cdot(\nabla\tilde{\mathbf{u}})^{T}-\tilde{\mathbf{u}}\cdot\nabla\tilde{\mathbf{u}}^{\dagger}-\nabla\cdot((\nu+\nu_t)\nabla_s\tilde{\mathbf{u}}^{\dagger})-\nabla\tilde{p}^{\dagger} & & \qquad (3.24)\\
+\tilde{\nu}^{\dagger}\nabla\tilde{\nu}+\nabla\cdot(\tilde{\nu}^{\dagger}\partial_{\nabla\tilde{\mathbf{u}}}s) &= \mathbf{0} \qquad (3.25)\\
\nabla\cdot\tilde{\mathbf{u}}^{\dagger} &= 0 \qquad (3.26)\\
-\tilde{\mathbf{u}}\cdot\nabla\tilde{\nu}^{\dagger}-\nabla\cdot(\eta\nabla\tilde{\nu}^{\dagger})+(\partial_{\tilde{\nu}}\eta)\nabla\tilde{\nu}^{\dagger}\cdot\nabla\tilde{\nu}+(\partial_{\tilde{\nu}}\nu_t)\nabla\tilde{\mathbf{u}}^{\dagger}:\nabla_s\tilde{\mathbf{u}} & &\\
-(\partial_{\tilde{\nu}}s)\tilde{\nu}^{\dagger}+\nabla\cdot(\tilde{\nu}^{\dagger}\partial_{\nabla\tilde{\nu}}s) &= 0, \qquad (3.27)
\end{aligned}$$

The boundary conditions for those equations on the solid wall are derived in a similar manner as before. We note that, since the variation of the turbulent variable at the wall is strictly null in the Lagrangian formalism, the adjoin boundary equation does not differ from those presented for the laminar case (equation (3.17)), the difference being the imposition of a homogeneous Dirichlet boundary condition for the variable $\tilde{\nu}^{\dagger}=0$ on the solid wall. We remark that, for both models (model 1 and 2), the adjoint equations are the same. Furthermore, the gradient expressions, in function of the model, are the same ones given in the previous chapter.

3.1 Results for turbulent backward-facing step

We now apply those data-assimilation techniques on the turbulent Backward-Facing Step, already presented in the last chapter. As we will see, contrarily to the laminar case where the final optimal state seems to be almost independent on the measurement of wall-pressure or friction, the final solution now can greatly depend on the measure performed and on the model (1 or 2). We will first present the results of the wall-friction measurement in detail, that seems to be the most difficult quantity to be assimilated. Then we will present the wall-pressure assimilation that, in a way, resembles the partial bulk-velocity assimilation, presented in the previous chapter.

3.1.1 Wall-Friction Information

We turn our attention now to the wall-friction measure. As before, we will consider a complete knowledge of the wall-friction at every point of our computational mesh, limited to the interval $x \in (0, 10)$, in which we find the whole re-circulation bubble. The results of such data-assimilation are plotted in figure 3.6 through 3.8. Starting with the model 2, we can see that the wall-friction assimilation leads to some nonphysical results. Indeed, even if the wall-friction is well represented (apart from an oscillation located at $x \approx 1$, see figure 3.6 (b)), the pressure is well over-estimated (figure 3.6 (a)). Another aspect of this solution, is that the velocity field (figure 3.7 (b)) seems to be 'squeezed' at the wall (all dark-blue lines are very near the wall, contrarily to the dotted lines or the other assimilation solutions) and this has an important reflection on the velocity-error, summarized on the table 3.1. As for the model 1, we can see that it almost exactly reconstructs the wall-friction (figure 3.6 (e)), producing however a small overestimation of the wall-pressure. Looking close to what goes on near the wall, we see that the boundary layer profiles for this model (black solid lines in figure 3.8 (a,b,c)) are distant from the DNS, specially the one at $x = 4.5$, where, in order to match the wall-friction, the velocity profile creates an inflection point and a negative-velocity zone near the wall, leaving the impression that the re-circulation bubble is actually bigger than it is. This problem is due to the particular form of the forcing term $\tilde{\mathbf{f}}_{\mathbf{x}}$, which is non null at the wall and very localized there (see figure 3.8 (d)). We note that this characteristics of the forcing term is not representative of what happens with the reference solution, since the actual term $\bar{\mathbf{f}} = -\nabla \cdot \underline{\tau}$ is null at the wall. This may be at the origin of this anomalous behavior of the wall-friction data-assimilation using the model 1. We note however that, even with this theoretical inconsistency on $\tilde{\mathbf{f}}_{\mathbf{x}}$, one can still produce a solution that is less distorted. For example, in the laminar case presented before, where the Reynolds number was much lower, the velocity field did not seem to produce such behavior. This is consistent with the fact that the viscosity was much higher, allowing the forcing term and its corresponding solution to be smoother. To simulate those conditions in this high-Reynolds number flow, we propose the penalization of the gradients of the forcing term on the cost functional, in the same way presented in the second chapter. We make clear, however, that this proposition does not at all solve the actual theoretical inconsistency problem discussed,

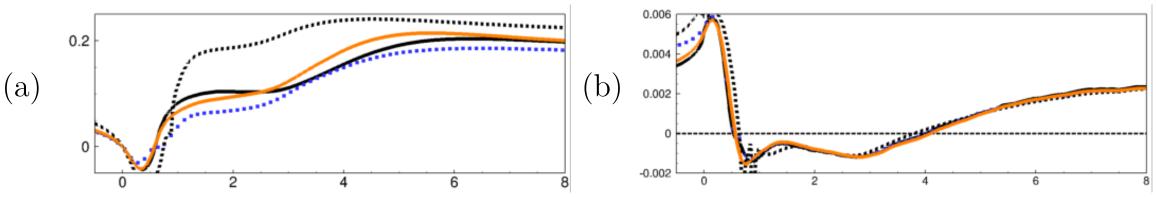

Figure 3.6: Wall-friction assimilation: wall-pressure (b) and wall-friction (c) for model 1 (black curve), model 2 (black dots) and $\gamma = 1$ penalized model 1 (orange curve). DNS results present in blue dots. The color-code will be retained in all the plots in the following

Table 3.1: Global results of wall-friction data-assimilation

	e_Ω			$(\mathcal{J}_\Gamma)_n/(\mathcal{J}_\Gamma)_0$		
	$\mathbf{f}_\mathbf{x}$	$\mathbf{f}_\mathbf{x}$, $\gamma^2 = 1$	$f_{\tilde{\nu}}$	$\mathbf{f}_\mathbf{x}$	$\mathbf{f}_\mathbf{x}$, $\gamma^2 = 1$	$f_{\tilde{\nu}}$
Optimal Solution	0.060	0.065	0.080	0.6%	1.2%	13%
RANS-SA		0.094			–	

as the inhomogeneous boundary conditions will still exist. It only attenuates one symptom of it at high Reynolds number flows. We can see that this strategy manages to correct the tendency of the velocity profiles (3.8), while keeping a good representation of the wall-friction. However, in terms of global velocity-error, the error increases slightly (see table 3.1). From this table, we can also see that neither of the proposed models managed to recover well the actual velocity field from the wall-friction. For this reason, we will set aside this measure and focus on the wall-pressure in the next section.

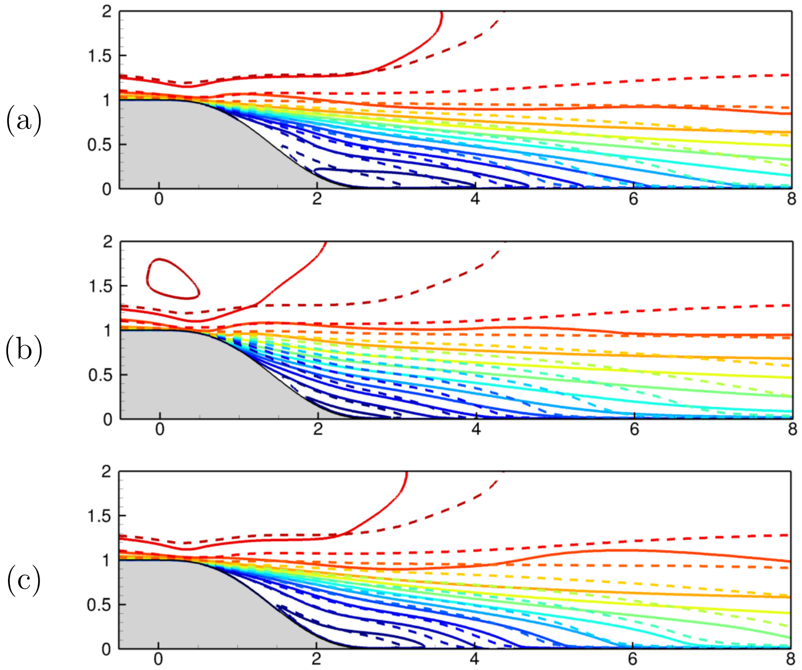

Figure 3.7: Wall-friction assimilation: u_x iso-contours for model 1 (a), model 2 (b) and $\gamma = 1$ penalized model 1 (c).

3.1.2 Wall-Pressure Information

We consider now the complete information of the pressure distribution along the wall as our measure, restricted to the interval $x \in (0, 10)$. The main results are presented in figures 3.9 and 3.10.

By performing such data-assimilation for the two models, we can see that, in terms of the reduction of the cost-functional, the model 1 manages to converge precisely to the exact pressure, whereas the second one, although it produces a very good agreement in most of the wall-region, fails to capture the exact pressure at $x \in (0.5, 1)$. It is as well at this region where the corresponding wall-friction (figure 3.9 (d)) oscillates the most, provoking a (rather small) secondary bubble. This is coherent with the relation between the pressure variation along the wall and the concavity of the wall-normal velocity-error (deduced in the second chapter) $\partial_{\mathbf{t}}(\tilde{p} - \overline{p}) = \nu \partial_{\mathbf{nn}}(\tilde{\mathbf{u}}_{\mathbf{t}} - \overline{\mathbf{u}}_{\mathbf{t}})$, allowing an (adverse) error in wall-pressure to gradually modify the inflection of the normal velocity profile, causing the separation. For the model 1, a (small) secondary bubble also occurs around $x \approx 1.9$. We can also see that both models make similar reconstructions for the re-circulation bubble

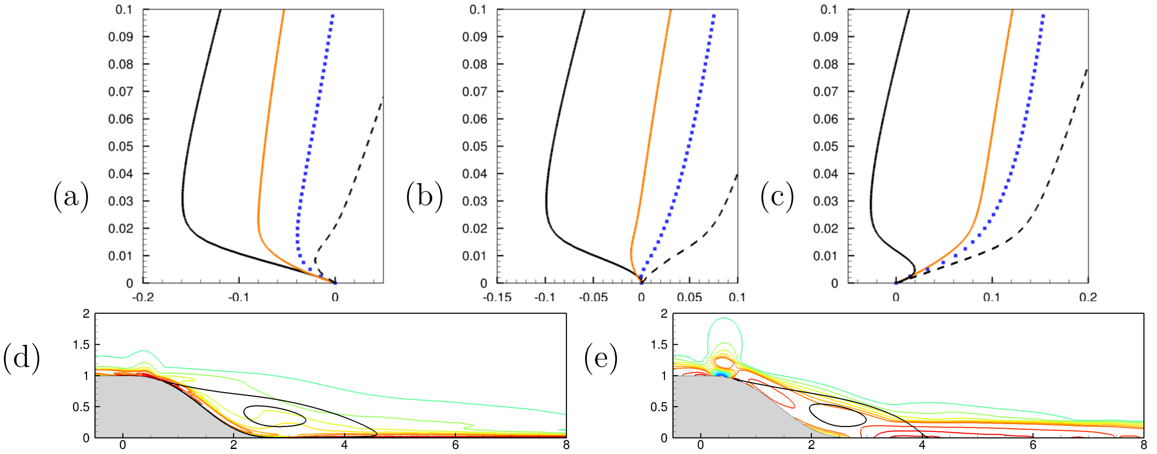

Figure 3.8: Wall-friction assimilation: profiles of $u_x(y)$ velocity at $x = 3.5$ (a), $x = 4$ (b) and $x = 4.5$ (c). Stream-wise component of the optimal forcing for (d) non-penalized case and (e) penalized case. Black lines represent two streamlines.

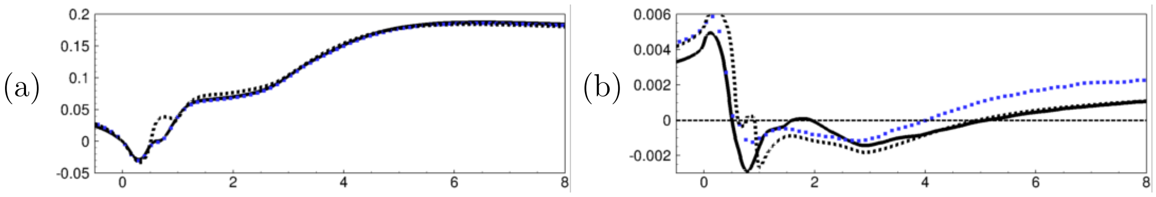

Figure 3.9: Wall-pressure assimilation: wall-pressure (a) and wall-friction (b) for models 1 and 2.

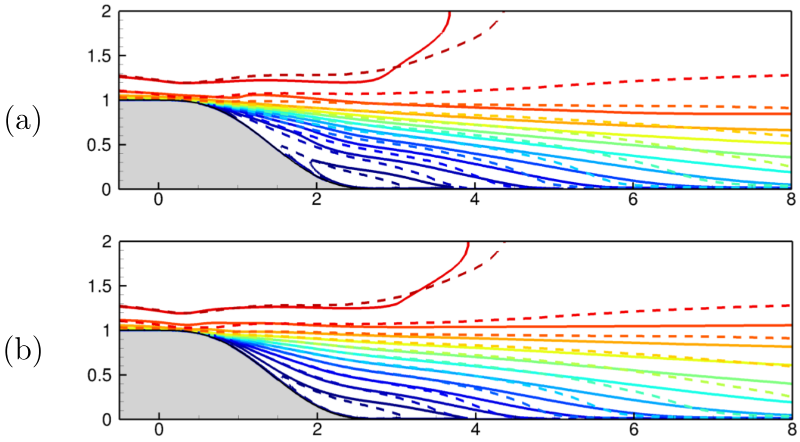

Figure 3.10: Wall-pressure assimilation: u_x iso-contours for model 1 (a) and model 2 (b).

length ($L \approx 5.2$), which corresponds to an intermediate value between RANS-SA and the DNS.

It is interesting to note as well that, for the model 1, the fact that $\tilde{\mathbf{f}}_{\mathbf{x}} \cdot \mathbf{n} \neq 0$ at the wall did not seem to affect the solution in the same way it did for the wall-friction measure. Indeed, the u_y velocity profiles obtained from data-assimilation of both models (3.11) are close to each other and do not present the ”anomalous” behavior as in the case of wall-friction measure. The explanation for this fact is given by the Helmholtz decomposition of the forcing term $\tilde{\mathbf{f}}_{\mathbf{x}}$, similarly to those used in the second chapter. Indeed, we have:

$$\tilde{\mathbf{f}}_{\mathbf{x}} = \tilde{\mathbf{f}}_{\mathbf{x}}^{s} + \nabla \phi \tag{3.28}$$

where $\tilde{\mathbf{f}}_{\mathbf{x}}^{s}$ is solenoidal, satisfying $\tilde{\mathbf{f}}_{\mathbf{x}}^{s} \cdot \mathbf{n} = 0$ at the solid wall Γ and ϕ is a scalar potential. We note that, although the forcing term $\tilde{\mathbf{f}}_{\mathbf{x}}$ is solenoidal (in the bulk), since it holds non-null value at Γ, it can produce a scalar potential different from zero $\phi \neq 0$. Indeed, by taking the divergence of this equation, we have that ϕ is given by a Laplace equation $-\Delta\phi = 0$ in Ω and $\mathbf{n} \cdot \nabla\phi = \tilde{\mathbf{f}}_{\mathbf{x}} \cdot \mathbf{n} \neq 0$ at the wall. This fully determines ϕ. Since ϕ is a scalar potential, its contribution, which holds the non-null component of $\tilde{\mathbf{f}}_{\mathbf{x}} \cdot \mathbf{n}$, can be merged into the pressure, leaving unchanged the actual forcing acting on the velocity field alone. For this reason, the fact that $\tilde{\mathbf{f}}_{\mathbf{x}} \cdot \mathbf{n} \neq 0$ does not seem to be an important issue. It is important to note that the present argument does not apply for the previous case of wall-friction assimilation, since the boundary condition for ϕ comes from the normal component of $\tilde{\mathbf{f}}_{\mathbf{x}} \cdot \mathbf{n}$, and not from the tangential component $\tilde{\mathbf{f}}_{\mathbf{x}} \cdot \mathbf{t}$. Having elucidated those theoretical aspects of the wall-pressure measure, we now proceed to the sparse-measure of the pressure. Here, we suppose that the we dispose of N_w measurements $\overline{\mathbf{m}}_i$. Those measurements are supposed to be localized at some discrete locations on the wall, given by the points $\mathbf{x}_i$. For numerical simplicity of the imposition of the adjoint boundary conditions, we will assume that those measures will correspond to the following quantity:

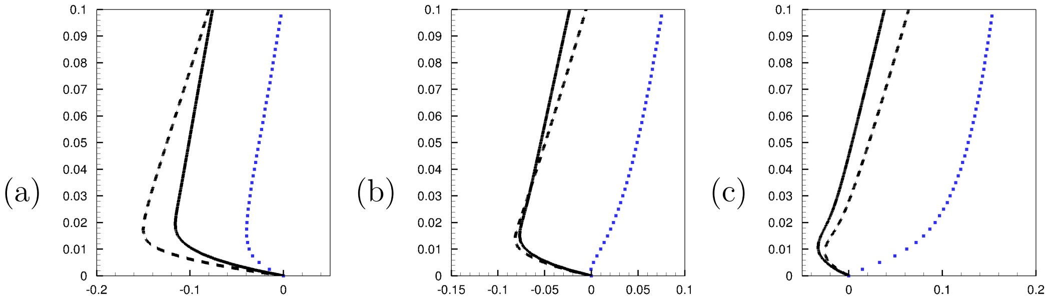

Figure 3.11: Wall-pressure assimilation: profiles of $u_x(y)$ velocity at $x = 3.5$ (a), $x = 4$ (b) and $x = 4.5$ (c).

$$\tilde{\mathbf{m}}_i = \int_{\Gamma} p(\mathbf{x}) g_{\mathbf{x}_i} d\Gamma \approx \int_{\Gamma} p(\mathbf{x}) \delta_{x_i} d\Gamma = p(\mathbf{x}_i) \tag{3.29}$$

where $g_{\mathbf{x}_i}$ is a normalized Gaussian function centered at $\mathbf{x}_i$. The typical standard-deviation of those Gaussian functions are of a few mesh-elements. The boundary conditions of the adjoint problem are given by:

$$\tilde{\mathbf{u}}^{\dagger} \cdot \mathbf{n} = -\sum_{i=1}^{N_w} g_{\mathbf{x}_i} (\tilde{\mathbf{m}}_i - \overline{\mathbf{m}}_i), \quad \tilde{\mathbf{u}}^{\dagger} \cdot \mathbf{t} = 0 \tag{3.30}$$

We also suppose that those measure points $\mathbf{x}_i$ will be evenly distributed in the interval $x \in (0, 10)$. The results of this data-assimilation are summarized in figure 3.12, where the wall-pressure and wall-friction are plotted, together with a vertical gray line, indicating the measure locations. We can see that the pressure is indeed satisfied at the measure locations for both models 1 and 2. We can also see that both models present a reluctance to reconstruct well the pressure around $x \in (0.5, 1)$. This was already observed by the dense wall-pressure measure for the model 2. Thus, this seems to be a difficult region for the pressure assimilation, probably because of the presence of the detachment point. We can also see from figure 3.12 that the resulting wall-friction is globally similar to the one obtained by the full information of the pressure. However, it is interesting to note that, due to the sharp/localized nature of the gradient, given by the measure, the forcing term will also be sharp and may lead to some oscillation on the wall-friction, especially for the model 1, where the forcing term acts more directly on the momentum equations. This issue can be attenuated, again, with the aid of the penalization term of the gradients of $\tilde{\mathbf{f}}_{\mathbf{x}}$. This leads to smoother wall-frictions, but not necessarily to better global velocity-field reconstructions (see table 3.2). From the point of view of the global velocity fields reconstruction, we can see that the model 2 produces systematically a better solution than the first model. We can see that, although the model 1 is more controlable (as presented throughout the second chapter), the information of the pressure is not sufficient for it to reconstruct a correct solution. On the other hand, since the second model is more rigid than the first one, we can, again, take advantage of this to produce better solutions, for example, with $N_w \geq 10$. We can see as well that, for the model 2, the value of the error is approximately the value obtained with velocity data-assimilation, presented in the second chapter.

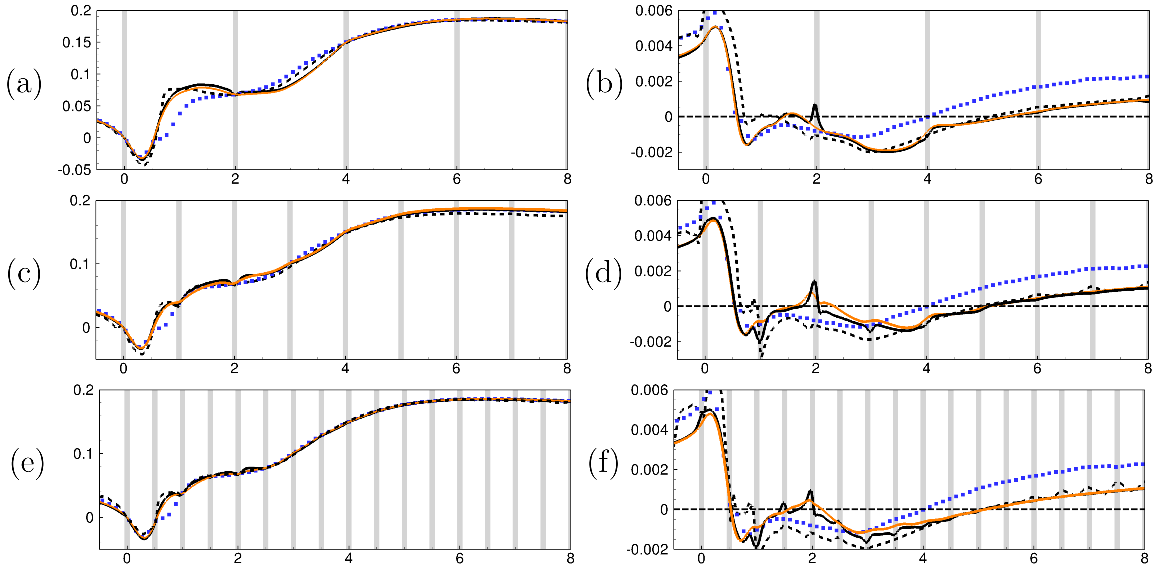

Figure 3.12: Wall-pressure assimilation for three numbers of probes $N_w = 5, 10$ and 20: wall pressure (a,c,e) and wall-friction (b,d,f). Blue points correspond to the DNS, solid lines correspond to model 1, dashed black lines to the model 2 and orange lines to the model 1 with $\tilde{\mathbf{f}}_{\mathbf{x}}$-gradient penalization. The location of pressure measurements are indicated by gray lines.

Table 3.2: Global results of wall-pressure data-assimilation

	e_Ω			$(\mathcal{J}_\Gamma)_n/(\mathcal{J}_\Gamma)_0$		
	$\mathbf{f_x}$	$\mathbf{f_x}, \gamma^2 = 1$	$f_{\tilde{\nu}}$	$\mathbf{f_x}$	$\mathbf{f_x}, \gamma^2 = 1$	$f_{\tilde{\nu}}$
Total Measure	0.057	–	0.037	$\sim 10^{-5}$	–	6%
$N_w = 5$	0.061	0.061	0.045	$\sim 10^{-5}$	$\sim 10^{-5}$	$\sim 10^{-5}$
$N_w = 10$	0.057	0.057	0.038	$\sim 10^{-5}$	$\sim 10^{-4}$	0.7%
$N_w = 20$	0.057	0.057	0.036	$\sim 10^{-5}$	0.2%	0.2%
RANS-SA		0.094			–	

3.1.3 Combined Measure: Wall-Pressure and Velocity

In the last paragraph, we investigated how the data-assimilation procedure performs under sparse wall-pressure assimilation. We could see that, for the models 2, the final assimilated solution was almost independent on the number of measure points. This solution was also very similar to the one obtained with velocity data-assimilation. For this reason, the combined measure will not be performed with this model. However, since the model 1 is fully controlable, it may be interesting to understand the compromise between the pressure or velocity assimilation. For this purpose, we perform two assimilations, one with minimal information ($N_x = 3$, $N_y = 5$ and $N_w = 5$) and other with maximum information ($N_x = 6$, $N_y = 20$ and $N_w = 20$). We can see from table 3.3 that the data assimilation results do not change much with the addition of the wall-pressure measure. This suggests that, from the point of view of the global error e_Ω, the wall-pressure does not include much extra information with respect to the velocity measurement. If, however, wall-pressure probes are preferred instead of bulk velocity ones, we may look for minimizing the amount of the last ones. For this purpose, we could look for optimally place those velocity probes on top of the wall-pressure ones such that the information provided by them is not redundant (see, for example, Mons et al. [2017a]), allowing us to explore directions not observed by wall-pressure probes. This will be considered however out of the scope of the present work.

Table 3.3: Global results of sparse wall-pressure/velocity data-assimilation for model 1 ($\tilde{\mathbf{f}}_\mathbf{x}$) with penalization $\gamma^2 = 1$

	e_Ω		
	Pressure/Velocity	Velocity alone	Pressure alone
$N_x = 3$, $N_y = 5$ $N_w = 5$	0.041	0.042	0.061
$N_x = 6$, $N_y = 20$ $N_w = 20$	0.027	0.027	0.057

4 Conclusions

We could see from the presented wall-measure data assimilation that indeed this information is not in general sufficient to reconstruct correctly the velocity field. In particular, the wall-friction assimilation produced some unexpected results. For example, the second model produced a much thinner re-circulation bubble, over-estimating the wall-pressure. Since the control parameter is completely associated with the turbulence model, we may expect that, by considering other models, we could obtain more appropriated results. As for the model 1, we could notice strong gradients of the forcing term near the wall, leading to a solution that oscillates there, creating a seemingly larger re-circulation bubble. This is shown to be related to some inconsistencies related to a inhomogeneous wall-boundary condition for the adjoint, leading to non-null values for forcing term, which is not the case of the actual forcing term coming from the Reynolds-stress tensor. A possible way to attenuate this effect is to penalize the gradients of the forcing term, smoothing it and consequently the final solution as well.

As for the wall-pressure assimilation, the results obtained are much more coherent with those obtained with the velocity-field, presented on chapter 2. The model 1 does not manage to reconstruct well the velocity field, and we conclude that the wall-pressure is not sufficient information for the capture of the velocity field. The second model, on the other hand, even with few pressure probes, manages to reconstruct the velocity field in the same manner as with the velocity-field data-assimilation. This issue was already explained in the second chapter, where the rigidity of this model was exposed and its use with very few measure was proposed.

Chapter 4

Reslvent-based Data-Assimilation

1 Introduction

The reconstruction of a time-averaged (mean) flow together with the fluctuation around it is a topical issue within data-assimilation in fluid mechanics. In the previous chapters, we explored the mean-flow data-assimilation only through a variational minimization algorithm, where some tuning parameter of the considered model for the mean flow is adjusted so that the resulting approximation of the mean-flow is as close as possible to some input data, coming from experiments of higher-fidelity numerical simulations. However, one common drawback of this approach is the fact that, in order for the final assimilated solution to represent well the reference one, a sufficiently large number of probes is needed. This issue is intimately related to the large number of tuning parameters considered then since, if few measurements are provided, the gradient of those spatially-distributed tuning parameters may be sharp and may demand regularization (such as the gradient penalization employed with sparse velocity measurements, present in the second chapter).

Although those techniques can lead to correct mean-flow reconstructions, they remain purely steady approaches, where mean-flow information is used to reconstruct the mean-flow itself. In the case where one desires to obtain a time-resolved assimilated field, one has to resort to unsteady data-assimilation techniques. Classically, those methods belong to two (possibly overlapping) categories: 3/4DVar (Le Dimet and Talagrand [1986]) and Ensemble Kalman filters (Evensen [2009]). Those methods have been extensively applied in the context of meteorology (Lorenc [1986], Liu et al. [2008]) and more recently in fluid mechanics (Mons et al. [2016], Gronskis A. [2013], D'Adamo et al. [2007]), and Meldi [2018]. This kind of data-assimilation requires some obvious extra numerical elements that were not needed before, such as a time-stepper code, and possibly a time-adjoint code (for 4DVar) and/or a statistical treatment of an ensemble of points in the phase space (for Ensemble Kalman filter).

Another approach proposed in this chapter, more suited for statistically-steady flows, and which does not require many extra numerical tools, relies on the analysis of the linear Navier-Stokes operator around the mean-flow. The origin of such analysis lies in the work of Barkley [2006], where he observed that, in the case of the supercritical flow around a circular cylinder, the stability analysis of the linearized Navier-Stokes operator on the time-averaged (mean) flow (instead of the steady solution, or the base-flow) gives rise to a mode whose eigenvalue is purely imaginary and its frequency matches well with the one coming from the nonlinear limit cycle. The theoretical explanation of this fact was further explored by Sipp and Lebedev [2007a] with the aid of weakly-nonlinear analysis near the bifurcation. They also noticed that the mode coming from this analysis should approximate the first harmonic of the nonlinear limit-cycle. Later, this lead Mantic-Lugo et al. [2015] to develop a self-consistent model describing the nonlinear saturation of such flow, where the steady solution evolves to the mean-flow through the action of the Reynolds-stress tensor, that can be reconstructed with the mode coming from the stability analysis. All the previously mentioned references concerned oscillator flows (see Chomaz [2005]). The extension of the mean-flow analysis to flows whose spectral content is arbitrarily spread in frequency has been conducted by Beneddine et al. [2016], who showed that, under some fairly general assumptions, the Fourier mode of the nonlinear fluctuation around the mean-flow at a given frequency is proportional to the leading mode of the Singular Value Decomposition of the Resolvent operator around the mean-flow, evaluated at that frequency. This means that with the knowledge of the mean-flow one can estimate the spatial structure of the fluctuation, at a given frequency, up to a (complex) multiplicative constant, modeling the energy and the phase of that mode. This constant could be estimated with some extra (sparse) information of the fluctuation field, typically some pointwise time-resolved signal. This procedure was validated for turbulent flow over a backward-facing step, in which case the mean-flow was obtained through a time-resolved simulation (ZDES, see Deck [2005]). The main drawback of this approach is of course the need for the *a priori* knowledge of the mean-flow. To overcome this issue, we propose here to combine the estimation of the fluctuation field from the Resolvent analysis in order to reconstruct the Reynolds-stress tensor and the mean-flow equation involving this Reynolds-stress tensor. By doing so, we obtain a nonlinear model whose unknown parameters are the amplitudes of the Resolvent modes at each considered frequency. For this reason, one could consider this model as an *a priori* reduced order model for the unsteady flow, which is then used for data-assimilation. For this reason, only very sparse information is needed for the reconstruction of not only the mean-flow but also the fluctuation around it.

In this chapter, this procedure will be applied to simple model problems of transitional laminar flows. The first configuration will be the flow around a square cylinder ($Re = 60, 100$), falling in the category of oscillator flows, for which the frequency spectrum is exhibits a peak and its harmonics. This means that only one Resolvent mode is needed at one frequency (and thus only one parameter) for the reconstruction procedure. We will show that the velocity at some point in the flow then suffices for the estimation of the mean-flow and the fluctuation field. The second configuration is the two-dimensional backward-facing step ($Re = 500$), a typical example of noise amplifier Hervé et al. [2012]. For this flow, the spectral content of the fluctuation will strongly depend on the external noise that triggers the dynamics. We will consider two different cases, one for which the noise is characterized by two incommensurable frequencies and the other for which it is a broadband white noise. The frequency content of the flow, which can be inferred from time-resolved signals at some points of the flow, can then be used as an input signal for the data assimilation process or as a reference signal to assess the quality of the reconstruction.

This chapter proceeds as follows. First (§2), we present the model combining the mean-flow and the perturbation equations. In particular, we show how to solve for this problem for a given set of Resolvent amplitudes. Then, we successively present the results of the data-assimilation procedure to the cylinder flow case (§3) and then to the backward facing step flow (§4).

2 Theory

This section is devoted to the derivation of the model that will be used in the data-assimilation procedure. We consider an incompressible flow, whose velocity and pressure fields are given by the Navier-Stokes equations :

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - Re^{-1} \Delta \mathbf{u} = \mathbf{0}, \quad \nabla \cdot \mathbf{u} = 0 \tag{4.1}$$

where $Re = U_\infty L/\nu$ is the Reynolds number based on an arbitrary velocity U_∞ and a characteristic length L, which are used to make all the variables non-dimensional. Those equations can be recast into a more compact form:

$$B\partial_t \mathbf{q} + L_\mathbf{0} \mathbf{q} + N(\mathbf{q}, \mathbf{q}) = \mathbf{0} \tag{4.2}$$

where the state vector $\mathbf{q}$ is composed by the concatenation of the velocity and pressure fields $\mathbf{q} = (\mathbf{u}, p)$ and the operators B, $L_\mathbf{0}$ (Stokes operator) and $N(\mathbf{q}_1, \mathbf{q}_2)$ (nonlinear convection, sometimes also denoted by $N(\mathbf{u}_1, \mathbf{u}_2)$ since it is independent on the pressure fields) are given by:

$$B = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}, \quad L_\mathbf{0} = \begin{bmatrix} -Re^{-1}\Delta() & \nabla() \\ \nabla \cdot () & 0 \end{bmatrix}, \quad N(\mathbf{q}_1, \mathbf{q}_2) = \begin{bmatrix} \mathbf{u}_1 \cdot \nabla \mathbf{u}_2 + \mathbf{u}_2 \cdot \nabla \mathbf{u}_1 \\ 0 \end{bmatrix} \tag{4.3}$$

Since we focus on statistically steady regimes, we may decompose the instantaneous state $\mathbf{q}(\mathbf{x}, t)$ into a mean and a fluctuating component, as

$$\mathbf{q}(\mathbf{x}, t) = \overline{\mathbf{q}}(\mathbf{x}) + \mathbf{q}'(\mathbf{x}, t) \tag{4.4}$$

where the time-averaged operator is defined as

$$\overline{(\cdot)} = \lim_{T \to +\infty} \frac{1}{T} \int_0^T (\cdot)(t) dt. \tag{4.5}$$

Introducing the decomposition (4.4) into the Navier-Stokes equations (5.1) and taking the time-average yields the mean flow equation:

$$N(\overline{\mathbf{q}}, \overline{\mathbf{q}}) + L_\mathbf{0} \overline{\mathbf{q}} = -\overline{N(\mathbf{q}', \mathbf{q}')} = P\overline{\mathbf{f}} \tag{4.6}$$

which differs from the fixed point solution equation by the presence of the right-hand side term $\overline{\mathbf{f}} = -\overline{\mathbf{u}' \cdot \nabla \mathbf{u}'}$ (acting on the equations through a prolongation $P = (I, 0)^T$) that expresses the time-averaged effect of the fluctuating velocity induced by the quadratic non-linearity (forcing term coming from the Reynolds stress tensor). The goal here is to model this term by approximating the fluctuation $\mathbf{u}'$ by singular vectors of the Resolvent operator (Beneddine et al. [2016]). To do so, we subtract the mean flow equation (5.10) from the instantaneous equations (5.1), leading to:

$$B\partial_t \mathbf{q}' + L_{\overline{\mathbf{q}}} \mathbf{q}' = P\mathbf{f}' \tag{4.7}$$

where the operator $L_{\overline{\mathbf{q}}} = L_\mathbf{0} + N_{\overline{\mathbf{q}}} = L_\mathbf{0} + N(\overline{\mathbf{q}}, \cdot) + N(\cdot, \overline{\mathbf{q}})$ is the linearized Navier-Stokes operator and $\mathbf{f}' = -\mathbf{u}' \cdot \nabla \mathbf{u}' + \overline{\mathbf{u}' \cdot \nabla \mathbf{u}'}$ represents the nonlinear fluctuations, acting as a forcing term. We then try to find solutions of this equation in the frequency space by using the ansatz $\mathbf{q}' = \hat{\mathbf{q}} e^{\mathrm{i}\omega t} + c.c.$ and $\mathbf{f}' = \hat{\mathbf{f}} e^{\mathrm{i}\omega t} + c.c.$, leading to :

$$(\mathrm{i}\omega B + L_{\overline{\mathbf{q}}})\hat{\mathbf{q}} = P\hat{\mathbf{f}} \tag{4.8}$$

This equations represents an (infinite-dimensional) input/output relation between the forcing term $\hat{\mathbf{f}}$ and its perturbation $\hat{\mathbf{q}}$. In our approach we would like to simplify this equation, retaining only the dynamics that maximizes the gain in kinetic energy:

$$G(\omega) = \frac{||\hat{\mathbf{u}}||^2}{||\hat{\mathbf{f}}||^2} = \frac{(P\hat{\mathbf{q}}, P\hat{\mathbf{q}})}{(\hat{\mathbf{f}}, \hat{\mathbf{f}})} = \frac{(R^\dagger(\overline{\mathbf{q}}; \omega) R(\overline{\mathbf{q}}; \omega)\hat{\mathbf{f}}, \hat{\mathbf{f}})}{(\hat{\mathbf{f}}, \hat{\mathbf{f}})} \tag{4.9}$$

where $R(\overline{\mathbf{q}}, \omega)$ is the so-called Resolvent operator, defined as $R(\overline{\mathbf{q}}, \omega) = P^T(\mathrm{i}\omega B + L_{\overline{\mathbf{q}}})^{-1}P$ and its adjoint $R^\dagger(\overline{\mathbf{q}}, \omega)$ given by the general deffinition $(R^\dagger \mathbf{a}, \mathbf{b}) = (\mathbf{a}, R\mathbf{b}), \forall \mathbf{a}, \mathbf{b}$. Here we consider the energy inner product $(\mathbf{a}, \mathbf{b}) = \int_\Omega \mathbf{a}^* \cdot \mathbf{b} d\Omega$.

We can show that the extraction of such dynamics is equivalent to the Singular Value Decomposition of the Resolvent operator $R(\overline{\mathbf{q}}, \omega)$, i.e:

$$R^\dagger(\overline{\mathbf{u}}; \omega) R(\overline{\mathbf{u}}; \omega)\hat{\mathbf{f}}_k = \mu_k^2(\omega)\hat{\mathbf{f}}_k, \quad ||\hat{\mathbf{f}}_k|| = 1 \tag{4.10}$$

where μ_k^2 are the Singular Values (ordered such that $\mu_0^2 \geq \mu_1^2 \geq \cdots \geq 0$), $\hat{\mathbf{f}}_k$ are the Right Singular Functions or the optimal forcing ($\hat{\mathbf{f}}_0$ being the optimal forcing term). The Left Singular Functions are defined as $\hat{\mathbf{u}}_k = \mu_k^{-1} R(\overline{\mathbf{u}}; \omega)\hat{\mathbf{f}}_k$. Alternatively, one can write the SVD problem for the variables $\hat{\mathbf{u}}_k$ directly. To do so, we apply the operator $\mu_k^{-1} R(\overline{\mathbf{u}}; \omega)$ on equation (4.10), leading to:

$$R(\overline{\mathbf{u}}; \omega) R^\dagger(\overline{\mathbf{u}}; \omega)\hat{\mathbf{u}}_k = \mu_k^2(\omega)\hat{\mathbf{u}}_k, \quad ||\hat{\mathbf{u}}_k|| = 1 \tag{4.11}$$

If the nonlinear term $\hat{\mathbf{f}}(\omega)$ is expanded in the orthonormal basis constituted of the Right Singular Vectors $\hat{\mathbf{f}}_k$ and if we apply the Resolvent operator on it, we obtain an expansion of $\hat{\mathbf{u}}(\omega)$ in terms of the Left Singular Vectors:

$$\hat{\mathbf{u}}(\mathbf{x}, \omega) = \sum_{k=0}^{+\infty} \mu_k(\omega)\hat{\mathbf{u}}_k(\mathbf{x}, \omega) \left(\hat{\mathbf{f}}_k(\omega), \hat{\mathbf{f}}(\omega)\right) \tag{4.12}$$

This expression indicates that, when the forcing Fourier modes $\hat{\mathbf{f}}(\omega)$ is not precisely known, the perturbation $\hat{\mathbf{u}}$ cannot be exactly determined. However, when the flow presents a strong convective unstable behavior, one dominant mode is expected such that $\mu_0 \gg \mu_k,\ k > 0$. In this case, a good approximation of the Fourier mode of the fluctuation is given by the first term in (4.12):

$$\hat{\mathbf{u}}(\mathbf{x}, \omega) \approx \mu_0(\omega)\hat{\mathbf{u}}_0(\mathbf{x}, \omega) \left(\hat{\mathbf{f}}_0(\omega), \hat{\mathbf{f}}(\omega)\right) = A(\omega)\hat{\mathbf{u}}_0(\mathbf{x}, \omega) \tag{4.13}$$

where the (complex) function $A(\omega)$ holds the information of the energy and phase of the mode at frequency ω. At this point, this approximation of the Fourier representation of the fluctuation is still not suited for numerical purposes as it is not yet 'discrete' in frequency space. In the following, we consider two distinct cases: first a simple case where the frequency content is mostly 'peaked' (oscillator flows) around one frequency and then, a more complex one, where it is more broadband (amplifier flow).

2.1 Time-periodic flows

Many flows present a periodic behaviour such as the flows around bluff bodies. In those cases the steady-state solution is unstable and the instability develops into a saturated periodic limit-cycle. Furthermore, it is not uncommon that most of the energy of this periodic limit-cycle is concentrated at the fundamental frequency ω_0 (see Turton et al. [2015a]). For those flows, a single-frequency description of the fluctuation may be acceptable:

$$\mathbf{u}' \approx A\hat{\mathbf{u}}_0 e^{\mathrm{i}\omega_0 t} + c.c. \tag{4.14}$$

where the parameter A is now a single complex scalar. Based on this, we can approximate the Reynolds-stress tensor and the force acting on the mean flow equation following $\tilde{\mathbf{f}} = -2|A|^2\mathrm{Re}\{\hat{\mathbf{u}}_0 \cdot \nabla\hat{\mathbf{u}}_0^*\}$. Considering the mean flow equation 5.10, an approximation $\tilde{\mathbf{u}}$ of the mean-flow $\overline{\mathbf{u}}$ may be obtained through:

$$N(\tilde{\mathbf{q}}, \tilde{\mathbf{q}}) + L_{\mathbf{0}}\tilde{\mathbf{q}} = P\tilde{\mathbf{f}} = -2|A|^2\mathrm{Re}\{N(\hat{\mathbf{u}}_0, \hat{\mathbf{u}}_0^*)\} \tag{4.15a}$$

$$R(\omega_0)R^\dagger(\omega_0)\hat{\mathbf{u}}_0 = \mu_0^2(\omega_0)\hat{\mathbf{u}}_0 \tag{4.15b}$$

where $R(\omega_0) = R(\tilde{\mathbf{q}}; \omega_0)$.

At this point, we remark that the system formed by the mean flow equation and the Resolvent analysis (4.15) is a closed set of nonlinear equations, if one knows the frequency of the nonlinear signal ω_0 and its energy $|A|^2$. Arguably, already from a data-assimilation point of view, one can know *a priori* the frequency of the flow from a time-resolved signal from the flow, for instance a hot-wire or a point-wise wall-pressure measurement. An additional measure (amplitude of the oscillations, length of re-circulation bubble, strength of backflow, ...) can then be used to determine the energy of the fluctuation $|A|^2$, since, for every value of $|A|^2$ one can solve for the nonlinear solution of 4.15 and therefore tune this parameter so that the solution matches to the actual measure. The output of this procedure is the full flow-field $\mathbf{u}$ composed of the reconstructed mean-flow $\tilde{\mathbf{u}}$ and the fluctuation $\mathbf{u}'$ given in equation (4.14). The strength of the following approach lies in the simplicity of the optimization problem, since it only involves a single parameter contrarily to the procedure described in Foures et al. [2014], where complete spatial fields had to be determined.

2.2 Broadband Flows

For broadband flows , the previous approximation is no longer suited. In those cases we have thus to consider a multiple (N) frequency approach $\{\omega_j\}_{j=1,\cdots,N}$, for which ($A_j = A(\omega_j)$):

$$\mathbf{u}' = \sum_{j=1}^{N} A_j\hat{\mathbf{u}}_0^j e^{\mathrm{i}\omega_j t} + c.c. \tag{4.16}$$

This approximation is only valid if the dominant frequencies of the unsteadiness are located within a frequency range where the assumption of separation of the first singular value holds. This is usually the case in flows dominated by shear-driven unsteadiness, such as jets, wakes, shear-layers, etc.

If we consider this approximation in the mean flow equation 5.10, we have the following equation governing the mean flow $\tilde{\mathbf{u}}$:

$$N(\tilde{\mathbf{q}},\tilde{\mathbf{q}})+L_{\mathbf{0}}\tilde{\mathbf{q}}=P\tilde{\mathbf{f}}=-2\sum_{j=1}^{N}|A_j|^2\mathrm{Re}\{N(\hat{\mathbf{u}}_0^j,\hat{\mathbf{u}}_0^{j,*})\} \tag{4.17a}$$

$$R(\omega_j)R^{\dagger}(\omega_j)\hat{\mathbf{u}}_0^j=\mu_0^2(\omega_j)\hat{\mathbf{u}}_0^j \tag{4.17b}$$

The choice of the discrete frequencies $\{\omega_j\}$ will depend on the knowledge of the physics of the problem, and can be determined from few point-wise measurements located at highly energetic regions of the flow. Moreover, similarly as before, the energy of each frequency $|A_j|^2$ can be tuned using some measurements of the flow field through the model 4.17. The major difference here lies in the fact that the optimization procedure now involves N parameters $(|A_0|,|A_1|,\cdots)\in\mathbb{R}^N$ to be tuned. The output of this model is as before the reconstructed mean flow $\tilde{\mathbf{u}}$ and the corresponding fluctuations at the sampling frequencies ω_j, given by 4.16.

2.3 Numerical resolution of models

In this paragraph we provide some elements on how we chose to solve the system of nonlinear equations given by 4.15 (or 4.17). One possible way is to iteratively solve the mean flow for a given SVD mode and then solve a new SVD problem with the new mean flow and keep alternating in this way up to the point where the solution remains fixed. This method was employed by Mantic-Lugo et al. [2015] and Mantic-Lugo and Gallaire [2016] on a similar model than ours. However, it was observed that for relatively high values of $|A|$, the method needs a quite low under-relaxation factor, making the convergence quite slow. Here we propose an alternative, which is to regard the set of equations 4.15 as a whole nonlinear system and apply the Newton method to it. To do so, first of all, we note that the SVD eigensystem (equation 4.11) is not yet suited for this purpose, since, to extract the complete Jacobian matrix, we need to calculate, for example, the variations of those equations with respect to the state $\tilde{\mathbf{q}}$, which does not show itself explicitly. To pose the same problem in a more suitable way, we write an eigensystem for the variables $(\hat{\mathbf{q}}_0,\hat{\mathbf{a}}_0)$:

$$\begin{aligned}&\lambda(\mathrm{i}\omega_0B+L_{\tilde{\mathbf{q}}})\hat{\mathbf{q}}_0=B\hat{\mathbf{a}}_0,\quad \hat{\mathbf{q}}_0^{*,T}B\hat{\mathbf{q}}_0=1\\&(\mathrm{i}\omega_0B+L_{\tilde{\mathbf{q}}})^*\hat{\mathbf{a}}_0=B\hat{\mathbf{q}}_0\end{aligned} \tag{4.18}$$

such that $\hat{\mathbf{u}}_0=P^T\hat{\mathbf{q}}_0$, $\hat{\mathbf{f}}_0=\mu_0^2P^T\hat{\mathbf{a}}_0$ and $\lambda=\mu_0^2$. The variable $\hat{\mathbf{a}}_0$ is defined as $\hat{\mathbf{a}}_0=(i\omega_0B+L_{\tilde{\mathbf{q}}})^{-1,*}B\hat{\mathbf{q}}_0$. We remark here that, for the sake of clarity, we keep the 'continuous' notation introduced in previous paragraphs, even if now those objects are discrete, since we are dealing with the numerical method. For example, $\tilde{\mathbf{q}}$, $\hat{\mathbf{u}}_0$, etc, are now vectors and $B, L_{\tilde{\mathbf{q}}}$, etc, matrices.

Together with those equations, we need to solve the mean-flow equation as well (for simplicity, we only show the single-frequency case given by equation 4.15). Its linearization, together with the linearization of the rewritten eigensystem leads to:

$$\begin{bmatrix} L_{\tilde{\mathbf{q}}} & |A|^2 N_{\hat{\mathbf{q}}_0^*}(\cdot) + |A|^2 N_{\hat{\mathbf{q}}_0}(\cdot)^* & 0 & 0 \\ \lambda N_{(\cdot)}\hat{\mathbf{q}}_0 & \lambda(\mathrm{i}\omega_0 B + L_{\tilde{\mathbf{q}}}) & (\mathrm{i}\omega_0 B + L_{\tilde{\mathbf{q}}})\hat{\mathbf{q}}_0 & -B \\ 0 & \hat{\mathbf{q}}_0^{*,T} B(\cdot) & 0 & 0 \\ N_{(\cdot)}^{*,T}\hat{\mathbf{a}}_0 & -B & 0 & (\mathrm{i}\omega_0 B + L_{\tilde{\mathbf{q}}})^{*,T} \end{bmatrix} \begin{bmatrix} \delta\tilde{\mathbf{q}} \\ \delta\hat{\mathbf{q}}_0 \\ \delta\lambda \\ \delta\hat{\mathbf{a}}_0 \end{bmatrix} = \begin{bmatrix} -L_{\mathbf{0}}\tilde{\mathbf{q}} - N(\tilde{\mathbf{q}},\tilde{\mathbf{q}}) - |A|^2 N(\hat{\mathbf{q}}_0,\hat{\mathbf{q}}_0^*) - |A|^2 N(\hat{\mathbf{q}}_0^*,\hat{\mathbf{q}}_0) \\ B\hat{\mathbf{a}}_0 - \lambda(\mathrm{i}\omega_0 B + L_{\tilde{\mathbf{q}}})\hat{\mathbf{q}}_0 \\ (1 - \hat{\mathbf{q}}_0^{*,T} B\hat{\mathbf{q}}_0)/2 \\ B\hat{\mathbf{q}}_0 - (\mathrm{i}\omega_0 B + L_{\tilde{\mathbf{q}}})^{*,T}\hat{\mathbf{a}}_0 \end{bmatrix} \tag{4.19}$$

The iterative solution of this system of equations and the update of the solution leads to the aforementioned Newton method. It was observed that the convergence of such method was often difficult and the reason for this is not fully understood. However, we could observe that, by changing slightly the equations, the convergence of the method was improved. First of all, we no longer suppose that λ is a real parameter, and we will let it free to evolve in the complex plane as the solution converges. Moreover, since now we have added a new variable to the system, namely the imaginary part of λ, we need a supplementary equation for the Jacobin of the system to be a square matrix. Note that if, instead of having an SVD problem, we had an eigenvalue one to be solved, involving a non-normal operator (such as, for example Mantic-Lugo et al. [2015]), the eigenvalue should not be considered to be real and this extra equation had to be considered. For this purpose, we remark that, from equations 4.10, 4.11 or 4.18, the Singular Modes are defined up to an arbitrary phase $e^{\mathrm{i}\phi}$. We construct then the extra equation that fixes this phase, which can be, for example the imposition that the imaginary part of some point-wise quantity of the mode is null. This strategy leads to a nonlinear system that converges much better than the previous one. We remark that, in order to have a valid solution of the SVD system, the imaginary part of λ needs to be verified to be null once the residual has been decreased. Indeed, for all computations here presented, the ratio between imaginary and real parts of λ is of order 10^{-12}. This linear system is then solved with an GMRES iterative solver, preconditioned with the lower triangular matrix (lower Gauss-Seidel), meaning that the equation for $\delta\tilde{\mathbf{q}}$ are solved first, followed by the equation for $\delta\hat{\mathbf{q}}_0$ and $\delta\hat{\mathbf{a}}_0$.

3 An "Oscillator" Flow - Squared-Section Cylinder

The two-dimensional flow around a squared-section cylinder is a typical example of oscillator flows. For Reynolds number larger than $Re \approx 50$ (see Sohankar et al. [1998] for discussions around the critical value), it naturally develops an unsteady behaviour, corresponding to a periodic limit cycle, characterized by a single frequency and its harmonics. Below, we will first (§3.1) present the details of the configuration and the numerical implementations. Then (§3.2), we will consider a single time resolved measurement in the wake of the cylinder to reconstruct both the mean-flow and the perturbation dynamics. Finally (§3.3), the case of other measurements solely based on mean-flow characteristics will be explored.

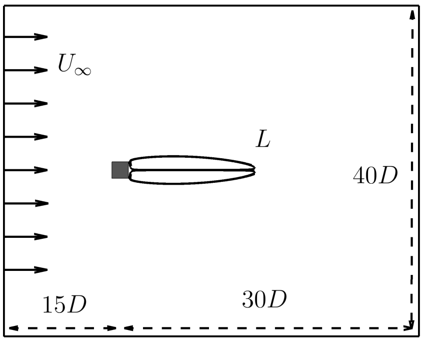

Figure 4.1: Sketch of the physical domain and inflow boundary condition for the square cylinder case. The re-circulation bubble's length is indicated in the figure and will be defined as L.

3.1 Configuration and numerical implementation

A sketch of the computational domain and the inflow boundary condition is provided in Figure 4.1. The inflow boundary is located 15 diameters away from the center of the cylinder and the velocity profile imposed is uniform. Two Reynolds numbers $Re = DU_\infty/\nu$ have been considered here, $Re = 60, 100$. The side walls are located 20 diameters away from the cylinder, on which we impose a slip condition. The outflow boundary condition is imposed 30 diameters downstream of the cylinder and reads $(pI - Re^{-1}\nabla\mathbf{u})\cdot\mathbf{n} = \mathbf{0}$. Both the Direct Numerical Solution and the models are solved on the same two-dimensional mesh. All codes are second order in space and are based on finite elements. The DNS solver employs a second-order semi-implicit temporal scheme for time-advancement. All implementations have been performed in the FreeFEM++ open source code (Hecht [2012]).

In Table 4.1, we show various quantities of interest to characterize the mean- and unsteady features of the flowfield: the re-circulation bubble's length L (for the steady solution and the time-averaged unsteady solutions), the fundamental frequency ω_0 and the amplitude $|A^{DNS}|$ of the Fourier mode at the fundamental frequency in the unsteady simulation. Results are given for 3 meshes of different density. All of them are strongly refined around the square and in its wake and are coarsened in the freeflow region. We can see that mesh 2 provides a good agreement with the finer mesh 3 for all observed quantities. Mesh 2 will therefore be chosen as the default mesh in all following computations. In particular, the model will be computed on this mesh.

3.2 Single point velocity measurement

Let us first assume that we only know the time evolution of (for example) the cross-stream velocity component at a single point in the wake of the cylinder, here $\mathbf{x}_m = (x, y) = (2, 0)$. This quantity is represented in figures 4.2(a,d) for the two Reynolds numbers. The frequency content of this signal may be extracted by considering harmonic averages Arbabi and Mezić [2017]:

$$\check{\psi}(\omega) = \lim_{T\to+\infty} \frac{1}{T}\mathcal{F}_T(\psi)(\omega), \tag{4.20}$$

Table 4.1: Mesh dependency of steady and unsteady simulations for square cylinder at $Re = 100$ - chosen mesh for the computations in boldface. Observed parameters are recirculation bubble's length L, nonlinear frequency ω_0, first harmonic energy $|A^{DNS}|$ and dominant gain, evaluated at the nonlinear frequency $\mu_0^2(\omega_0)$.

		Steady Solution		DNS		
	# triangles	L	$\mu_0^2(\omega_0)$	L	ω_0	$\|A^{DNS}\|^2$
Mesh 1	43000	8.42	213775	2.43	0.913	2.131
Mesh 2	32000	8.42	211745	2.43	0.912	2.130
Mesh 3	12000	8.40	206077	2.44	0.912	2.128

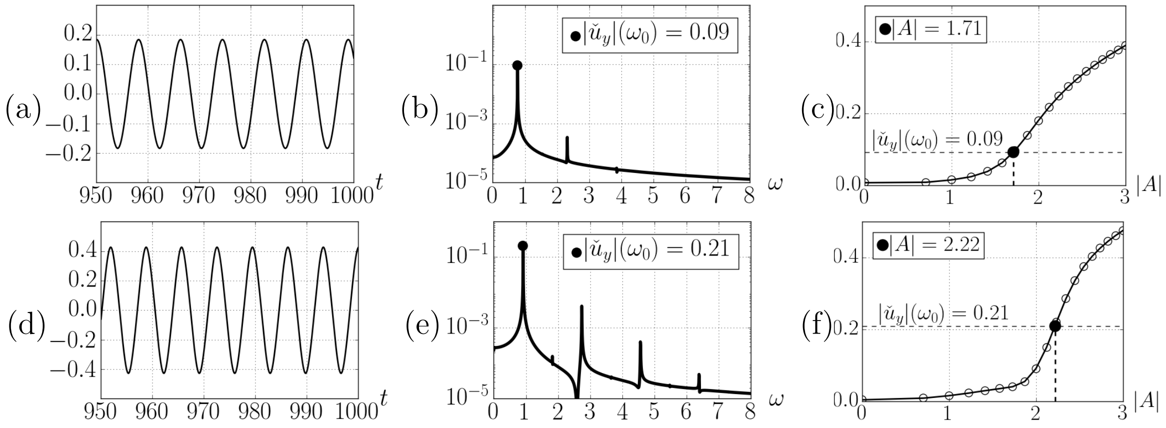

Figure 4.2: Square cylinder flow at $Re = 60$ (a,b,c) and $Re = 100$ (d,e,f): (a,d) measured signal corresponding to the cross-velocity u_y at $(x, y) = (2, 0)$, (b,e) Discrete Fourier Transform of signal in (a,d), giving the modulus of the Fourier coefficient ($|\check{u}_y|(\omega_0) = 0.09/0.21$) at the fundamental frequency $\omega_0 = 0.77/0.91$ and (c,f) comparison of this quantity with the one given by model (4.15) for several values of $|A|$, the optimal one being $|A| \approx 1.71/2.23$.

where $\mathcal{F}_T(\psi)(\omega)$ represents the Fourier transform on the finite time interval $(0, T)$:

$$\mathcal{F}_T(\psi)(\omega) = \int_0^T e^{-\mathrm{i}\omega t}\psi(t)dt. \tag{4.21}$$

The quantity $\check{\psi}(\omega)$ is finite and independent of the value of T. The amplitude $|\check{u}_y(\mathbf{x}_m, \omega)|$, represented as a function of ω in figures (b,e), indicates that the fundamental mode at ω_0 strongly dominates those of the harmonics $p\omega_0$ with $p \geq 2$. This suggests that a reconstruction with a single frequency may be sufficient. The fundamental frequencies of the periodic limit cycles are equal respectively to $\omega_0 = 0.77$ and 0.91 for $Re = 60$ and $Re = 100$, while the amplitudes correspond to $|\check{u}_y(\mathbf{x}_m, \omega_0)| = 0.09$ and 0.21.

Since there is only one free parameter $|A|$ in the model 4.15, we may explore its reconstruction ability by evaluating how the solution depends on $|A|$ and comparing it with the actual measurement data shown in figures 4.2(a,b,d,e). From a numerical point of view, this does not necessarily represent a much high computational effort since one may proceed by progressively increasing the value of $|A|$ and restart each new computation from the solution corresponding to the highest available $|A|$ (if no solution is available, we may start from $|A| = 0$, the steady solution). By doing so and by changing the value

$|A|$ slowly within the nonlinear solver, each nonlinear solver converges in a few iterations. This convergence deteriorates when the gap between two values of $|A|$ becomes large. For example, if we desire to solve the model for a relatively high value of $|A|$ from the steady solution ($|A| = 0$), we may need, for example, 10 iterations to solve it for $|A| = 1,5$ or 20 iterations for $|A| = 2.2$.

Once those solutions obtained, in order to compare the model's solution in function of $|A|$ with the available information, we apply the operator $\check{(\cdot)}$ to the fluctuation predicted by model (4.14):

$$|\check{u}_y(\omega_0)| = \left| \lim_{T \to +\infty} \frac{1}{T} \int_0^T e^{-\mathrm{i}\omega_0} \mathbf{u}'(\mathbf{x}_m, t) dt \right| = |A| |\hat{\mathbf{u}}_0(\mathbf{x}_m)| \tag{4.22}$$

We can now investigate which value of $|A|$ provides a solution for which the quantity $|A||\hat{\mathbf{u}}_0(\mathbf{x}_m)|$ is closest to the measurement $|\check{u}_y(\omega_0)|$. To do so, we have represented in figure 4.2 (c,f), the measurement of the model $|A||\hat{\mathbf{u}}_0(\mathbf{x}_m)|$ as a function of the parameter $|A|$. The dashed horizontal line indicates the value linked to the measurement. We can see that the prediction of the amplitude by the model increases monotonically with $|A|$ and that it crosses the measurement value for a particular value of parameter $|A|$: $|A| = 1.71$ (resp. 2.23) for $Re = 60$ (resp. 100). These values slightly overestimate the actual DNS values, since $|A^{DNS}| = 1.68$ (resp. 2.13) for $Re = 60$ (resp. 100). This is due to the fact that the model only considers the fundamental harmonic and lacks a representation of the higher order harmonics. The effect of the latter modes on the mean-flow distortion is taken into account by a slight overestimation of the amplitude of the fundamental Fourier mode.

The reconstructed solution of the model 4.15 at the crossing point is respectively shown in figure 4.3 (resp. 4.4) for $Re = 60$ (resp. 100). We can see that the reconstructed streamwise component of the mean-flow (a), the reconstructed streamwise component of the fundamental mode at frequency ω_0 (b) and the reconstructed streamwise mean-component of the Reynolds stress tensor induced force (c) compare extremely well with their DNS counterparts (b,d,f). This is confirmed if we focus on the reconstruction of these quantities on the symmetry line $y = 0$: figure 4.5 shows that for both considered Reynolds numbers the streamwise profiles of mean-velocity, first-harmonic and force induced by Reynolds stress tensor closely follow the results obtained by the DNS.

3.3 Other measurements

Other quantities can be considered as inputs to tune the perturbation amplitude $|A|$. For example, this parameter can also be identified with mean quantities such as the mean-re-circulation length, the mean-drag or the mean-velocity field. In figure 4.6, we have assessed the quality of the reconstruction considering the above mentioned mean-measurements.

We can see that the mean measurements based on the re-circulation length and mean-drag vary monotonically (either increasing or decreasing) with $|A|$, showing that the stronger the fluctuations, the smaller the re-circulation bubble and the higher the average drag. As before, if we know the actual value of these mean-measurements, we can find a particular perturbation amplitude $|A|$ that matches the measurement value. It turns out that all of these predictions for $|A|$ are consistent and also correspond to the particular value of $|A|$ for which the error on the full mean-velocity field is minimal. However, since these are mean-quantities, it is not possible to predict the fundamental frequency. To asses

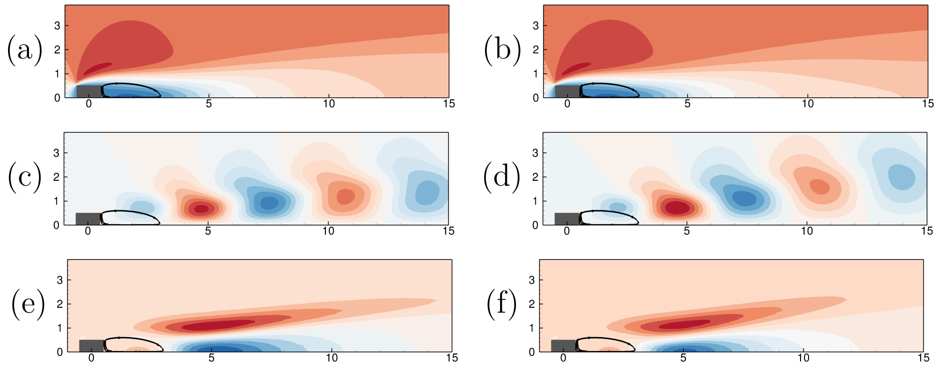

Figure 4.3: Comparison between reconstructed unsteady flows (left column) and DNS (right column) for the square-cylinder configuration at $Re = 60$. Stream-wise component of (a,b) mean flow velocities (c,d) first-harmonic velocities $A\hat{\mathbf{u}}_0$ and (e,f) mean force induced by the Reynolds stress tensor $\tilde{\mathbf{f}} = 2|A|^2\mathrm{Re}\{\hat{\mathbf{u}}_0 \cdot \nabla\hat{\mathbf{u}}_0^*\}$ with $|A| = 1.71$. The black curves delimit the re-circulation regions of the reconstructed mean flow. The frequency of the periodic flows is respectively $\omega_0 = 0.77$ and the amplitude of the Fourier mode at the fundamental frequency is $|A^{DNS}| = 1.68$.

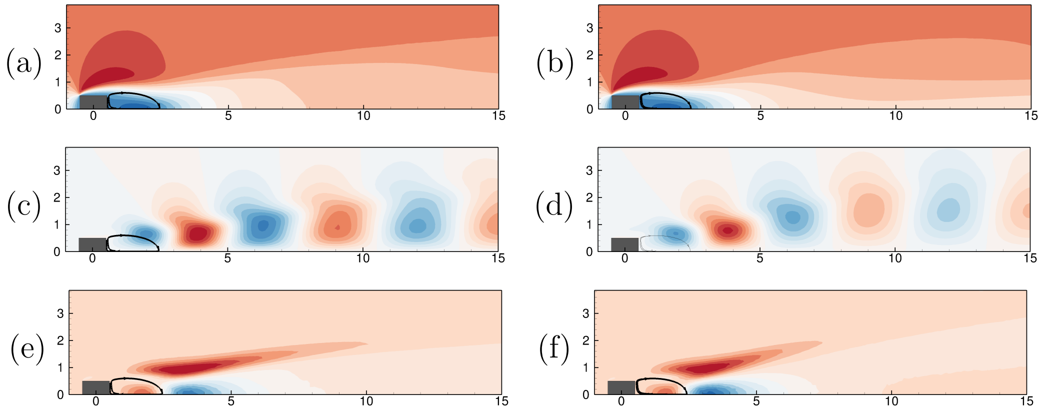

Figure 4.4: Comparison between reconstructed unsteady flows (left column) and DNS (right column) for the square-cylinder configuration at $Re = 100$. Stream-wise component of (a,b) mean flow velocities (c,d) first-harmonic velocities $A\hat{\mathbf{u}}_0$ and (e,f) mean force induced by the Reynolds stress tensor $\tilde{\mathbf{f}} = 2|A|^2\mathrm{Re}\{\hat{\mathbf{u}}_0 \cdot \nabla\hat{\mathbf{u}}_0^*\}$ with $|A| = 2.23$. The black curves delimit the re-circulation regions of the reconstructed mean flow. The frequency of the periodic flows is respectively $\omega_0 = 0.91$ and the amplitude of the fundamental frequency is $|A^{DNS}| = 2.13$.

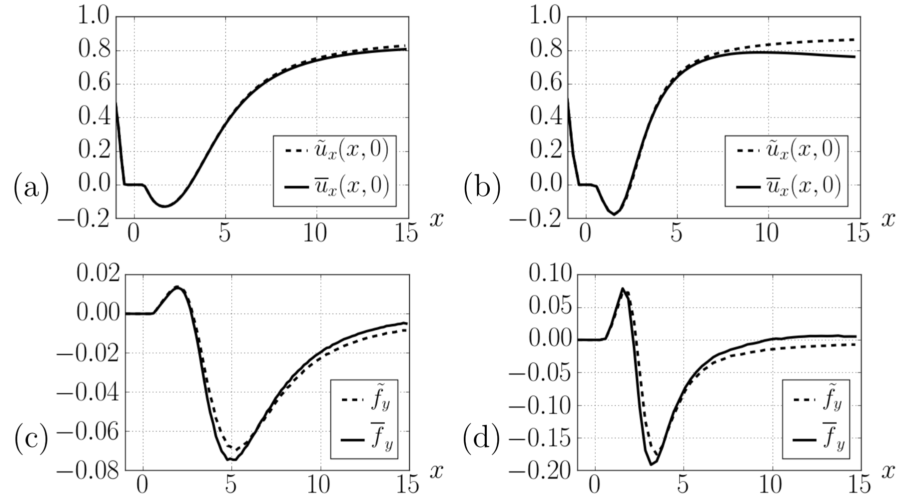

Figure 4.5: Details of reconstructed quantities for $Re = 60$ (a,c) and $Re = 100$ (b,d). Mean velocity field (a,b) and Reynolds-stress divergence forcing (c,d) $\tilde{\mathbf{f}}$ at the centerline $y = 0$.

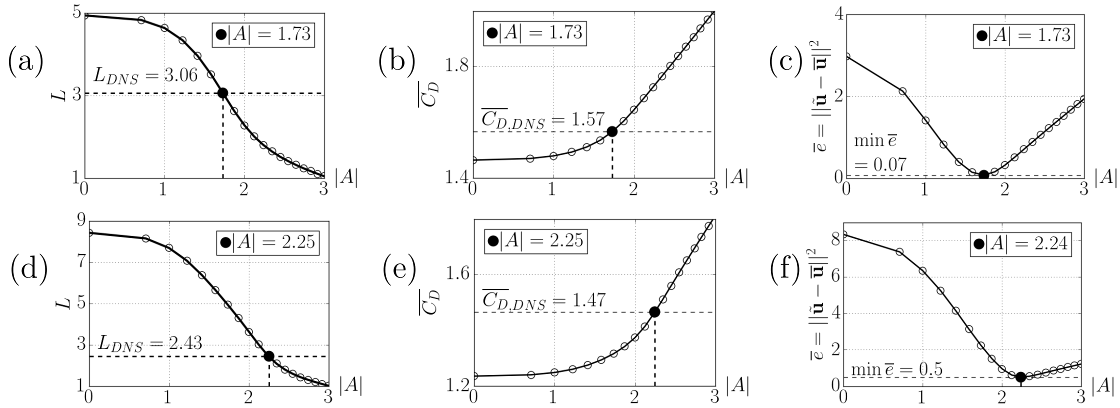

Figure 4.6: Other possible measures: for $Re = 60$, we have (a) the length of the recirculation bubble, (b) Mean drag coefficient (c) global error $\|\tilde{\mathbf{u}} - \overline{\mathbf{u}}\|^2$. For each different measure, we may have a different (but close) value of the estimated $|A|$. Idem for $Re = 100$, (b,d,f).

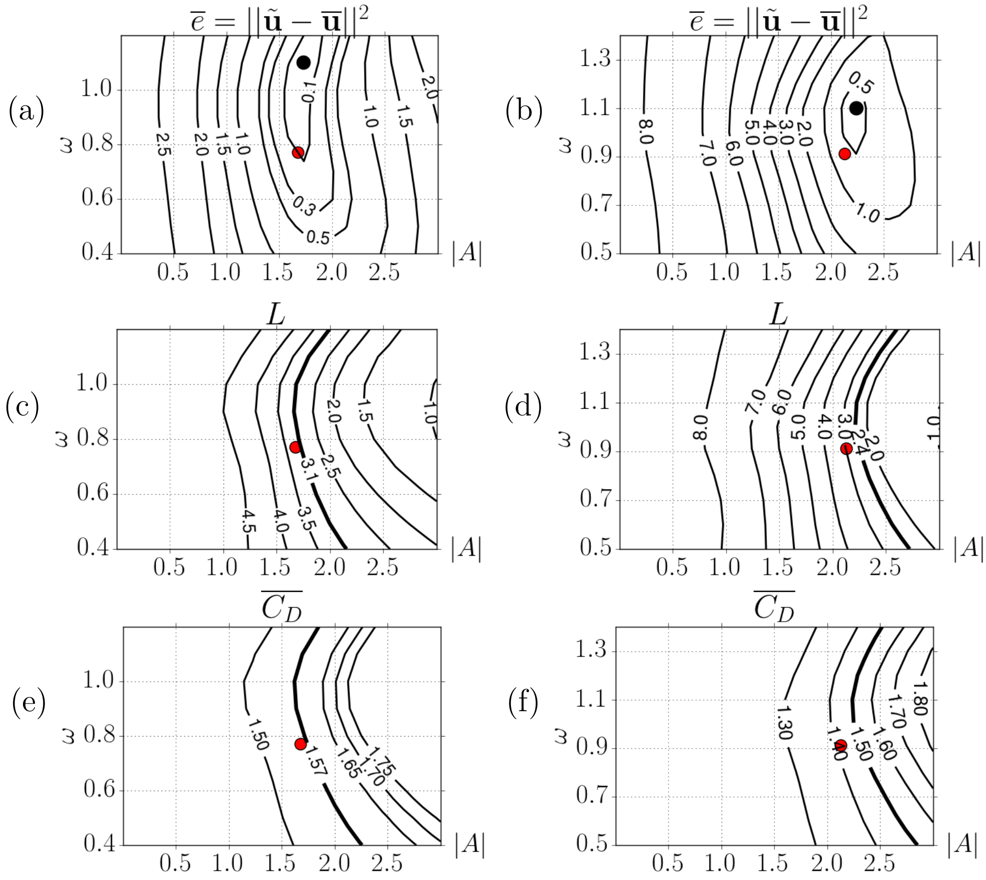

Figure 4.7: $(\omega, |A|)$ parametric study : (a,b) global error (minimal value marked by a black dot), (c,d) bubble's length and (e,f) mean drag. Results presented for $Re = 60, 100$, respectively. Reference (DNS) measured quantities marked in thick line and reference frequency and first-harmonic's amplitude $(\omega, |A|)$ represented by the red dot.

the dependency of those quantities with respect to the chosen fundamental frequency, we can perform a parametric study in the $(|A|, \omega)$ plane. Results are shown in figure 4.7 for Reynolds numbers 60 (left column) and 100 (right column).

We can see that the re-circulation bubble's length (2nd row) and the mean drag (3rd row), at a fixed frequency, exhibit a behavior that is similar to the results shown before with $\omega = \omega_0$: all curves are monotonic and they all eventually cross the actual value from the DNS. This means that we cannot uniquely determine a frequency from this data. If we observe now the global error in the mean-velocity field (1st row), we can see an elongated 'valley' in the frequency direction. This suggests that a precise knowledge of the chosen frequency is not crucial for the mean-flow estimation. The minima in these 'valleys' (black solid symbols) lie on points that do not necessarily correspond to the DNS frequency ω_0 (red solid symbols), since these minima are reached for $(|A|, \omega) = (1.68, 1.01)$ at $Re = 60$ and $(|A|, \omega) = (2.2, 1.08)$ at $Re = 100$. Note finally, that as discussed above, the amplitude parameter $|A|$ that enables a fit to the measured DNS quantities (black solid symbols and thick solid lines) always slightly overestimate the amplitude of the Fourier mode in the DNS. This is due to the fact that the model lacks a representation of the higher order harmonics, whose effect on the mean-flow is taken into account by a slight overestimation of the amplitudes of the fundamental Fourier mode.

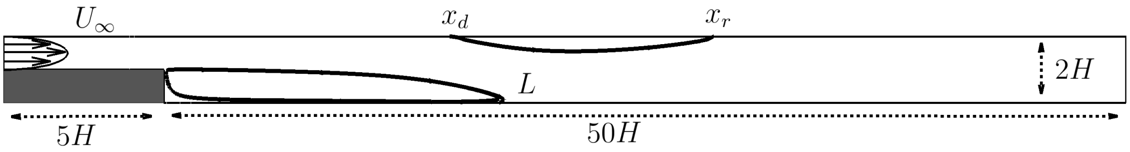

Figure 4.8: Sketch of the physical domain and inflow boundary condition for the backward-facing step case. The length of the main re-circulation bubble, the detachment and reattachment points are indicated in the figure and defined to be L, x_r and x_a, respectively.

At this point, we remind that a similar model was proposed by Mantic-Lugo et al. [2015] for the flow around a circular cylinder, where instead of using the Resolvent formalism the classic stability analysis was used for the approximation of the harmonics. However, instead of taking external data to provide a criterion for the choice of $|A|$, their model stopped for a value of $|A|$ such that the eigenmode was critically stable. This condition is justified whenever the effect of the second harmonics on the first is negligible. However, even if this method is similar to the here presented one, it is almost exclusive to "oscillator" flows, where its owns dynamics triggers the nonlinear saturation. For this reason, in the next section we show the generality of the proposed formalism on a backward-facing step flow, where the fluctuation is triggered by external noise, amplified through linear mechanisms (Marquet and Sipp [2010]).

4 Noise Amplifier Flows - Backward-facing step

In this section, we consider the (more challenging) backward facing step flow configuration. After a brief presentation of the configuration (§4.1), we assess the reconstruction procedure in two distinct noise environements. The first case (§4.2) corresponds to a quasi-periodic flow with two incommensurable frequencies, while the second one (§4.3) refers to a broadband white noise excitation.

4.1 Configuration

The configuration chosen for representing the noise amplifier type flow is the Backward-facing step described in Barkley et al. [2002], Hervé et al. [2012]. The Reynolds number, based on the height of the step H and the maximum velocity U_∞ of the Poiseuille flow imposed at the inlet, is fixed at $Re = U_\infty H/\nu = 500$. The inflow boundary condition is located at $5H$ upstream of the step and the outflow is located at $50H$ downstream, where, again, the boundary condition $(pI - Re^{-1}\nabla\mathbf{u})\cdot\mathbf{n} = \mathbf{0}$ is imposed. The other boundaries of the domain correspond all to solid walls on which a no-slip boundary condition is imposed. A sketch of the computational domain is given in Figure 4.8.

For the present flow conditions, there exists a steady-state solution that is globally stable (see Barkley et al. [2002], Blackburn et al. [2008]) but which exhibits strong transient mechanisms. The latter can be triggered by considering upstream stochastic noise, whose amplitude is tuned so that the dynamics of the perturbations is nonlinear and develops a

Table 4.2: Mesh dependency of steady and unsteady simulations for backward-facing step ($|A_1^{DNS}|$ and $|A_2^{DNS}|$ represent the amplitudes of the Fourier modes $\check{\mathbf{u}}(\omega_1 = 0.48)$ and $\check{\mathbf{u}}(\omega_2 = 0.7)$ in the case of 2-frequencies external noise) - chosen mesh for the computations in boldface

		Steady Solution			DNS - 2 frequencies				
	# triangles	L	x_d	x_r	L	x_d	x_r	$\lvert A_1\rvert$	$\lvert A_2\rvert$
Mesh 1	35000	10.875	8.688	17.508	7.455	5.765	10.390	0.448	0.282
Mesh 2	17000	10.870	8.685	17.508	7.461	5.761	10.405	0.461	0.292

mean-flow. In the present study, we pick:

$$\mathbf{f}(\mathbf{x}, t) = \begin{pmatrix} 0 \\ \phi(\mathbf{x}) \end{pmatrix} w(t), \tag{4.23}$$

where $w(t)$ contains the temporal dependency (and thus frequency content) of the noise and $\phi(\mathbf{x}) = e^{-|\mathbf{x}-\mathbf{x}_G|^2/\sigma^2}$ is a Gaussian function centered at $\mathbf{x}_G = (-0.5, 0.25)$ and of width $\sigma = 0.1$ (for further details, see Hervé et al. [2012]). The precise choice of $w(t)$ will be given below and will enable us to consider different flow environments.

In table 4.2, we provide some brief mesh-convergence results for the steady solution and the case of two incommensurable frequencies. Two meshes with different refinement levels have been generated. Both meshes are characterized by a higher refinement near the step location $\mathbf{x} = (0, 0)$ and a smooth coarsening towards the outflow (downstream of the main re-circulation bubble, which ends around $x \approx 10$). We can see that for both meshes all observed quantities are almost identical (up to $\sim 3\%$). For this reason, we found reasonable to keep the coarser mesh (Mesh 2) as the default mesh for all subsequent computations.

4.2 Upstream noise with two incommensurable frequencies: $\omega = 0.48$ and $\omega = 0.70$

In this section, we investigate the case where the dynamics of the backward-facing step flow is triggered with two distinct frequencies. The first frequency is the one that maximizes the Resolvent gain of the base-flow, $\omega_1 = 0.48$. The second one is chosen such that it lies in the zone where the Resolvent mode is still important and such that it is incommensurable with ω_1. The second frequency is chosen to be $\omega_2 = 0.70$. The time-signal presented in the external noise can thus be rewritten as:

$$w(t) = E_1 e^{\mathrm{i}\omega_1 t} + E_2 e^{\mathrm{i}\omega_2 t} + c.c. \tag{4.24}$$

where the amplitudes of the forcing are equal to $E_1 = E_2 = 0.5$.

Some characteristics of the unsteady flow are presented in figure 4.9 and in table 4.2. Compared to the base-flow, the mean-flow (figure 4.9 (a)) exhibits a shorter main re-circulation bubble ($L = 7.5$ versus 10.9 for the base-flow) and also a smaller secondary bubble at the top wall ($x_d = 5.8$ and $x_r = 10.4$). The spectra of the cross-velocity $\check{u}_y$, probed at $y = 0.2$ and $x = 1, 4$ and 15 show that as the perturbation is convected downstream, higher-order harmonics which are different from the two original frequencies

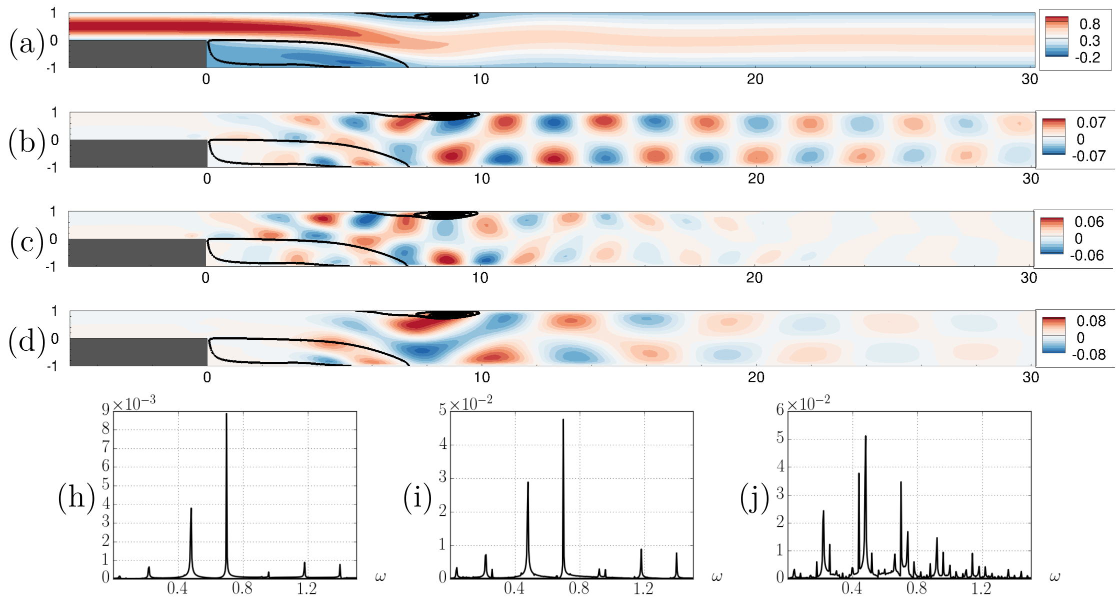

Figure 4.9: (a) Exact time-averaged flow (with $L = 7.5$, $x_d = 5.8$ and $x_r = 10.4$), (b) Fourier mode at $\omega = 0.48$, (c) at $\omega = 0.70$, (d) at $\omega = 0.70 - 0.48 = 0.22$ and spectrum of the cross-spectra field at (e) $(x, y) = (1, 0.2)$, at (f) $(x, y) = (4, 0.2)$ and (g) $(x, y) = (15, 0.2)$.

are generated. The Fourier modes of the streamwise $\check{u}_x$ and crossstream velocity $\check{u}_y$ fields are shown in figures 4.9 (b,c,d,e,f,g) for three different frequencies, ω_1, ω_2 and $\omega_2 - \omega_1 = 0.22$.

The reconstruction is achieved using only the two incommensurable frequencies ω_1 and ω_2, even if the signals contain many more peaks. This is justified by the fact that these two peaks are stronger than those generated by the nonlinear interactions in almost all measured signals. In the following, we will denote $|A_1|$ and $|A_2|$ the amplitudes of the normalized Resolvent modes. Since we only have two tuning parameters, we can easily explore the full parameter-space by performing a (rather coarse, but enlightening) $(|A_1|, |A_2|) \in \mathbb{R}^2$ parametric study. In the same time, we will monitor the evolution of the main quantities (measurements) that may be used to tune these two parameters. The results are presented in figure 4.10. Each figure is relative to a different measure, the thick solid line in figs (b-h) denoting the values measured in the DNS and the red solid symbol designating the true DNS amplitudes $(|A_1|, |A_2|)$ of the normalized Fourier modes.

Focusing first on the mean-flow measurements (figs a-d), we observe that the DNS re-circulation bubbles' length and positions (L, x^d,x^r) can be recovered by tuning either $|A_1|$ and / or $|A_2|$. This is consistent with the elongated 'valley' obtained for the total mean-flow error $\overline{e} = ||\tilde{\mathbf{u}} - \overline{\mathbf{u}}||^2$ (fig. a). This means that any values of $(|A_1|, |A_2|)$ within the valley will produce correct mean-flow bubbles' sizes, the minimum being obtained for $(|A_1|, |A_2|) \approx (0.77, 0.44)$. If we compare these values with those of the DNS, we see that the model needs to overestimate the amplitudes of the modes to have a good representation of the mean-flow. This was also the case for the flow over the square cylinder, however, to a less extent. This can easily be explained by the fact that the model is based on only two-frequencies while the reference flow is characterized by many more peaks. Hence, the model needs to compensate the absence of these peaks by overestimating those at

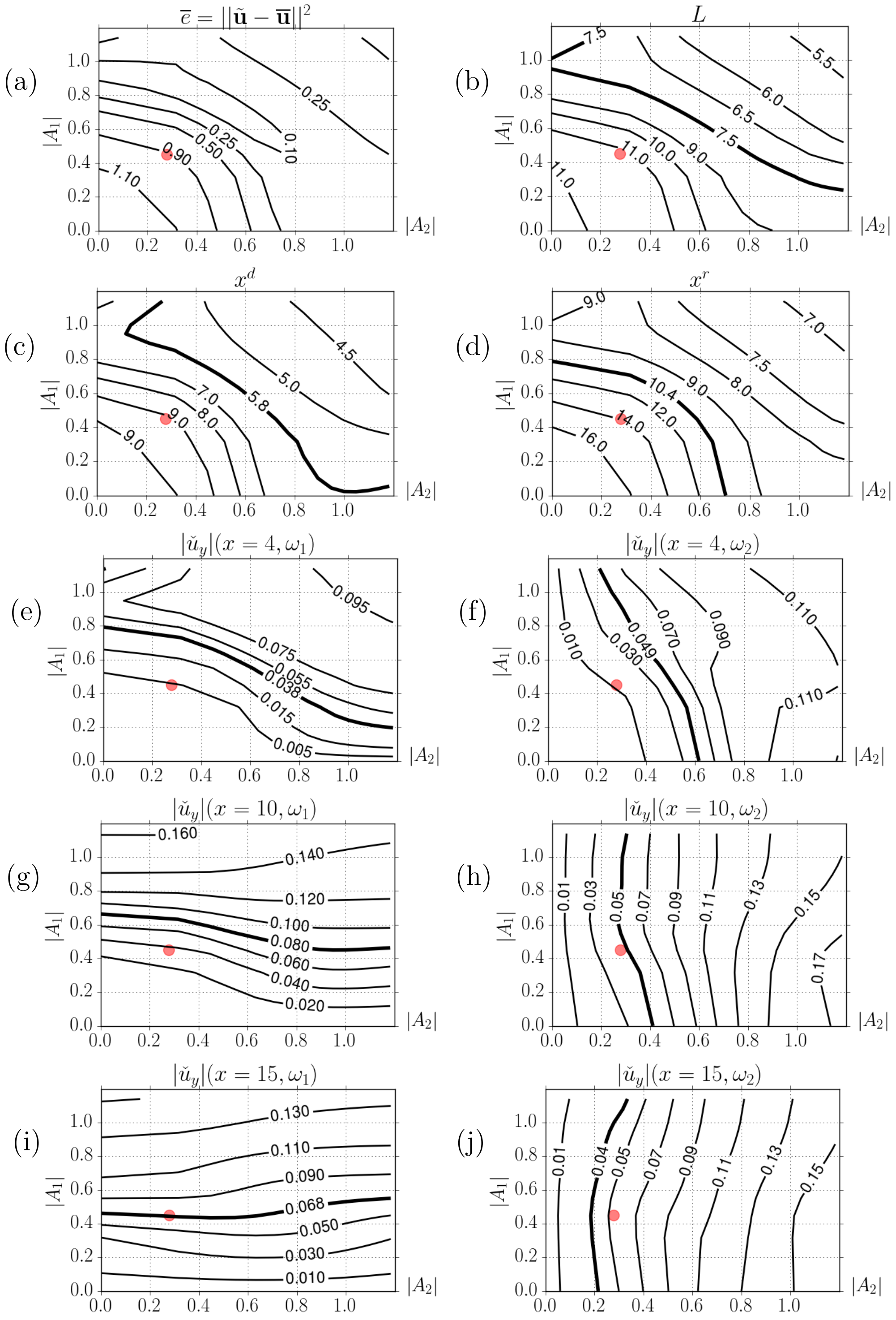

Figure 4.10: $(|A_1|, |A_2|)$ parametric study: (a) total L^2 error of the mean-flow $\overline{e} = ||\tilde{\mathbf{u}} - \overline{\mathbf{u}}||^2$, (b) main re-circulation bubble's length L, (c) detachment position x^d and (d) reattachment position x^r of the secondary re-circulation bubble. The mean-flow is optimally reconstructed at $(|A_1|, |A_2|) = (0.77, 0.44)$, while the actual energy of the Fourier modes at those frequencies are $(|A_1|, |A_2|) = (0.45, 0.28)$ (red dots). Fourier coefficients for ω_1 and ω_2 at three locations of the flow at (e,f) $x = 4$, (g,h) $x = 10$ and $x = 15$. Thicker lines correspond to exact quantities, coming from the DNS.

ω_1 and ω_2. The reconstructed flowfield (mean-flow and Fourier modes) for the optimal parameters $(|A_1|, |A_2|) = (0.77, 0.44)$ corresponding to the minimum of $\bar{e}$ and their exact DNS counterparts are presented in figure 4.11. We can see that, even if the energy of the mode for ω_1 is over-predicted, its spatial structure is well represented by the resolvent mode. This is also the case for ω_2, especially in the high-amplitude region close to the bubbles. Comparing the model's Reynolds stress-tensor ($\overline{u'_x u'_y}$ component) with the one of the reference flow, we see that they have similar spatial structure and amplitude, confirming that the contribution of harmonics that are not considered in the model may be accounted for by overestimating the amplitudes of the modes retained in the model.

If we now consider measures involving the fluctuation amplitudes at given points, figures 4.10(e-j) indicate that the reconstructed fluctuation amplitude at frequency ω_1 (resp. ω_2) exhibits nearly a linear behaviour as a function of $|A_1|$ (resp. $|A_2|$) and does only weakly depend on $|A_2|$ (resp. $|A_1|$), especially when $x \geq 10$ (see figs (g-j)). Deviations from this behavior observed with the measurements at $x = 4$ may be attributed to the fact that, at those considered frequencies, the resolvent modes are only energetic if the re-circulation bubble's length L is sufficiently short. This can be seen through the similarity of the plots in figures 4.10 (b) and (e). Considering the measurements at $x = 15$ (see figs. i,j), the fluctuation amplitude at ω_1 obtained in the DNS (thick solid line in fig. i) is best-recovered for $|A_1| \approx 0.4$, while the best-fit for the amplitude at frequency ω_2 is obtained for $|A_2| \approx 0.2$ (see fig. j). Those amplitudes are not far from the DNS ones (red solid symbols). However, the mean-flow is not well represented with these parameter values (see figs a,b,c,d). This means that one cannot, with the two frequencies model, well represent at the same time the quantities related to the mean-flow and its fluctuations. This nevertheless points out an interesting feature: if one is interested only in the mean-flow and Reynolds-stress tensor reconstruction, one may consider a model with few frequencies and tune the few parameters by mean-flow measurements (such as L, x_r or x_d). This may be considered as a very efficient way to perform mean-flow data-assimilation with extremely sparse data. However, if, in addition to the mean-flow, one desires to have a good estimate of the fluctuation, more frequencies need to be included in the model. This will be shown in the next section dealing with the BFS forced with white-noise external perturbation.

4.3 Upstream noise with broadband white noise forcing

We consider now the case where the dynamics is triggered with a (time) white-noise $w(t)$. Since the noise-amplification mechanism associated to the linearized Navier-Stokes equations in the case of the present configuration is a broadband low-frequency band, we expect the output fluctuations to exhibit the same feature. This is in stark contrast with respect to the cylinder flow and the two-frequencies backward-facing step flow cases discussed before. The harmonic averages used above, that extract Koopman modes, now converge to zero for all frequencies. Following Arbabi and Mezić [2017], a more suited quantity for the frequency analysis consists in considering the two-time auto-correlation function. For a given point $\mathbf{x}$ in space, this function can be given by:

$$\mathbf{T}_{\mathbf{x},\mathbf{x}}(t) = \mathbb{E}[u'(\mathbf{x},\tau)u'(\mathbf{x},\tau+t)] = \lim_{T\to+\infty} \frac{1}{T}\int_0^T u'(\tau)u'(\tau+t)d\tau. \tag{4.25}$$

Since no Koopman modes (i.e. quasi-periodic component) exist in the present configuration, this function $\mathbf{T}_{\mathbf{x},\mathbf{x}}(t)$ tends to 0 for large t, so that it is square integrable. Hence,

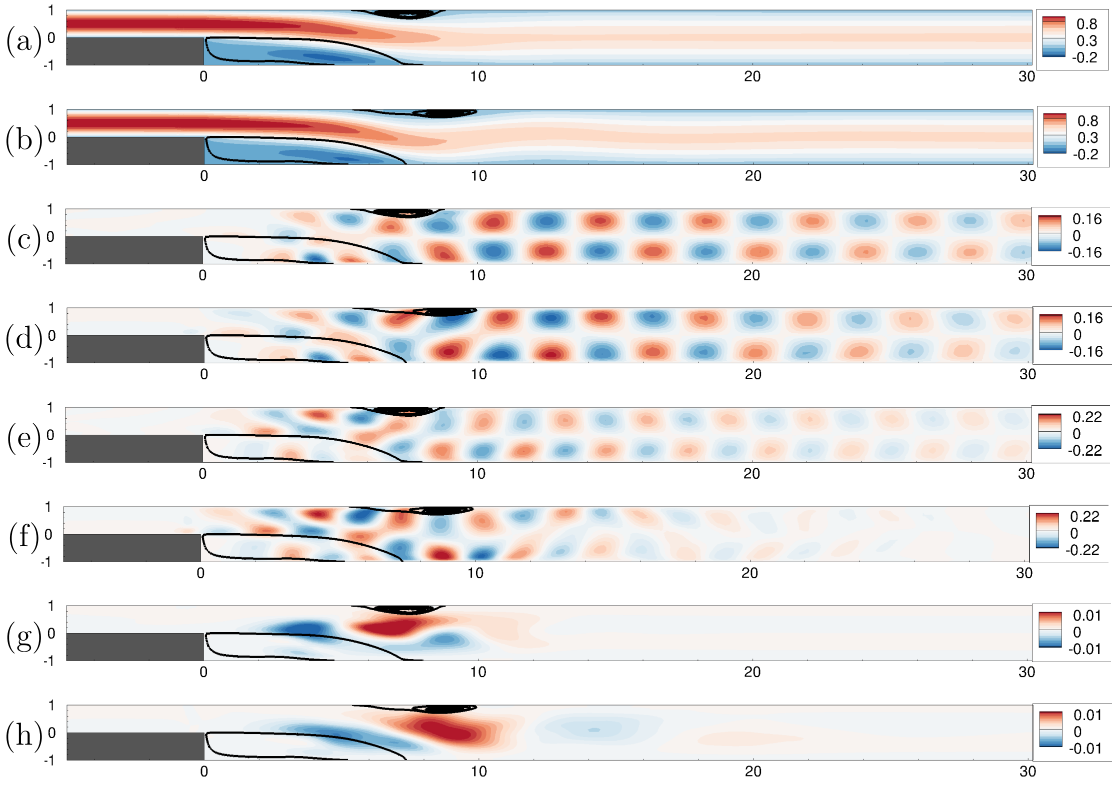

Figure 4.11: Reconstructed fields with optimal parameters $(|A_1|, |A_2|) = (0.77, 0.44)$ minimizing $\bar{e}$: reconstructed (a) and exact (b) mean flows, reconstructed (normalized) Fourier modes at $\omega_1 = 0.48$ (c) and (e) $\omega_2 = 0.7$ and reference ones (d,f). Reynolds stress tensor $\overline{u_x' u_y'}$ for the model (g) and DNS (h). Identical colour-bars have been used for reconstructed and DNS fields to help comparison.

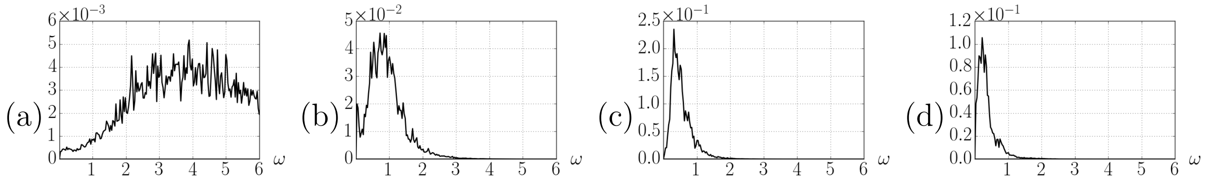

Figure 4.12: DNS results for white-noise external forcing: Power-Spectral Density (PSD) of the cross-wise velocity field at (a) $(x, y) = (0, 0.2)$, (b) $(x, y) = (4, 0.2)$, (c) $(x, y) = (10, 0.2)$ and (d) $(x, y) = (15, 0.2)$, obtained with the Welch algorithm from a 5000 non-dimensional time-interval, subdivided in 99 bins with 50 % overlap

one may express the tensor under the form of a Fourier decomposition:

$$\mathbf{T}_{\mathbf{x},\mathbf{x}}(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}} e^{i\omega t} d\omega, \tag{4.26}$$

where the Fourier modes may be obtained from:

$$\mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}} = \lim_{T \to +\infty} \frac{1}{T} \mathbb{E}[|\mathcal{F}_T(u')(\omega)|^2]. \tag{4.27}$$

For $t = 0$, we obtain that the time-averaged energy of the signal is equal to:

$$\mathbf{T}_{\mathbf{x},\mathbf{x}}(0) = \lim_{T \to +\infty} \frac{1}{T} \int_0^T u'(\tau)^2 d\tau = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}} d\omega = \frac{1}{\pi} \int_0^{+\infty} \mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}} d\omega, \tag{4.28}$$

which establishes that $\mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}}$ is also the Power Spectral Density (PSD) of the signal.

Coming back to the description of the flow unsteadiness, if we probe the resulting unsteady flow (presented in figure 4.12) at several flow locations, we can see that, as expected, the further we go downstream, the stronger and the more dominant are the low-frequencies in the PSD, shown in figure 4.12 (a-d).

4.3.1 Objective functional

We now turn to the model used for the reconstruction of the flow using partial measurements. The auto-correlation function in the case where the fluctuation field is approximated by eq. (4.16) reads:

$$\mathbf{T}_{\mathbf{x},\mathbf{x}}(t) = \sum_{j=1}^{N} |A_j|^2 |\hat{\mathbf{u}}^j_0(\mathbf{x})|^2 \left(e^{-\mathrm{i}\omega_j t} + e^{\mathrm{i}\omega_j t} \right). \tag{4.29}$$

For $t = 0$, the time-averaged energy of the signal therefore reads:

$$\mathbf{T}_{\mathbf{x},\mathbf{x}}(0) = \sum_{j=1}^{N} 2|A_j|^2 |\hat{\mathbf{u}}^j_0(\mathbf{x})|^2 \tag{4.30}$$

We therefore note that the PSD in the reference signal is dense (see ((4.28)) and figure 4.13 (a)), while the PSD in the model (see ((4.30)) and figure 4.13 (b)) is discrete. For the model, that contains a finite number of discrete Fourier modes, to be representative of the whole energy contained in the DNS, the two PSDs need to match. To establish this

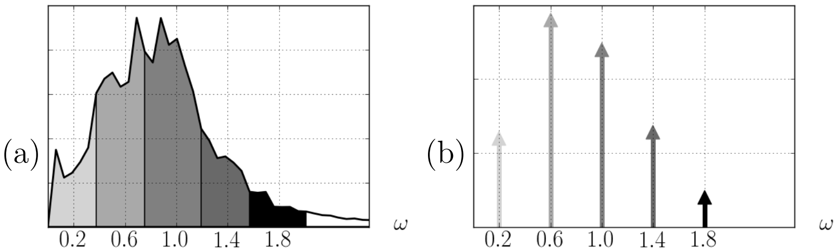

Figure 4.13: Power Spectral Density (PSD) for a broadband flow (a). Its integral on five ($N = 5$) frequency intervals are compared with the discrete-in-frequency energies coming from the model (b).

connection, we associate each term $2|A_j|^2|\hat{\mathbf{u}}_0^j(\mathbf{x})|^2$ (in equation (4.3.1)) to the corresponding integral of $\mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}}$ in frequency domain (in equation (4.28)). To do so, we discretize the frequency space by considering $\{\Omega_j\}_{j=1,N+1}$, where $\omega_j = (\Omega_j + \Omega_{j+1})/2$ falls into the center of the interval (Ω_j, Ω_{j+1}). In our case, we fix the first frequency to be $\Omega_0 = 0$ and $\Omega_{N+1} = 2$, since Fourier modes corresponding to frequencies higher than $\Omega_{N+1} = 2$ are weak and located upstream only (figure 4.12), near the location of the external forcing. Moreover, we pick a uniform frequency discretization such that $\Delta\omega = \Omega_{j+1} - \Omega_j$ for all j. This choice is in accordance with the actual PSDs of the DNS signal, in which only one 'bump' at low frequencies is visible. However, in the case where several 'bumps' in frequency would exist, a non-uniform distribution of Ω_js could be considered. The energy of the signal at point $\mathbf{x}$ contained within the frequency band (Ω_j, Ω_{j+1}) is:

$$I_{\mathbf{x},\omega_j} = \int_{\Omega_j}^{\Omega_{j+1}} \mathbf{T}^{\omega}_{\mathbf{x},\mathbf{x}} d\omega = 2\pi|A_j|^2|\hat{\mathbf{u}}_0^j(\mathbf{x})|^2 \tag{4.31}$$

With this, one can, as previously, compare the energy of the fluctuation at a measure point $\mathbf{x}_m$ predicted by the model to the one measured. The measure of the error associated to this quantity is given by:

$$J_{\omega_j,\mathbf{x}_m} = \left(\frac{2\pi|A_j|^2|\hat{\mathbf{u}}_0^j(\mathbf{x})|^2}{I_{\mathbf{x},\omega_j}} - 1\right)^2. \tag{4.32}$$

We note that this new objective functional to be minimized, contrarily to the mean-flow related ones (such as L, x_r, x_d), should be capable to discriminate efficiently between the different frequency contributions associated to each $\{|A_j|\}_{j=1,N}$. Indeed, following the recommendations of Beneddine et al. [2016], we will choose the measurement points at regions where the Resolvent modes are energetic to maximize the robustness of the reconstruction. The spatial distribution of the energy of the Resolvent modes is shown in figure 4.14. We can see that for high frequencies ($\omega \geq 1$), the resolvent modes are concentrated near the step location above the main re-circulation bubble, where for low frequencies the Resolvent modes tend to be more spread towards the outflow. For this reason, we choose to consider for data-assimilation the frequency content of different points of the domain. For high frequencies, we will consider the signal at $x = 4$, while for low frequencies, we consider two different possible points $x = 10, 15$. This leads to two cost functions to quantify the error for the data-assimilation procedure:

$$J^1 = \sum_{\omega_j<1} J_{\omega_j,x=10} + \sum_{\omega_j\geq 1} J_{\omega_j,x=4}, \quad J^2 = \sum_{\omega_j<1} J_{\omega_j,x=15} + \sum_{\omega_j\geq 1} J_{\omega_j,x=4} \tag{4.33}$$

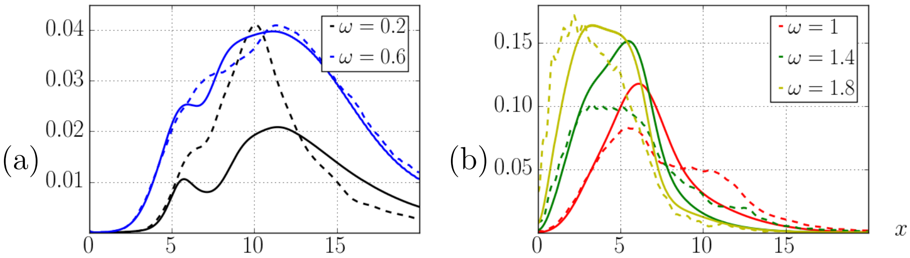

Figure 4.14: Spatial distribution of the energy, evaluated on the line $y = 0.2$, of the resolvent modes on the mean flow (solid lines) and Fourier modes (dashed lines) at several frequencies (both modes are normalized).

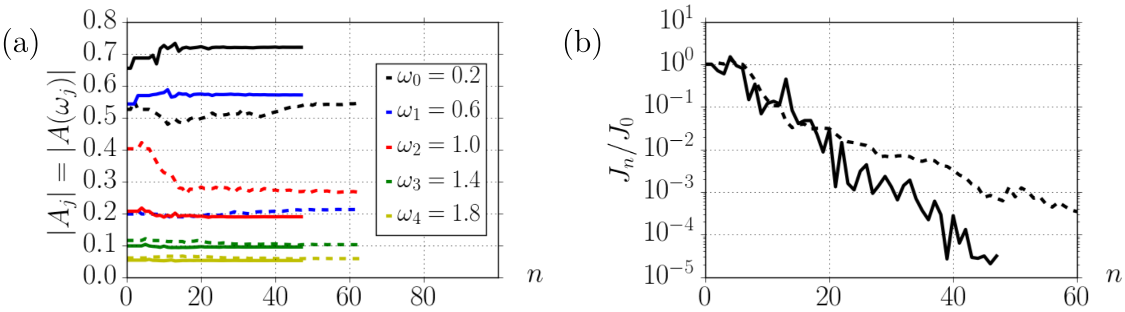

Figure 4.15: Convergence of the amplitudes $|A_j|$ and the cost functional J_n/J_0 for the cost functional J^1 (solid lines) and J^2 (dashed lines). Case $N = 5$.

4.3.2 Optimization procedure

Now that the cost functionals have been precisely defined, we can proceed with the optimization procedure. We note that this white-noise case is richer in frequency domain. For this reason, it is interesting to evaluate how the assimilated solution will change as a function of the number of considered frequencies N. For that purpose, we will consider the cases $N = 2, 5, 10$. We can see now that a parametric study (similar to the one performed for the two-frequencies case) is no longer feasible. Hence, we need to resort to an automatized optimization algorithm, where the cost functional $J(|A_j|)$ is minimized. For simplicity, here we choose a gradient-free algorithm, which provides good performance if the number of frequencies is not too large. We picked here the COBYLA algorithm, a simplex-based algorithm, where the optimal condition is evaluated within a $N-$simplex evolving in the parameter space. For further details on this algorithm, see Powell [1994]. To speed up the convergence for the cases $N = 5$ and 10, we first perform the optimization for the case $N = 2$, which is quick. Then, we initialize the $|A_j|$ coefficients for $N = 5$ by computing the Resolvent modes based on the mean-flow determined with $N = 2$ and by setting $|A_j| = \sqrt{I_{\mathbf{x}_m,\omega_j}/2\pi}/|\hat{\mathbf{u}}_0^j(\mathbf{x}_m)|$, where $\mathbf{x}_m$ is the measurement point. We repeat the same procedure to initialize the optimization for $N = 10$ based on the mean-flow obtained for $N = 5$.

4.3.3 Reconstruction results

In figure 4.15 we represent the convergence of the optimization process for the case $N = 5$ for two cost functionals described. We can see that starting from the case $N = 2$

Table 4.3: Reconstructed parameters in function of number of frequencies considered N.

	L		x_d		x_r		$\bar{e} = \|\|\tilde{\mathbf{u}} - \bar{\mathbf{u}}\|\|^2$		$\mathcal{A} = \sqrt{\frac{1}{T}\int_0^T \int_\Omega \|\mathbf{u}'\|^2}$	
DNS	7.07		5.62		9.99		–		1.07	
	J^1	J^2	J^1	J^2	J^1	J^2	J^1	J^2	J^1	J^2
$N = 2$	7.42	9.39	5.91	7.57	9.31	13.34	0.10	0.70	1.20	0.81
$N = 5$	6.25	8.67	5.24	6.93	8.16	12.01	0.05	0.40	1.33	0.92
$N = 10$	6.02	7.78	5.35	6.77	7.80	10.40	0.09	0.20	1.41	1.11

provides indeed a good initial guess, especially for J^1, since the optimization did not drift apart too much from the initialization and the optimisation did not take more than 20 iterations to converge up to a precision of $J_n/J_0 \approx 10^{-2}$. For the cost functional J^2, the convergence needed slightly more iterations, mainly because of a poor estimation of the initial amplitude for $\omega_2 = 1$. Furthermore, we observe that the estimation of the amplitude for $\omega_0 = 0.2$ and $\omega_1 = 0.6$ is considerably lower than those predicted for J^1. The rate of convergence of the cost functional J_n/J_0 is almost exponential for both cost functionals considered.

In table 4.3, we have compared, for the 2 cost functionals and 3 numbers of frequencies N, some overall features of the reconstructed flow with those of the DNS. Those features are the detachment and reattachment points, the mean-flow error and the the overall amplitude of the fluctuations, defined as $\mathcal{A} = \sqrt{\frac{1}{T}\int_0^T \int_\Omega |\mathbf{u}'|^2}$, which is equal to $\mathcal{A} = \sqrt{\sum_j 2|A_j|^2}$ for our model. We could observe that, by refining the frequency discretization (increasing N), the value of $\mathcal{A}$ increased for both cost functionals. This increase of this value lead to a mean-flow that was more saturated, leading to a decrease of almost all detachment / attachment points, with exception of x_f for the cost functional J^1. Indeed, this variable seemed less sensitive to the increase in $\mathcal{A}$. Globally, the cost functional that met the best value for $\mathcal{A}$ was J^2, whereas for best mean-flow approximation was J^1. This is explained, in part, by the mismatch between the resolvent modes and the actual Fourier (Spectral Proper Orthogonal Decomposition, SPOD, as it will be explained shortly) modes. Indeed, in Figure 4.14, it is seen that the Resolvent modes only approximately reconstruct the Fourier modes and that significant discrepancies may exist. For example, for $\omega = 0.2$, at the point used to tune the amplitudes ($x = 10$ or 15), the amplitude of the normalized Resolvent mode is either lower ($x = 10$) or higher ($x = 15$) than the normalized Fourier mode. In particular, tuning the Resolvent amplitude with the value of the Fourier mode at $x = 10$ will lead to an over-prediction of the fluctuation level and thus higher values of $\mathcal{A}$.

We turn now our attention to the comparison of the assimilated fields with the ones coming from the DNS, namely the mean-flow and the Fourier modes. The definition of the Fourier modes is here based on the more general two-point, two-time correlation tensor (see Towne et al. [2018]), whose components are the (time) correlation between the state vector at two distinct points in space:

$$\mathbf{T}_{i,j}(t) = \mathbb{E}[u'(\mathbf{x}_i, \tau) u'(\mathbf{x}_j, \tau + t)].$$

Since there are no Koopman modes, all components of the correlation tensor tend to 0 when t becomes large, so that one may define the Fourier transform:

$$\mathbf{T}^{\omega}_{i,j} = \mathcal{F}_{\infty}(\mathbf{T}_{i,j})(\omega). \tag{4.34}$$

The diagonal entries (for which $\mathbf{x}_i = \mathbf{x}_j$) of this tensor correspond to the PSDs defined earlier. This tensor can also be expressed in terms of the Fourier transforms of $\mathbf{u}'$ as:

$$\mathbf{T}^{\omega}_{i,j} = \lim_{T \to +\infty} \frac{1}{T} \mathbb{E}[\mathcal{F}_T(u')(\mathbf{x}_i, \omega) \mathcal{F}^*_T(u')(\mathbf{x}_j, \omega)]. \tag{4.35}$$

The explicit computation of this tensor from the DNS data is performed in the following way: we divide our total time-signal into N_b 'bins' (with 50% of overlap), each one with a (sufficiently large) time-duration T, and we perform the average of each Fourier transform of each 'bin':

$$\mathbf{T}^{\omega} \approx \frac{1}{T} \frac{1}{N_b} \sum_{k=1}^{N_b} \mathcal{F}^k_T(u')(\omega) \mathcal{F}^{k,*}_T(u')(\omega) \tag{4.36}$$

To extract the most coherent structure over all bins, ie the Fourier mode, we look for the largest eigenvalue of $\mathbf{T}^{\omega}$:

$$\mathbf{T}^{\omega} \mathbf{y}^{\omega} = (\lambda^{\omega}_0)^2 \mathbf{y}^{\omega}_0, \quad ||\mathbf{y}^{\omega}_0|| = 1. \tag{4.37}$$

The Fourier mode is then:

$$\check{\mathbf{u}}^{SPOD}(\omega) = \lambda^{\omega}_0 \mathbf{y}^{\omega}_0. \tag{4.38}$$

In figure 4.16 we compare the reconstructed fields for J^2 and $N = 5$ with the DNS. We can see that the mean-flow is well represented, as well as the Fourier modes. This shows the validity of the presented data-assimilation procedure, where, from point-wise time-resolved sensors, we manage to estimate the whole steady and unsteady features of a given flowfield.

5 Conclusions and Perspectives

In this chapter, we presented a data-assimilation technique based on Resolvent modes around the mean-flow. The technique was based on a model composed of a mean-flow equation coupled with Resolvent modes corresponding to left singular vectors of the Resolvent operator. The model is not closed since the amplitudes of those modes are unknown and need to be tuned with the help of measurements. This technique is essentially different from what was presented in the second and third chapters since the model has typically few degrees of freedom to be tuned, corresponding to the energies of each mode. This feature allows one to assimilate also very sparse data, without, for example, the need for supplementary penalization or regularization. This data can come from various types of measurements, such as mean-flow characteristics (length of separation) or point-wise time-resolved information (easily obtainable with, for example, a Pitot probe). Furthermore, the optimization / tuning procedure is much simpler, not involving necessarily an adjoint solver. This technique was applied on an oscillator flow, whose frequency content was essentially monochromatic. For this reason, the data-assimilation results showed that we are indeed capable of recovering not only the mean-flow but also the fluctuation around it. This technique was also applied on a backward-facing step where the dynamics is driven by an external forcing. Two cases were considered: the first where the forcing

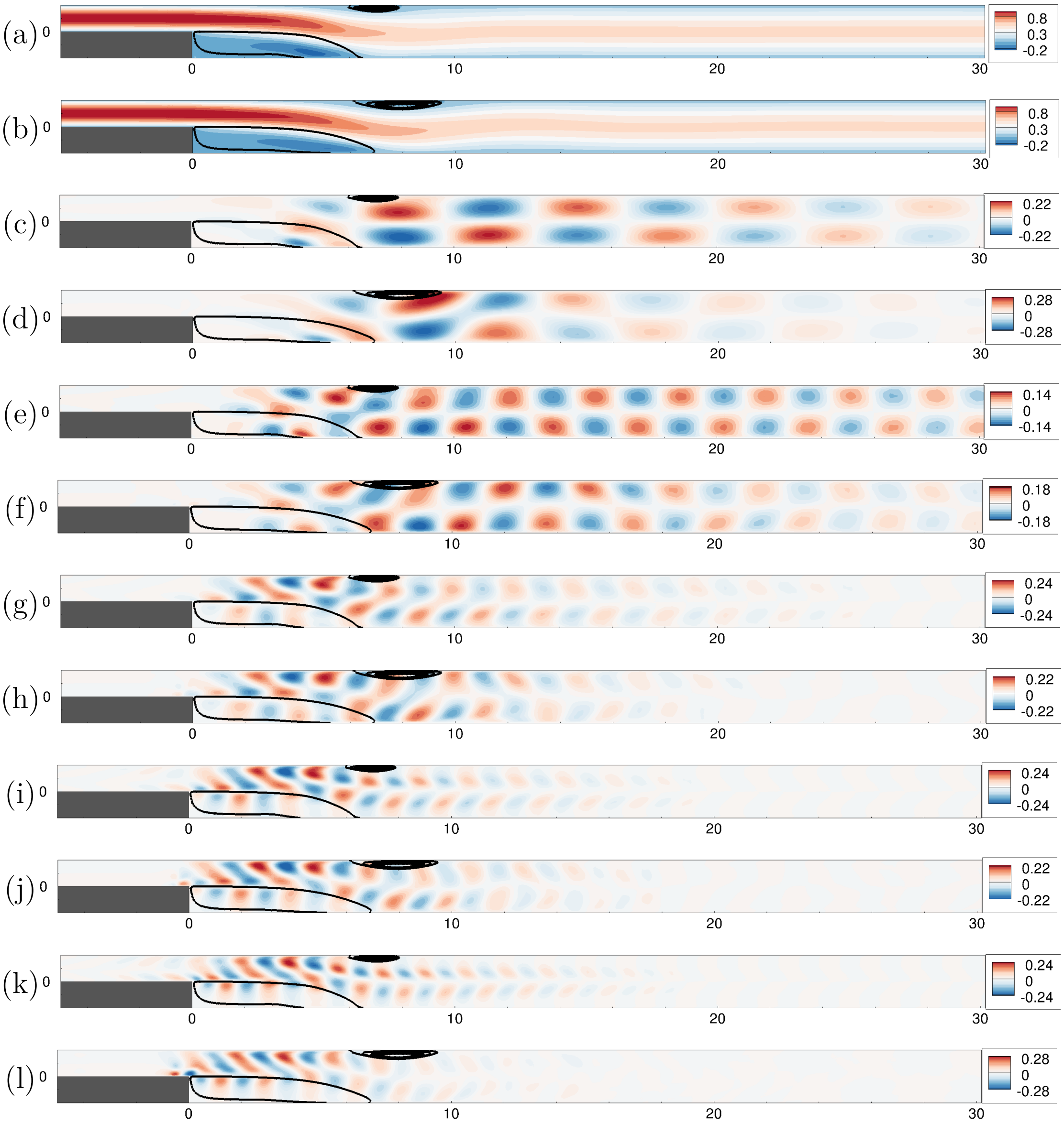

Figure 4.16: Comparison between reconstructed (a,c,e,f,i,k) and reference (b,d,f,h,j,l) fields: mean-flow (a,b) and (normalized) Fourier modes at $\omega = 0.2$ (c,d), $\omega = 0.6$ (e,f), $\omega = 1.0$ (g,h), $\omega = 1.4$ (i,j), $\omega = 1.8$ (h,l).

spectrum was discrete in frequency and composed of two incommensurable frequencies and, the second, continuous, with white-noise broadband forcing. In the first case, even if the data from the DNS showed more than those two peaks in frequency, we chose to use a two-frequencies model. With this model, we could only obtain partial information about the complete flow, meaning that we could either accurately reconstruct the unsteady features (the amplitudes of the actual Fourier modes) or the mean-flow-related quantities. A natural way of completing the model and solving those issues would be to add higher-order harmonics as done in Symon et al. [2019a]. However, this would lead to a much more complex model than the one proposed by them, since in their work the mean-flow equation was not solved, and the mean-flow was estimated from variational data-assimilation and frozen for all subsequent Resolvent analysis. The second case dealt with white-noise external perturbation, that lead to a dense (or broadband) frequency content. This is an essentially different test-case since before, the nature of the fluctuation was discrete in frequency (given by the Koopman modes, see Arbabi and Mezić [2017]), and so was our model. For this reason, a connection between this broadband fluctuation and our model needed to be established. This connection is based on the integral of the Power-Spectral Density, which is a finite quantity either for the actual reference flow and for our model. By enforcing our model to satisfy those quantities at certain point locations, where the velocity field was measured, we were able to obtain a solution that reproduced this time both mean-flow quantities and the fluctuation as well. The reconstruction results show an increase of the fluctuation energy content when the frequency-refinement is increased. We could also see that the assimilated solution depends somewhat on the spatial location of the measurements. This comes from a mismatch between the SPOD modes and the Resolvent ones. However, as pointed out by Towne et al. [2018], those two modes are the same if the projection coefficient of the actual nonlinear forcing onto the Left-Singular Vectors (Resolvent-forcing) are uncorrelated between the optimal mode and its suboptimal ones. This suggests that, even in this broadband case, nonlinear interactions may be produced and may be important.

Chapter 5

Mean-flow analysis for the mixed periodic/broadband fluctuations on turbulent flows

1 Introduction

The mean-flow stability analysis was formally introduced for transitional flows, particularly the flow around a circular cylinder, where the unsteady behavior was produced by a global instability that saturates into a periodic limit cycle. In this case, Barkley [2006] showed that the eigensystem formed by the linearized Navier-Stokes equations around the time-averaged (mean) flow lead to an eigenmode whose eigenvalue was (almost) purely imaginary, predicting very well the nonlinear frequency of the periodic limit-cycle for a range of Reynolds numbers near the bifurcation. Further applications (Turton et al. [2015b]) and theoretical elucidations (Sipp and Lebedev [2007b], Mezić [2013]) lead to the conclusion that, if the nonlinear forcing acting on the, say, first harmonic equation was negligible, this mean-flow analysis could reproduce the main features of the nonlinear limit-cycle, namely the frequency (from the eigenvalue) and the structure of this harmonic (from the eigenmode). In the laminar/transitional regime this concept has widely been successfully employed, for example for control purposes (Sipp et al. [2010]) or for designing models for the self-sustained dynamics of periodic limit-cycles (Mantič-Lugo et al. [2014]). In a turbulent context, for example in bluff-body configurations, this concept has also been applied with success in determining the frequency and the spatial structure of the vortex-shedding mode (see Mettot et al. [2014b], Meliga et al. [2012].

Although the mean-flow stability analysis (based on the solution of the eigensystem of the linearized equations) was successful for determining dynamics in the context of oscillator flows, whose corresponding fluctuation is characterized by a discrete frequency distribution, noise-amplifier-like flows that hold a broadband-frequency representation, do not, in general, exhibit a neutral/unstable mode (Ehrenstein and Gallaire [2005], Ehrenstein and Gallaire [2008]). Rather, they may exhibit several (quasi-neutral) stable modes. In those cases, a more suitable formalism is given by the Resolvent operator that establishes the linear input/output transfer function from the linearized equations around the mean-flow. The Singular-Value Decomposition of this operator provides then a framework (McKeon and Sharma [2010]) where the most energetically-amplified input/output

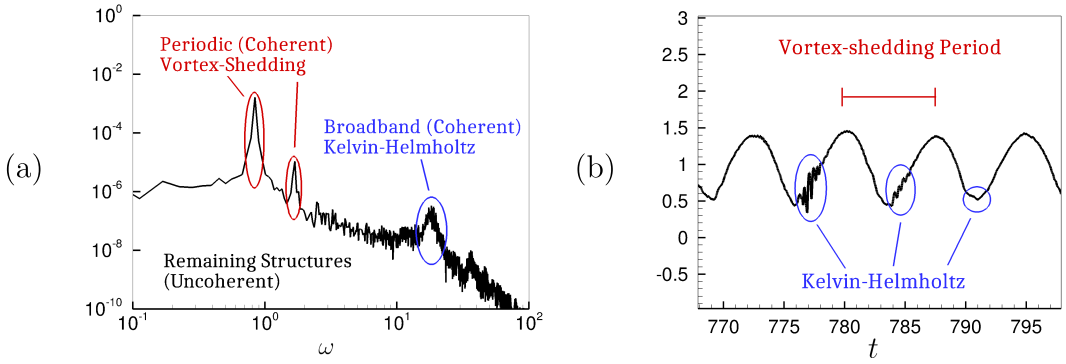

Figure 5.1: Sketch of the frequency representation of a flow's signal containing the vortex-shedding and the Kelvin-Helmholtz (coherent contributions) and the remaining (uncoherent) fluctuations (a). The time-signal from which this sketch was produced is given in (b) and corresponds to a point on the oscillating shear-layer.

dynamics are evidenced. Furthermore, it was shown (Beneddine et al. [2016]) that, if those Singular-Values are well separated from one-another, the flow dynamics can be described by a low-rank approximation of the (infinite dimensional) Resolvent operator. This technique has been extensively applied in several contexts: to highlight coherent turbulent mechanisms (Cossu et al. [2009], Pujals et al. [2009], Schmidt et al. [2018]), as a tool for data-assimilation (Symon et al. [2019a]) and control (Leclercq et al. [2019]), to mention a few.

In this chapter, we apply those concepts to the flow around a squared-section cylinder at $Re = 22000$ on the reference solution coming from a DNS run. This flow is characterized by a periodic component at a well defined frequency, corresponding to the vortex-emission, a common feature of flows around bluff-bodies, and a broadband-frequency one. This last contribution holds both a coherent part, corresponding to the well-defined structure of the Kelvin-Helmholtz instabilities happening at the shear-layer of the separation region, and an uncoherent/turbulent part. A sketch of such signal is depicted in figure 5.1 (a).

In what follows, we wish to recover both coherent components, namely the vortex-shedding and Kelvin-Helmholtz, through mean-flow stability or Resolvent analysis. To do so, we apply one version of the triple decomposition to the complete fully-developed turbulent signal. This decomposition was first introduced by Reynolds and Hussain [1972] and used to extract a periodic-wave component from the background turbulence. Further work, employed other versions of this decomposition in order to extract (not necessarily periodic) coherent structures (Hussain [1983], Hussain [1986]). Since then, several studies have been carried out using this framework in numerous contexts, for example, in turbulence modeling (Speziale [1991], Palkin et al. [2016]) or in experimental data-treatment (Lyn et al. [1995]). More recently, Mezić [2013] proposed another version of this decomposition based on the Koopman operator formalism where the flow is separated in a quasi-periodic component (holding possibly multiple well-defined frequencies) from a broadband one. In this work, the latter will be preferred since its rigorous application leads automatically to two separate equations, one for the (quasi-)periodic component and the second one for the broadband component.

The first one corresponds to a nonlinear set of equations, where the influence of the broadband fluctuation (coherent and uncoherent) is given classically through the

Reynolds-stress tensor, which is modeled by an eddy-viscosity (Boussinesq hypothesis), leading to a similar formalism as in URANS. This eddy-viscosity is then computed from treating the statistics from a DNS run. Upon further decomposition of the periodic component into mean-flow and periodic fluctuation, we can establish the mean-flow equations, where the eignvalue analysis is performed to recover the vortex-shedding mode.

The second one can be viewed as forced linear equation, where the forcing term comes from the quadratic nonlinearities of the broadband component itself and the linear operator corresponds to the linearized Navier-Stokes equations around the periodic component, in a similar way as in Floquet stability theory. In order to establish the input/output relation (Resolvent) in this Floquet-like framework, we consider that both the forcing term and the perturbation depend not only on the imposed frequency but also on the vortex-shedding time scale. By doing so, we are able to recover the dependency of the high-frequency Kelvin-Helmholtz modes on the phase of the slow and periodic movement of the shear layer, given by the vortex-shedding frequency. It is interesting to note that this phenomenon is observed in experiments (see Brun et al. [2008]) and also in the DNS performed in this work. For example, if we probe the unsteady flow at some point within the oscillatory motion of the shear-layer (figure 5.1 (b)), we can see that, according to the phase of the periodic movement, we can observe the Kelvin-Helmholtz structures or not at all.

This chapter will be structured as follows: in the section §2, we will introduce and apply the concepts of the triple decomposition, giving rise to the equations governing periodic and broadband components. Then, in section §3 we will present the DNS in more detail and the mean-flow analysis deduced in the previous section. Finally, we provide some conclusions and discussions.

2 Triple Decomposition and Mean-Flow analysis

The goal of this section is to provide the theoretical framework for Resolvent analysis in general turbulent flows. We suppose here that those flows are governed by the incompressible Navier-Stokes equations:

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nabla \cdot (\nu \nabla_s \mathbf{u}) = \mathbf{0}, \quad \nabla \cdot \mathbf{u} = 0, \tag{5.1}$$

where ∇_s is a shortcut for $\nabla + \nabla^T$. In general, a turbulent flow may hold a contribution whose frequency content is discrete (quasi-periodic part) and a contribution that is continuously spread over a frequency band. In so far, following Mezić [2013], we may decompose any fully-developed velocity / pressure field into a quasi-periodic component, $\langle \mathbf{u} \rangle$ and $\langle p \rangle$, and a broadband one $\mathbf{u}'$ and p' such that:

$$\mathbf{u}(\mathbf{x}, t) = \langle \mathbf{u} \rangle (\mathbf{x}, t) + \mathbf{u}'(\mathbf{x}, t), \quad p(\mathbf{x}, t) = \langle p \rangle (\mathbf{x}, t) + p'(\mathbf{x}, t). \tag{5.2}$$

The extraction of the quasi-periodic component $\langle \mathbf{u} \rangle (t)$ of any given signal $\mathbf{u}(t)$ can be formally expressed in terms of the Koopman operator theory. According to Arbabi and Mezić [2017], Mezić [2013], there exists a countable set of frequencies $\{\omega_k\}$ such that the harmonic averages, defined as

$$\hat{\mathbf{u}}(\omega_k) = \lim_{T \to +\infty} \frac{1}{T} \int_0^T e^{-\mathrm{i}\omega_k t} \mathbf{u}(t) \, dt, \tag{5.3}$$

converge to a finite non-zero quantity. Typically this set of discrete frequencies is composed of few fundamental (incommensurable) frequencies, that arise, for example, from global instability mechanisms (absolute instabilities, Huerre and Monkewitz [1985],Sipp et al. [2010]), and their nonlinear interactions (see, for example Leclercq et al. [2019]). The resulting modes $\hat{\mathbf{u}}(\omega_k)$ are called Koopman modes. The quasi-periodic component is then reconstructed as the Koopman Mode Decomposition (KMD):

$$\langle \mathbf{u} \rangle (t) \equiv \sum_{\omega_k} e^{\mathrm{i}\omega_k t} \hat{\mathbf{u}}(\omega_k) + c.c., \tag{5.4}$$

where, here we can see that the notation $\langle \cdot \rangle$ is used to group together the extraction of the quasi-periodic component of $\mathbf{u}$ described by equations (5.3) and (5.4). The remainder of the original signal $\mathbf{u}' = \mathbf{u} - \langle \mathbf{u} \rangle$ holds a broadband-like distribution in frequency, as the corresponding harmonic averages of $\hat{\mathbf{u}}'(\omega)$ converge to zero for any frequency ω. As stated by Mezić [2013] the Koopman approach provides an extension of the triple decomposition concept, that was initially introduced by Reynolds and Hussain [1972] and based on phase-average of the time-signals. While the latter requires an a priori knowledge of the period defining the conditional averaging process, the Koopman approach provides an intrinsic definition for $\langle \mathbf{u} \rangle$, and is therefore more general. We remark here that this decomposition does not necessarily separate the flow field into a *coherent* and an *uncoherent* part, since there are many well-known physical instabilities that manifest in the form of coherent structures over a continuous broadband frequency range. A typical example of this is the convective Kelvin-Helmholtz instability occurring in shear-layers (see Beneddine et al. [2016]). In the following, we aim at modeling not only the coherent part linked to the quasi-periodic component but also the coherent contribution associated to the broadband component.

To proceed, we write the equations respectively governing the quasi-periodic component $\langle \mathbf{u} \rangle$ and the broadband fluctuation $\mathbf{u}'$:

$$\begin{aligned}
\partial_t \langle \mathbf{u} \rangle + \quad \langle \mathbf{u} \rangle \cdot \nabla \langle \mathbf{u} \rangle \quad + \nabla \langle p \rangle - \nabla \cdot (\nu \nabla_s \langle \mathbf{u} \rangle) &= -\nabla \cdot \langle \mathbf{u}' \otimes \mathbf{u}' \rangle , && (5.5)\\
\partial_t \mathbf{u}' + \langle \mathbf{u} \rangle \cdot \nabla \mathbf{u}' + \mathbf{u}' \cdot \nabla \langle \mathbf{u} \rangle \quad + \nabla p' - \nabla \cdot (\nu \nabla_s \mathbf{u}') &= \mathbf{f}', && (5.6)
\end{aligned}$$

where $\mathbf{f}' = \nabla \cdot (\langle \mathbf{u}' \otimes \mathbf{u}' \rangle - \mathbf{u}' \otimes \mathbf{u}')$ represents the nonlinear interactions of the fluctuation $\mathbf{u}'$. The derivation of those equations relies on specific properties of the operator $\langle \cdot \rangle$, that extracts the quasi-periodic component of a given flowfield. Those properties are documented in Reynolds and Hussain [1972] and are verified to be true even in the Koopman framework [Mezić, 2013].

In the following, we will separetely consider the modelling of the quasi-periodic component $\langle \mathbf{u} \rangle$ and the broadband component $\mathbf{u}'$.

2.1 Modelling of quasi-periodic component

In equation (5.5) governing the quasi-periodic component, the term $\langle \mathbf{u}' \otimes \mathbf{u}' \rangle$ corresponds to the Reynolds stress tensor stemming from the broadband contribution $\mathbf{u}'$. Even though $\mathbf{u}'$ may involve a part related to coherent structures, we model this tensor using the Boussinesq hypothesis:

$$-\langle \mathbf{u}' \otimes \mathbf{u}' \rangle + \frac{2}{3} \langle \kappa \rangle I \approx \langle \nu_t \rangle \nabla_s \langle \mathbf{u} \rangle , \tag{5.7}$$

where $\langle \nu_t \rangle = \langle \nu_t \rangle (t)$ is an eddy viscosity that is, in general, quasi-periodic as well, and thus the slightly unusual notation for it around the brackets $\langle \cdot \rangle$. This approximation results in the following equation:

$$\partial_t \langle \mathbf{u} \rangle + \langle \mathbf{u} \rangle \cdot \nabla \langle \mathbf{u} \rangle + \nabla \langle p \rangle - \nabla \cdot ((\nu + \langle \nu_t \rangle) \nabla_s \langle \mathbf{u} \rangle) = \mathbf{0}, \tag{5.8}$$

where, in an abuse of notation, the turbulent kinetic energy $\langle \kappa \rangle = \langle \mathbf{u}' \cdot \mathbf{u}' \rangle /2$ has been incorporated in the pressure term $\langle p \rangle$.

This equation governs the quasi-periodic part of the flowfield. If a turbulence model is provided $f(\langle \nu_t \rangle, \langle \mathbf{u} \rangle) = 0$, then this equation is closed and may be solved for by time-stepping. In general, such an equation may exhibit various types of solutions: fixed-points, periodic limit-cycle solutions linked for example to Hopf bifurcations of these fixed-points, or even multi-frequency quasi-periodic solutions linked to secondary instabilities of the periodic limit-cycles. Hence, classical stability analysis may be performed around fixed points [Crouch et al., 2009, Meliga et al., 2012, Mettot et al., 2014a] or Floquet type stability analysis around limit cycles. Yet, one has to keep in mind that only quasi-periodic solutions to this equation are "strictly" relevent.

2.1.1 Modelling with mean-flow marginal global modes

As explained in Mettot et al. [2014b], a mean-flow approach may be favoured to characterize the unsteady quasi-periodic component of $\langle \mathbf{u} \rangle$ from the sole knowledge of the time-averaged mean-flow $\bar{\mathbf{u}}$. Indeed, from Mezić [2013], it is known that the Koopman mode corresponding to the fundamental harmonic of a time-periodic limit-cycle associated to equation (5.8) corresponds to a marginal eigenvalue of the linearized operator around the mean $\bar{\mathbf{u}}$. There are actually two conditions for this result to hold: either the limit-cycle should be purely harmonic (only the fundamental) or the amplitudes of all Koopman modes (Fourier modes) be weak. We thus further decompose the quasi-periodic component into its time-average (denoted with $\overline{(\cdot)}$) and its corresponding fluctuation (denoted with $\tilde{(\cdot)}$), completing the triple decomposition. We remark that the eddy-viscosity $\langle \nu_t \rangle$, having the same time-dependency as $\langle \mathbf{u} \rangle$, may have, in general, a fluctuation associated to it. Introducing the decompostion

$$\langle \mathbf{u} \rangle (\mathbf{x}, t) = \bar{\mathbf{u}}(\mathbf{x}) + \tilde{\mathbf{u}}(\mathbf{x}, t), \quad \langle p \rangle (\mathbf{x}, t) = \bar{p}(\mathbf{x}) + \tilde{p}(\mathbf{x}, t), \quad \langle \nu_t \rangle (\mathbf{x}, t) = \bar{\nu}_t(\mathbf{x}) + \tilde{\nu}_t(\mathbf{x}, t) \tag{5.9}$$

in the phase-average equation (5.8) yields the equations governing the mean-flow (with in particular additional Reynolds stresses stemming from the quasi-periodic component $\tilde{\mathbf{u}}$):

$$\bar{\mathbf{u}} \cdot \nabla \bar{\mathbf{u}} + \nabla \bar{p} - \nabla \cdot ((\nu + \bar{\nu}_t) \nabla_s \bar{\mathbf{u}}) = -\overline{\tilde{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}}} + \nabla \cdot \overline{(\tilde{\nu}_t \nabla_s \tilde{\mathbf{u}})}, \tag{5.10}$$

and the linearized equations around the mean-flow:

$$\partial_t \tilde{\mathbf{u}} + \bar{\mathbf{u}} \cdot \nabla \tilde{\mathbf{u}} + \tilde{\mathbf{u}} \cdot \nabla \bar{\mathbf{u}} + \nabla \tilde{p} - \nabla \cdot ((\nu + \bar{\nu_t}) \nabla_S \tilde{\mathbf{u}}) - \nabla \cdot (\tilde{\nu}_t \nabla_s \bar{\mathbf{u}}) = \mathbf{0}, \tag{5.11}$$

If a turbulence model is at hand, we may explicitly relate the fluctuations $\tilde{\nu}_t$ to variations of the closure model. If not, we neglect this term, which we will assume in the following to make it simple. We finally look for eigenvalues of these linearized equations in the form:

$$\tilde{\mathbf{u}} = \hat{\mathbf{u}} e^{\lambda t}, \tag{5.12}$$

so that (with a slight abuse of notations since the pressure term has been omitted)

$$\lambda\hat{\mathbf{u}} + L(\overline{\mathbf{u}}, \overline{\nu}_t)\hat{\mathbf{u}} = \mathbf{0}, \tag{5.13}$$

where the operator $L(\overline{\mathbf{u}}, \overline{\nu}_t)$ represents the linearized Navier-Stokes operator for a given mean-flow $\overline{\mathbf{u}}$ and a given mean eddy-viscosity $\overline{\nu}_t$ and $\lambda = \sigma + i\omega$ is the eigenvalue, dictating how the mode will decay (with σ) and oscillate (with ω). We can see that, in a natural manner, the mean-eddy-viscosity is incorporated in this operator. This is in accordance with Meliga et al. [2012], where, to obtain the vortex-shedding mode, the eddy-viscosity, coming from the URANS (Spalart-Allmaras) model, was taken into account. In Mettot et al. [2014b], it was argued that in most of the cases, the frequency of the eigenmode was not really sensitive to $\bar{\nu}_t$, so that a crude approximation $\bar{\nu}_t = 0$ could also be considered in some cases.

According to Mezić [2013], if one of the two conditions mentioned above are met, then the eigenvalue λ of (5.13) should be marginal with a frequency equal to the fundamental frequency of the limit cycle. Also, the eigenvector should correspond to the Koopman mode at the fundamental frequency. Note also that considering eigenvalues of the linearized operator (rather than resolvent modes) complies with the discrete nature of the Koopman modes.

2.1.2 How to obtain ν_t

The mean eddy-viscosity $\bar{\nu}_t$ may be obtained through different ways. If a turbulence model is used, then a straightforward time-average of $\langle \nu_t \rangle$ may be performed. If the fluctuation field $\mathbf{u}$ is obtained by a DNS, then by taking the time-average of equation 5.8, we obtain:

$$-\overline{\mathbf{u}'\otimes\mathbf{u}'} + \frac{2}{3}\overline{\kappa}I = \overline{\nu}_t\nabla_s\overline{\mathbf{u}} + \overline{\tilde{\nu}_t\nabla_s\tilde{\mathbf{u}}}. \tag{5.14}$$

Applying the hypothesis $\tilde{\nu}_t = 0$ mentioned above leads to the following mean eddy-viscosity:

$$\overline{\nu}_t = \frac{\nabla_s\overline{\mathbf{u}} : (-\overline{\mathbf{u}'\otimes\mathbf{u}'} + 2/3\overline{\kappa}I)}{\nabla_s\overline{\mathbf{u}} : \nabla_s\overline{\mathbf{u}}}. \tag{5.15}$$

Note that a DNS computation generally straightforwardly provides the mean-flow $\overline{\mathbf{u}}$, the periodic fluctuation $\tilde{\mathbf{u}}$ (by performing harmonic-averages) and the statistics $\overline{\mathbf{u}\otimes\mathbf{u}}$. These statistics may be related to the $\overline{\mathbf{u}'\otimes\mathbf{u}'}$ through: $\overline{\mathbf{u}'\otimes\mathbf{u}'} = \overline{\mathbf{u}\otimes\mathbf{u}} - \overline{\mathbf{u}}\otimes\overline{\mathbf{u}} - \overline{\tilde{\mathbf{u}}\otimes\tilde{\mathbf{u}}}$.

The hypothesis ($\tilde{\nu}_t = 0$) may suffice as a first-order approximation, that leads to qualitatively correct global modes. However, a finer definition of $\tilde{\nu}_t$ could be found either by a more detailed study of the snapshots from the DNS (if available) or through a data-assimilation technique (see Camarri et al. [2017]).

2.2 Resolvent analysis for broadband component

For the recovery of the broadband component, a frequency domain input-output approach, such as the resolvent formalism, is more suited. However, we remark that the linearized operator involved in the equations modeling the broadband component $\mathbf{u}'$, equation (5.6), is quasi-periodic, since the field around which the equations are linearized, $\langle \mathbf{u} \rangle (t)$, is quasi-periodic:

$$\partial_t\mathbf{u}' + L(\langle \mathbf{u} \rangle (t))\mathbf{u}' = \mathbf{f}'. \tag{5.16}$$

Here, for the sake of clarity, we will suppose it is only periodic. This formalism is similar to the one appearing in Floquet stability theory, where we suppose that the solution may be written in the form

$$\mathbf{u}' = \check{\mathbf{u}}(t)e^{\lambda t},$$

where $\check{\mathbf{u}}(t)$ is a T-periodic mode that grows/decays with the Floquet exponent λ. Here, instead of analyzing the stability of the operator (as done in Floquet stability theory), we study its input/output relation. To do so, and similarly to Floquet stability analysis, we look for solutions of the forcing term and its corresponding response under the separated form:

$$\mathbf{f}' = \check{\mathbf{f}}(t)e^{\mathrm{i}\omega t} + c.c., \quad \mathbf{u}' = \check{\mathbf{u}}(t)e^{\mathrm{i}\omega t} + c.c. \tag{5.17}$$

where $\check{\mathbf{u}}(t), \check{\mathbf{f}}(t)$ are T-periodic fields and ω is a given frequency. Here we suppose that the frequency ω does not coïncide with a Floquet exponent eigenvalue $\lambda = \mathrm{i}\omega$ of the periodic operator $\partial_t + L(\langle \mathbf{u} \rangle (t))$, which is the case if stable periodic limit-cycles are considered. We may then rewrite equation (5.16) as:

$$\partial_t \check{\mathbf{u}} + \mathrm{i}\omega\check{\mathbf{u}} + L(\langle \mathbf{u} \rangle (t))\check{\mathbf{u}} = \check{\mathbf{f}}. \tag{5.18}$$

This equation establishes the input/output relation mentioned before, since, for to a given T-periodic forcing term $\check{\mathbf{f}}$, corresponds a T-periodic response $\check{\mathbf{u}}$, upon integration of 5.18. We remark however that only a T-periodic solution to this equation makes sense. One way of recovering this periodic solution can be to resort to feedback control techniques, where an extra forcing term is considered, penalizing the non-periodic component (similarly to Shaabani-Ardali et al. [2017]). Another way is to enforce periodicity by construction. This is the basic idea of, for example, the Floquet-Fourier-Hill Method (FFHM, see Deconinck and Kutz [2006]), where the solution $\check{\mathbf{u}}$, the forcing term $\check{\mathbf{f}}$ and the operator $L(\langle \mathbf{u} \rangle (t))$ are expanded as Fourier series, i.e.:

$$\check{\mathbf{u}}(t) = \sum_n \check{\mathbf{u}}_n e^{in\omega_0 t}, \quad \check{\mathbf{f}}(t) = \sum_n \check{\mathbf{f}}_n e^{in\omega_0 t}, \quad L(\langle \mathbf{u} \rangle (t)) = \sum_n L_n e^{in\omega_0 t}, \tag{5.19}$$

where $\omega_0 = 2\pi/T$ and $(\check{\mathbf{u}}_n, \check{\mathbf{f}}_n, L_n)$ are the different harmonics. With this, equation (5.18) can be recast in the form:

$$(\mathrm{i}\omega\mathcal{I} + \mathcal{L})\mathcal{U} = \mathcal{F}, \quad \mathcal{L} = \begin{bmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \cdots & L_0 + \mathrm{i}\omega_0 I & L_{-1} & L_{-2} & \cdots \\ \cdots & L_1 & L_0 & L_{-1} & \cdots \\ \cdots & L_2 & L_1 & L_0 - \mathrm{i}\omega_0 I & \cdots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{5.20}$$

where the vectors $\mathcal{U}, \mathcal{F}$ represent collections of the harmonics of the perturbation and forcing terms. The matrix $\mathcal{L}$ is the so-called Hill matrix (Lazarus and Thomas [2010]) and plays the same role as the operator $L(\overline{\mathbf{u}})$ in the case of mean-flow analysis. We can see thus that, in this case, there exists a generalization for the Resolvent operator, where:

$$\mathcal{R}(\omega) = (\mathrm{i}\omega\mathcal{I} + \mathcal{L})^{-1}. \tag{5.21}$$

Equation (5.20) formally establishes the existence of the T-periodic response $\check{\mathbf{u}}(t)$ if ω does not coincide with any of the eigenvalues of $\mathcal{L}$ (since in such a case the matrix $\mathrm{i}\omega\mathcal{I} + \mathcal{L}$

is invertible). The gain to be optimized under the constraint $\mathcal{U} = \mathcal{R}\mathcal{F}$ should be viewed as an energy-ratio between the fluctuation and the forcing term, both integrated over the period T. A possible form of such a gain is given by:

$$G(\omega) = \frac{\frac{1}{T}\int_0^T ||\check{\mathbf{u}}(t)||^2 dt}{\frac{1}{T}\int_0^T ||\check{\mathbf{f}}(t)||^2 dt} = \frac{\sum_n ||\check{u}_n||^2}{\sum_n ||\check{f}_n||^2} \equiv \frac{|||\mathcal{U}|||^2}{|||\mathcal{F}|||^2}, \tag{5.22}$$

where the last equality is the application of Parseval's theorem and $||\cdot||$ denotes the usual space energy norm $||\cdot||^2 = \int_\Omega |\cdot|^2 d\Omega$. We can see that this energy gain is a natural extension of the usual steady one.

This system is costly to solve, as it requires inversions of the operator $\mathrm{i}\omega\mathcal{I} + \mathcal{L}$, if the FFHM is preferred or many forward/backward time-integrations, if time-methods are preferred. In what follows, we propose two simplifications, namely the steady-flow $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$ and the quasi-static approximations.

2.2.1 Steady flow approximation: $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$

In the case $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$, we may avoid the complexity of equations (5.18) or (5.20)and recover the more usual Resolvent formalism. In this approximation, the modes $\check{\mathbf{f}}, \check{\mathbf{u}}$ would evolve on top of the mean-flow and their time-dependency would be eliminated: $\partial_t \check{\mathbf{u}} = 0$. We recover then the classical Resolvent operator:

$$\check{\mathbf{u}}(\omega) = R(\overline{\mathbf{u}}, \omega)\check{\mathbf{f}}(\omega), \quad R(\overline{\mathbf{u}}, \omega) = (\mathrm{i}\omega I + L(\overline{\mathbf{u}}))^{-1}. \tag{5.23}$$

Following Beneddine et al. [2016], we are interested in the pair $(\check{\mathbf{f}}, \check{\mathbf{u}})$ that maximizes the energy gain:

$$G(\omega) = \frac{||\check{\mathbf{u}}||^2}{||\check{\mathbf{f}}||^2}. \tag{5.24}$$

This is equivalent to studying the Singular-Value Decomposition of $R(\cdot) \equiv R(\omega, \overline{\mathbf{u}})(\cdot) = \sum_i \mu_i \hat{\mathbf{u}}_i(\check{\mathbf{f}}_i, \cdot)$, where the modes $(\check{\mathbf{f}}_i, \check{\mathbf{u}}_i)$ are obtaining by solving:

$$RR^\dagger \check{\mathbf{f}}_i = \mu_i^2 \check{\mathbf{f}}_i, \quad R^\dagger R \check{\mathbf{u}}_i = \mu_i^2 \check{\mathbf{u}}_i, \quad ||\check{\mathbf{u}}_i|| = ||\check{\mathbf{f}}_i|| = 1 \tag{5.25}$$

Note that the approximation $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$ may be valid as soon as the forcing $\check{\mathbf{f}}_i$ and response $\check{\mathbf{u}}_i$ (or more precisely the wavemaker region associated to $(\check{\mathbf{f}}_i,\ \check{\mathbf{u}}_i)$, see Brandt et al. [2011]) are located in a region where $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$.

2.2.2 Quasi-static Approximation

Another possible simplification of equations (5.18) is possible when the evolution time-scale of the periodic component $\langle \mathbf{u} \rangle$ is slow in comparison with the excitation frequency ω. In those cases, we may postulate that the modes $\check{\mathbf{u}}, \check{\mathbf{f}}$ also evolve on the slow time-scale. This hypothesis can be formally translated by the introduction of the slow time-scale $\tau = \varepsilon t$ that is involved in the fields $(\langle \mathbf{u} \rangle (\tau), \check{\mathbf{u}}(\tau), \check{\mathbf{f}}(\tau))$. Therefore $\partial_t \check{\mathbf{u}} = \varepsilon \partial_\tau \check{\mathbf{u}}$ and focusing on the leading-order equation, yields:

$$\mathrm{i}\omega\check{\mathbf{u}} + L(\langle \mathbf{u} \rangle (\tau))\check{\mathbf{u}} = \check{\mathbf{f}} \quad \Rightarrow \quad \check{\mathbf{u}}(\tau) = H_\tau \check{\mathbf{f}}(\tau). \tag{5.26}$$

We can see that the resulting equation is nothing but the Resolvent operator linearized around the instantaneous phase-averaged flow $H_\tau \equiv R(\omega, \langle \mathbf{u} \rangle (\tau))$. This means that, in this model, the modes $\check{\mathbf{f}}$ and $\check{\mathbf{u}}$ coincide with the classical Resolvent analysis based on the high frequency ω and snapshots of $\langle \mathbf{u} \rangle (\tau)$ taken along the slowly evolving quasi-periodic component. For this reason, we will refer to this approach as the quasi-static approximation. Furthermore, the gains from the SVD of the operator H_τ

$$G(\omega, \tau) = \frac{||\check{\mathbf{u}}(\tau)||^2}{||\check{\mathbf{f}}(\tau)||^2}, \tag{5.27}$$

that classically are solely dependent on the frequency ω are now dependent on τ as well and denoted by $\mu_i^\tau(\omega)$.

2.3 Remarks on the choice of averaging procedure and eddy-viscosity

In the previous sections, we proposed a theoretical decomposition of the flow variables based on the Koopman formalism, leading to two separate equations, one for the quasi-periodic component $\langle \mathbf{u} \rangle$ and one for the broadband one $\mathbf{u}'$. We remark, however, that the broadband component $\mathbf{u}'$, may itself be composed of a coherent and an uncoherent part $\mathbf{u}' = \mathbf{u}'_{co} + \mathbf{u}'_{unco}$, the first one related to spatio-temporal coherent structures driven by instability mechanisms, the second one related rather to chaotic/turbulent motion. In equation (5.6) governing $\mathbf{u}'_{co}$, this latter uncoherent broadband term may be modelled by an additional eddy-viscosity term, say ν'_t which is a priori different from $\langle \nu_t \rangle$. Such a term then appears in equations (5.16) and (5.20) to take into account the effects of the broadband uncoherent part $\mathbf{u}'_{unco}$ on $\mathbf{u}'_{co}$. Such a modelling strategy for a broadband uncoherent part was already discussed by Morra et al. [2019], who found, in a case where no quasi-periodic component existed ($\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$), that considering an eddy-viscosity term in the linearized equations for the fluctuations improved the prediction of the Resolvent modes. Furthermore, such an eddy viscosity term has been extensively used, see for example, Cossu et al. [2009], Pujals et al. [2009], to model broadband large scale coherent structures in shear-flows. However, in few cases [Beneddine et al., 2016], it appears sufficient to consider solely the molecular viscosity (as done in the previous section, 3 of this chapter) to reconstruct the coherent broadband fluctuation field. This may stem for example from a lower Reynolds number so that $\mathbf{u}' \approx \mathbf{u}'_{co}$, while at higher Reynolds numbers the coherent part is always accompanied by an uncoherent part. The modelling strategy described may be sketched in figure 5.2.

3 Mean-Flow analysis on DNS data

In this section, we illustrate the theory in the case of a squared-section cylinder at moderately high Reynolds number. First, we will present the Direct Numerical Simulation (DNS) results, exhibiting a well defined periodic component, corresponding to the vortex-shedding phenomenon, and a broadband part, comprising in particular small Kelvin-Helmholtz type vortices arising in the shear-layer that issues from the detachment point. We will then analyse the fluctuation dynamics of the periodic and broadband components with the Resolvent based techniques presented in the previous section.

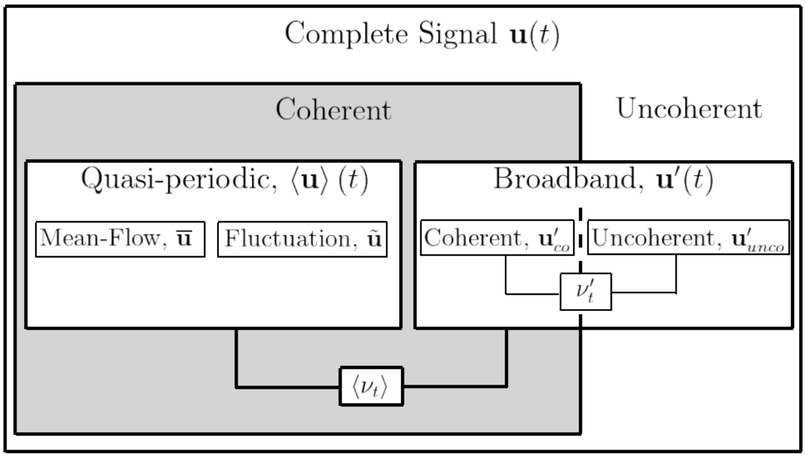

Figure 5.2: Proposed scenario for a decomposition and modeling of the flow field.

3.1 Direct Numerical Simulation solution

The configuration corresponds to the flow around a squared-section cylinder at Reynolds number $Re = U_\infty D\rho/\mu = 22000$, based on the cylinder's diameter. It corresponds to a benchmark study case for numerical methods in turbulent flows (Trias et al. [2015], Nakayama and Vengadesan [2002], Verstappen and Veldman [1997]) and the assessment of turbulence models (Bosch and Rodi [1998], Bao et al. [2011], Minguez et al. [2011]).

The DNS code solves the three-dimensional compressible Navier-Stokes equations, with a standard Sutherland law for the dynamic viscosity. The code is run at a low inflow Mach number, $M = 0.1$, representing a near-incompressible flow regime. The spatial discretisation used in the solver corresponds to a finite-volume method based on a modification of the AUSM+(P) scheme, which is second-order accurate. The time-integration corresponds to the second-order accurate backward scheme of Gear, with a non-dimensional time step (based on the inflow velocity and in the diameter of the cylinder) of 3×10^{-3}, which is around 100 times smaller than the highest observed frequency in the flow corresponding to the Kelvin-Helmholtz instability. For further details, see Dandois et al. [2018] and references herein. We use the FastS code, developed by ONERA, which is a highly optimised flow solver for high performance computing clusters.

The time-averaged flow is shown in figure 5.4 (b). A re-circulation length of $L = 0.97$ (measured from the center of the square) is obtained, which is in close agreement with the literature (for example, $L = 1.04$ by Trias et al. [2015]). The main physical unsteady feature of the flow corresponds to the vortex shedding phenomenon, as also observed in other bluff-body configurations. The frequency of this phenomenon can be seen in, for example, figure 5.3 (a), which represents the Fourier Transform of the cross-stream velocity signal at a point on the symmetry axis in the wake of the cylinder. We can see that the peak in frequency is sharp and corresponds to the frequency $\omega_0 \approx 0.83$ ($St \approx 0.133$, which is in a good agreement with Rodi [1997], Bosch and Rodi [1998], Trias et al. [2015]). The spectrum also exhibits smaller sharp peaks at the harmonics, due to the nonlinear interactions associated to the fundamental mode at ω_0. Coming back to the triple decomposition framework introduced in the last section, we will consider this phenomenon as being discrete in frequency, and therefore accounted for in $\langle\mathbf{u}\rangle$. The structure of the Koopman mode at $\omega = \omega_0$, given by the harmonic averaging procedure, is

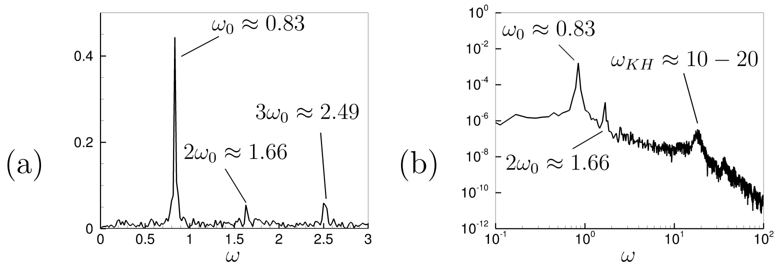

Figure 5.3: DNS: FFT of signal for the (a) cross-wise velocity component at $(0, 2.5)$ (second green dot in figure 5.4) and (b) Power-Spectral Density for the stream-wise velocity on the point $(-0.45, 0.63)$ (first green dot)

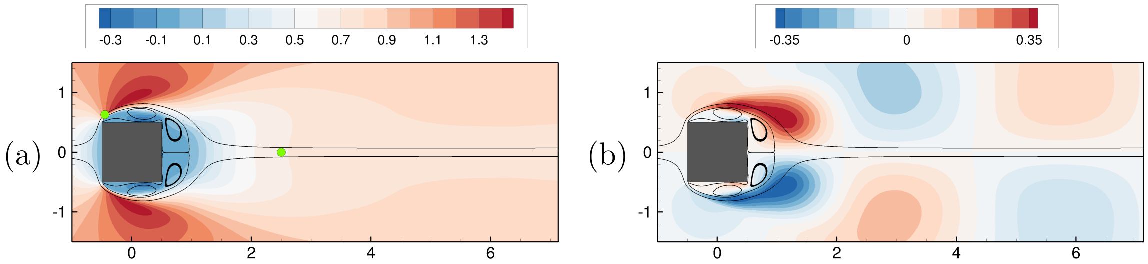

Figure 5.4: DNS: time average flow field (a) and Fourier mode of frequency $\omega_0 \approx 0.83$ (b).

shown in 5.4 (b). The flowfield also holds a contribution at higher frequencies, shown as a 'bump' in figure 5.3 (b), which is relative to the frequency spectrum of the stream-wise velocity near the upstream corner of the square. This broadband bump corresponds to small-scale convective Kelvin-Helmholtz instabilities (see Brun et al. [2008], Trias et al. [2015]) evolving in the shear-layer produced by the separation. This instability manifests itself over a broad range of frequencies $\omega_{KH} \approx 10-20$ approximately. For this reason, we will consider this phenomenon belonging to the broadband component $\mathbf{u}'$.

3.2 Mean-Flow Stability Analysis: periodic component

In this section, we perform the stability analysis on the mean-flow operator to recover the periodic component of the flow. As mentioned before, this analysis takes into account a mean-eddy-viscosity $\overline{\nu}_t$ which models the broadband fluctuations and will be computed from the DNS statistics. This quantity is given by equation 5.15. We can see from this definition that its spatial distribution is dependent on the numerator $prod = \nabla_s\overline{\mathbf{u}} : (-\overline{\mathbf{u}'\otimes\mathbf{u}'} + 2/3\overline{\kappa}I)$, also referred to as the production term, and the denominator $dest = \nabla_s\overline{\mathbf{u}} : \nabla_s\overline{\mathbf{u}}$, referred to as the production term. The destruction (plotted in 5.5 (a)) is a positive quantity and is only dependent on the gradients of the mean-flow. We can see that this quantity is strong at the shear-layers and decays exponentially away from the cylinder. The production term (plotted in 5.5 (b)) also follows a similar spatial dependency. We remark that that at some regions of the flow near the corners of the cylinder and at the far wake, the production is negative. At those regions, in the computation of the eddy-viscosity, the production value is set to zero. This is done to avoid ill-posedness problem for the operator $L(\overline{\mathbf{u}}, \overline{\nu}_t)$. The eddy viscosity is then given by the ratio between the production and the destruction. However, at regions where the destruction term is small, the value of $\overline{\nu}_t$ can be large. To avoid this issue, we use instead a denominator

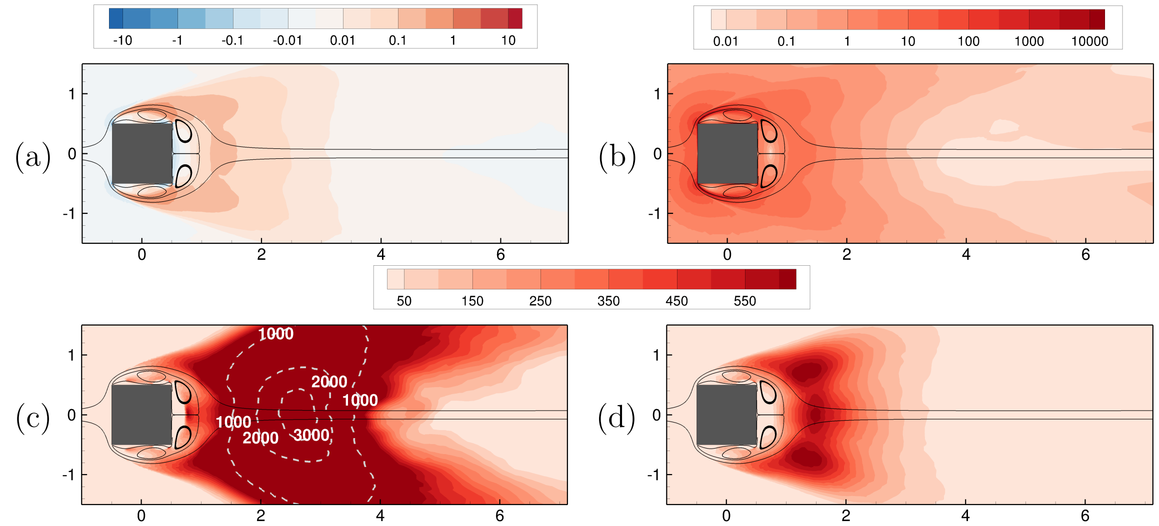

Figure 5.5: Mean-eddy-viscosity computation: (a) production term (numerator in equation 5.15, $prod = \nabla_s \overline{\mathbf{u}} : (-\overline{\mathbf{u}' \otimes \mathbf{u}'} + 2/3\overline{\kappa} I)$, (b) destruction term (denominator in equation 5.15), $dest = \nabla_s \overline{\mathbf{u}} : \nabla_s \overline{\mathbf{u}}$ and (c,d) computed eddy viscosity $\overline{\nu}_t$ with penalisation terms equal to $\varepsilon = 0.01/10$, respectively. The eddy-viscosity computed with $\varepsilon = 0.01$ (c) is saturated with the maximal value of the eddy-viscosity computed with $\varepsilon = 10$ (d) and completed with iso-contours (dashed white lines).

under the form $\sqrt{dest^2 + \varepsilon}$, where ε is a regularizing parameter. We plot the resulting $\overline{\nu}_t$ for $\varepsilon = 0.01$ and $\varepsilon = 10$ in figure 5.5 (c,d). We can see that for $\varepsilon = 0.01$, the eddy-viscosity is indeed large, leading to local Reynolds numbers of order 10. For this reason, we will assume, if not mentioned otherwise, that the quantity $\overline{\nu}_t$ was computed with $\varepsilon = 10$, although some tests will be performed with $\varepsilon = 0.01$ as well.

With this definition of the eddy-viscosity we are now in means of performing the stability analysis for the recovery of the periodic component. In figure 5.6 we provide those results for three different strategies: the laminar one (assuming $\overline{\nu}_t = 0$, as in Mettot et al. [2014b]), the frozen eddy-viscosity (as presented in section 2) and a third one which takes into account the fluctuation of $\tilde{\nu}_t$. To do so, we assume that the eddy-viscosity is governed by a URANS turbulence model (here the Spalart-Allmaras, see Spalart and Allmaras [1992]) and suppose that the mean-flow and the eddy-viscosity computed from the DNS are near to the ones issuing from a URANS run for this model. This hypothesis, at least for the mean-flow $\overline{\mathbf{u}}$, could be supported since mean-flow URANS results have been reported as been comparable with experiments and with the present DNS, for example in Meliga et al. [2016], where their re-circulation bubble was around $L = 1.11$.

We can see that, for the laminar analysis (figure 5.6 (a,b)), although the frequency is very well predicted ($\omega = 0.80$), the unstable mode is far away from the neutral line ($\sigma = 0.38$), and it corresponds to a mode whose spatial distribution is localized at a thin shear-region around the square and almost nonexistent towards the outflow. However, we can see that, taking into account the eddy-viscosity $\overline{\nu}_t$ in the analysis (figure 5.6 (c,d)), we are able to recover the overall features of the DNS Koopman mode (figure 5.4 (b)). Its eigenvalue is much closer to the neutral line ($\sigma = 0.03$), presenting now a lower frequency than before ($\omega = 0.69$). In order to investigate if this effect comes from a lack of modeling of the eddy-viscosity fluctuation, we perform as well the stability analysis considering the

Spalart-Allmaras model (figure 5.6 (e,f)), as described before. We can see that, indeed, the predicted frequency now is bigger than before ($\omega = 0.75$), suggesting that, indeed, the fluctuation of $\tilde{\nu}_t$ can have a positive effect on the eigenmode. We remark that in a URANS formalism, this fluctuation can be straightforwardly taken into account by the consideration of the complete Jacobian of the nonlinear system. Lastly, another question we may ask ourselves is the effect of the chosen eddy-viscosity between the ones computed with the two regularization parameters introduced before $\varepsilon = 10/0.01$. To answer this, we provide the modes coming from the eddy-viscosity computed with $\varepsilon = 0.01$ on the eigen-spectra in figure 5.6 (c,e) by a red dot (the case $\varepsilon = 10$ was already discussed). We can see that those modes tend to have a lower frequency and they tend to be closer to the neutral axis, which is coherent with the fact that this eddy-viscosity is stronger than the previous one. However, even with this (quite important) change in the eddy-viscosity choice, the modification of the mode was not too important. This suggests that the region most sensible to the quality eigenvalue is not at the wake of the cylinder, but at the shear layers, where the destruction term *dest* is very important and the regularization procedure is not active.

3.3 Resolvent analysis: broadband component

We now turn our attention to the reconstruction of the broadband component. First, we will provide the results for the quasi-static approximation, then we will compare them with the steady-flow one. As mentioned in section 2, this approach is valid if the frequency ω is much higher than the one coming from the periodic flow, $\omega_0 = 0.83$. Thus, we place ourselves in the interval $\omega \in (5, 35)$. The Resolvent gains for $\omega > 35$ quickly decay to zero and no more energetic input/output dynamics is recovered. In this interval, the Resolvent gains are given in figure 5.7 (a-d) for four different times (or phases) of the baseline period, corresponding to $\tau/T = 0, 1/4, 2/4$ and $3/4$, where the initial time was set such that the y-velocity of the periodic flow at $(0.5, 1)$ is maximal. From those plots, we can see a clear "bump" on the frequency range $\omega \approx 10 - 20$, which corresponds well to the region in frequency where the probes on the DNS accused the Kelvin-Helmholtz instabilities. We can see from the gains as well that the two first modes have similar energies while the third and fourth modes are several orders of magnitude less energetic. This suggests that they do not participate actively in the nonlinear fluctuation of the actual signal. It is interesting to note that this "bump" corresponding to the leading energy gain is also present in the steady-flow analysis, $\langle \mathbf{u} \rangle \approx \overline{\mathbf{u}}$, (blue/red curves on figure 5.7 (e), although a little bit less energetic. We can say then that the present analysis, considering the time-periodic Resolvent operator corresponds to a refinement of the classical mean-flow procedure.

It is interesting to note that the baseline periodic flow has a special space/time symmetry, which reflects itself on the modes. Indeed, the fact that the Koopman mode (figure 5.4 (b)) is anti-symmetric with respect to the mean-flow implies that at a given phase τ, the field $\langle \mathbf{u} \rangle (\tau)$ is the mirror-symmetry of the field $\langle \mathbf{u} \rangle (\tau + T/2)$. This means that the same physical mechanisms observed at a given phase τ will also be present in the phase $\tau + T/2$, meaning also that the same Resolvent modes found at τ should be found as well at $\tau + T/2$ as their mirror-symmetric counterpart. Indeed this phenomenon is observed. We can see that the upper modes, given for the frequencies $\omega = 10/20$, (marked in red in figure 5.7 (f-i,k-n)) are mirror-symmetry version of the lower (blue) ones after a semi-

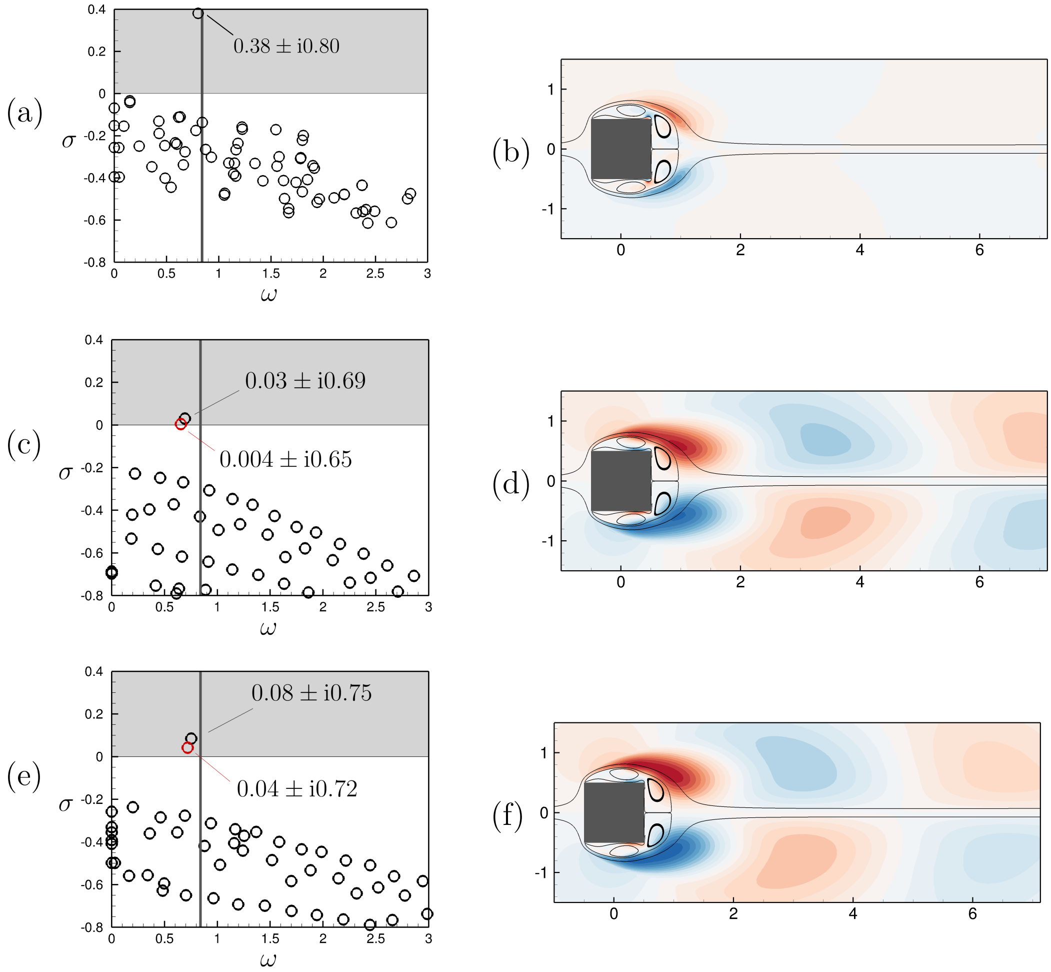

Figure 5.6: Stability analysis: (a,c,e) eigen-spectra and (b,d,f) unstable mode for laminar viscosity approach (a,b), frozen eddy-viscosity (from figure 5.5 (d)) approach (c,d) and variable eddy-viscosity approach (e,f), where $\tilde{\nu}_t$ is modeled with the Spalart-Allmaras. Red dots correspond to the unstable mode computed from the $\varepsilon = 0.01$ mean-eddy-viscosity.

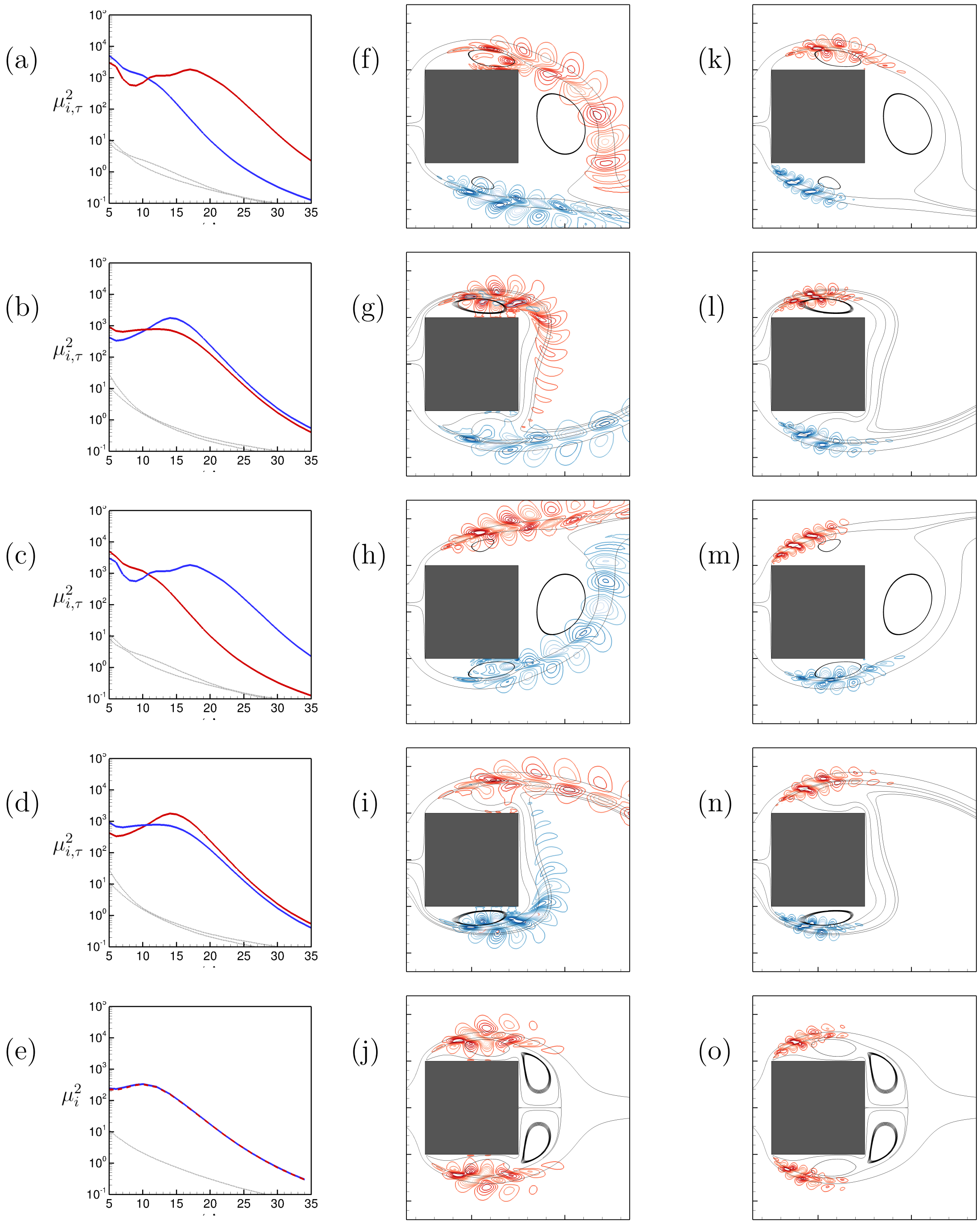

Figure 5.7: Resolvent analysis for the broadband high-frequency component. Quasi-static approximation (a,b,c,d) for $\tau/T = 0$ (b), 1/4 (c), 2/4 (d) and 3/4 (e): upper/lower modes Resolvent gains (first column) and their respective modes at $\omega = 10$ (second column) and $\omega = 20$ (third column). Red and blue modes correspond to upper/lower modes. (e) Steady-flow approximation: optimal/sub-optimal gains (first column) and leading modes for $\omega = 10$ (second column) and $\omega = 20$ (third column).

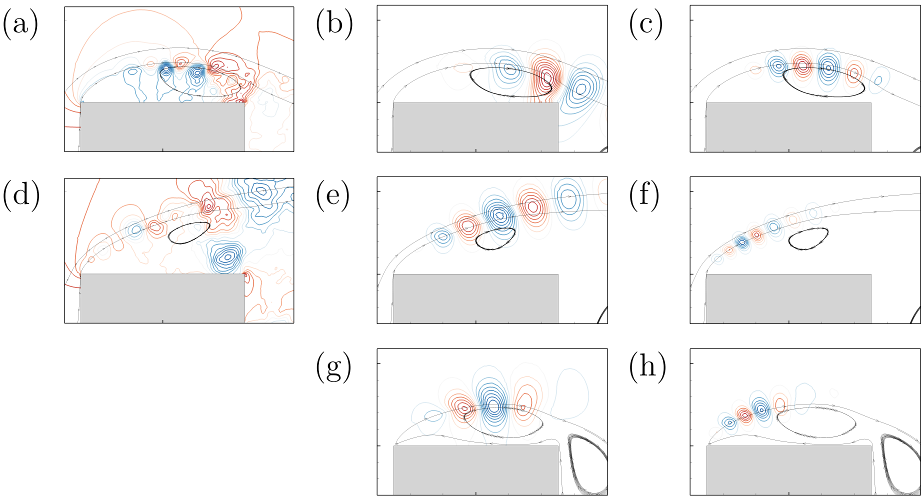

Figure 5.8: Pressure $p' = p - \langle p \rangle$ snapshots (a,d) at two different phases of the baseline period, corresponding to $\tau/T = 0$ and $1/2$ (same phase as in figure 5.7 (a,f,k) and (c,h,m)). The pressure of the modes $\check{p}$ at those phases are also plotted for $\omega = 10$ (b,e) and $\omega = 20$ (c,f). Streamlines correspond to $\langle \mathbf{u} \rangle$. Steady-flow modes for $\omega = 10$ and $\omega = 20$ are also provided in (g,h) for comparison.

period $T/2$. This is also true for their corresponding Resolvent gains, corresponding to the same colour code as the modes. As for the static approach, we can see that, since the mean-flow is symmetric, the two upper/lower modes do not exist anymore. Instead, we have two identical modes, having the same Resolvent gain, corresponding this time to a symmetric or an anti-symmetric structure. Their anti-symmetric counterpart only is plotted in figure 5.7 (j,o), also for $\omega = 10/20$.

We compare now those modes, obtained with the quasi-static approximation, with (span-wise-averaged) snapshots of the actual DNS flow at two different phases, namely $\tau/T = 0$ and $1/2$. This is achieved in figure 5.8. We can see that, indeed, the broadband fluctuation, represented there by the pressure $p' = p - \langle p \rangle$, depends on the phase of the periodic velocity field. Furthermore, this fluctuation responds almost instantaneously with changes of $\langle \mathbf{u} \rangle$, validating the quasi-static approximation in this case. An interesting observed feature is the fact that at the phase $\tau/T = 0$, we can identify some structures on the pressure snapshot that have the same wavelength as the mode obtained for $\omega = 20$ and maybe some larger structures corresponding to the lower frequencies towards the rightmost corner. This is coherent with the energy gains observed by the upper mode at that phase (red curve in figure 5.7 (a)), that presents a peak near $\omega = 20$. As for the phase $\tau/T = 1/2$, we can see that, most of the fluctuation on the shear layer is comparable with the lower frequency $\omega = 10$ instead. This is also coherent with the energy gains (red curve in figure 5.7 (c)), which presents higher values for lower frequencies and almost no amplification for frequencies such as $\omega = 20$. Also, by comparing the fluctuations with the modes coming from the steady-flow approximation, we can see that they do not match so well, as those modes do not oscillate with the shear-layers.

4 Conclusion

In this chapter, we performed flowfield reconstructions, based on linearized Navier-Stokes equations, for a turbulent flow holding both a periodic large-scale low-frequency vortex shedding component and a broadband small-scale high-frequency component (Kelvin-Helmholtz type instabilities) evolving on top of the vortex-shedding mode. The triple decomposition [Mezić, 2013] separates in an accurate and rigorous manner these two components. The approaches employed for flow reconstruction are different for both components.

For the periodic component, a mean-flow stability analysis with an eddy-viscosity modelling had to be taken into account. The reconstruction accuracy even improved when augmenting the Jacobian operator with a linearized Spalart-Allmaras model governing the eddy-viscosity fluctuations.

For the broadband component, input-output linearized Navier-Stokes equations around the vortex-shedding component naturally step in. This leads to optimal forcing and response modes whose structure depends on the phase of the vortex-shedding cycle. We note that with the triple-decomposition of the flow-field used here, no eddy-viscosity is taken into account for those fluctuations; however, we believe that, if finer statistics of the fluctuations were provided, an eddy-viscosity associated with the uncoherent/chaotic component of the broadband spectrum could be taken into account and could improve reconstruction accuracy. The general reconstruction problem involves an expensive input-output problem to solve. We proposed two approximations, for which the complex input-output problem simplifies into a classical resolvent analysis: the quasi-static and the steady-state approximations. It was shown that the quasi-static approach gave finer results than the steady-state approach, and that the structure of the quasi-static resolvent modes very accurately compared with DNS snapshots at different phases of the vortex shedding cycle.

Chapter 6

Resolvent-based Data-Assimilation for the turbulent flow around a squared-section cylinder

1 Introduction

In the fourth chapter, we explored the data-assimilation based on a model composed by the mean-flow equation, coupled with the mean-flow resolvent modes, modeling the fluctuation. In that context, the Reynolds number was small and the full complexity of a general turbulent flow could not be assessed or discussed. Furthermore, in those cases, we studied flows having either a quasi-periodic fluctuation, in which case only one or two well-defined frequencies were taken into account by the model, or a (coherent) broadband one, for which a frequency-discretization was employed. Generally, for turbulent flows, we potentially have both behaviors together, as discussed in the fifth chapter. For this reason, in the present chapter, we attempt to extend the idea of resolvent-based data-assimilation to turbulent flows. More precisely, we wish to apply those ideas on the same physical configuration considered in the fifth chapter, namely the turbulent flow around a squared-section cylinder, where the same DNS is used as the reference solution. To do so, we need a model composed by the mean-flow equation coupled with unsteady modes coming from mean-flow analysis, whose amplitudes will be used as tuning parameters. However, we can see now that designing a model for this configuration is a much more challenging task than before. Indeed, in order to obtain the correct vortex-shedding structure, necessary for the proper reconstruction of the Reynolds-stress tensor, we need an eddy-viscosity, as shown in fifth chapter, which is a spatially-dependent function that models the broadband nonlinear interactions. One could argue that this quantity could be computed from from the resolvent modes modeling the the broadband component, in a similar way we did for the white-noise forced BFS, presented in the fourth chapter. This would however represent a difficult task since now the broadband signal is richer, presenting not only coherent structures, as the Kelvin-Helmholtz modes but also uncoherent ones. Furthermore, if we look to capture as well the time-dependency of those modes with the vortex-shedding frequency (present in the quasi-static approximation), this time scale needs also to be discretized, increasing even more the complexity of the model. For those reasons, we choose instead to rely on an URANS turbulence model, here the Spalart-Allmaras. This is useful since it naturally provides a physically plausible eddy-viscosity. Moreover, as we will see, this model predicts correctly the periodic component from the DNS.

This chapter will be structured as follows: in the section §2, we will design the model, based on URANS equations and its resolvent analysis, used in the data-assimilation context. Then, in section §3 we will present the data-assimilation results, followed by conclusions.

2 Development of the model for Data-Assimilation

In this section we will present an extension of the model used for data-assimilation process discussed in the fourth chapter in the laminar/transitional case. As discussed before, the inclusion of broadband fluctuations in the model, coupling them with the mean-flow equation through the Reynolds-stress tensor and with the vortex-shedding mode through the eddy-viscosity leaves the model too complex. For this reason, we will drastically simplify the problem by considering a URANS turbulence model. The chosen model will the the Spalart-Allmaras model, as introduced in the second chapter. The solution of this model, together with its steady version (RANS) will be presented in the next paragraph, together with a model for the data-assimilation.

2.1 RANS and URANS Spalart-Allmaras Solutions

The starting point of our data-assimilation for this turbulent case will be the evaluation of the RANS-SA solution, presented in figure 6.1. This helps to quantify how far we are from the actual DNS solution by solving a spatial-only problem. This solution $(\mathbf{U}, P, \tilde{N})$ is supposed to represent, in a way, the mean-flow $\mathbf{U} \approx \overline{\mathbf{u}}$. The RANS-SA equations can be summarized in:

$$\begin{aligned} \mathbf{U} \cdot \nabla \mathbf{U} + \nabla P - \nabla \cdot \Big((\nu + \nu_t(\tilde{N})) \nabla_s \mathbf{U} \Big) &= \mathbf{0} && (6.1) \\ \nabla \cdot \mathbf{U} &= 0 && (6.2) \\ \mathbf{U} \cdot \nabla \tilde{N} - \nabla \cdot (\eta(\tilde{N}) \nabla \tilde{N}) &= s(\tilde{N}, \nabla \tilde{N}, \nabla \mathbf{U}), && (6.3) \end{aligned}$$

where $s = s(\tilde{N}, \nabla \tilde{N}, \nabla \mathbf{U})$ is the nonlinear source term for the turbulent variable. For a better description of it, see Appendix A. We can see however that this model fails to capture its overall characteristics, presenting a much larger re-circulation bubble $L \approx 3.8$. This is actually a known fact that RANS modeling fails to predict the mean-flow in bluff-body-like configuration. This comes from the fact that this kind of model was designed for capturing small-scale turbulent motion. However, in this configuration, an important large-scale motion exists, the vortex-shedding. Indeed, this flow is unstable and, if we allow it to evolve in time (in an URANS manner), it will produce a much shorter mean-re-circulation bubble (see Bosch and Rodi [1998]). Those URANS equations are given by:

$$\begin{aligned} \partial_t \mathbf{U} + \mathbf{U} \cdot \nabla \mathbf{U} + \nabla P - \nabla \cdot \Big((\nu + \nu_t(\tilde{N})) \nabla_s \mathbf{U} \Big) &= \mathbf{0} && (6.4) \\ \nabla \cdot \mathbf{U} &= 0 && (6.5) \\ \partial_t \tilde{N} + \mathbf{U} \cdot \nabla \tilde{N} - \nabla \cdot (\eta(\tilde{N}) \nabla \tilde{N}) &= s(\tilde{N}, \nabla \tilde{N}, \nabla \mathbf{U}) && (6.6) \end{aligned}$$

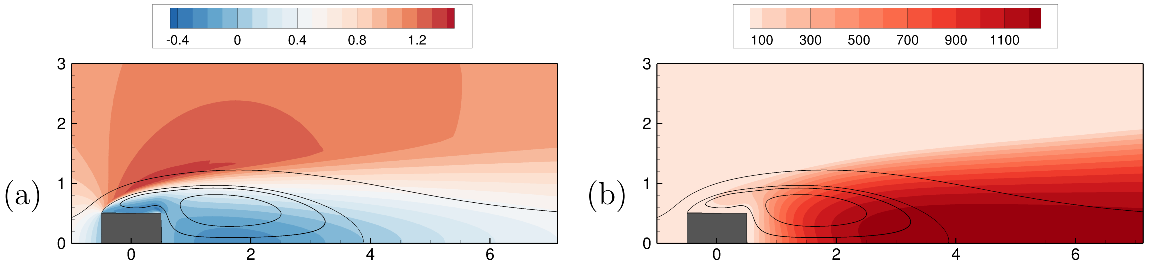

Figure 6.1: RANS solution: U_x velocity field (a) and eddy-visocisty $\tilde{N}/\nu$ (b).

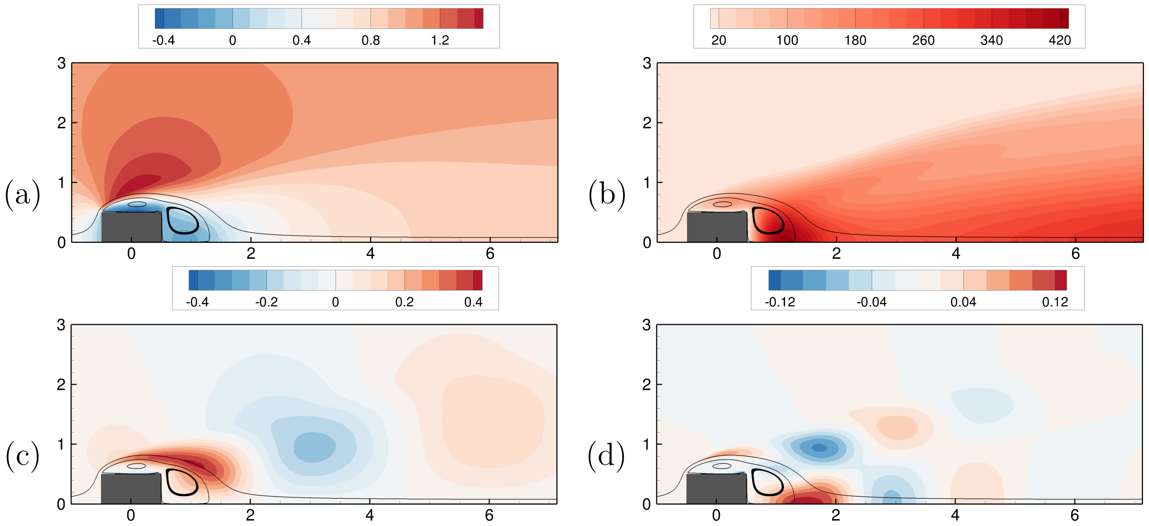

Figure 6.2: URANS solution: $\overline{U}_x$ mean-velocity field (a), mean-eddy-visocisty $\tilde{N}/\nu$ (b), first (c) and second (d) harmonic of velocity field. The ratio between their norms is $|A_0^{URANS}|/|A_1^{URANS}| = 6.7$

The solution of those equations corresponds to a saturated periodic limit-cycle, whose mean-flow (both velocity and eddy-viscosity) and nonlinear fluctuation, given by the first and second harmonics are presented in figure 6.2. The mean-flow similar to the one coming from the DNS, with $L \approx 1.4$. It is interesting to see that the nonlinear frequency of this solution is very close to the one coming from the DNS ($\omega \approx 0.83$). Also, the structure of the first harmonic is very similar as well to the DNS one. Those results are in accordance to the work of Bosch and Rodi [1998]. We remark that most of the energy of the fluctuation is concentrated on the first harmonic (at ω_0), the second one having a norm 6.7 times weaker than the first one. We can see thus that this URANS framework provides a good basis for capturing the vortex-shedding mode and for modeling the broadband structures. For this reason, it will be used as guideline for the construction of the model in the following. This model will focus on coupling the mean-flow equations with the vortex-shedding mode only, in a very similar way we did for the laminar square-cylinder case, present in the fourth chapter. However, before doing so, we investigate the predictability power of the resolvent formalism in this case.

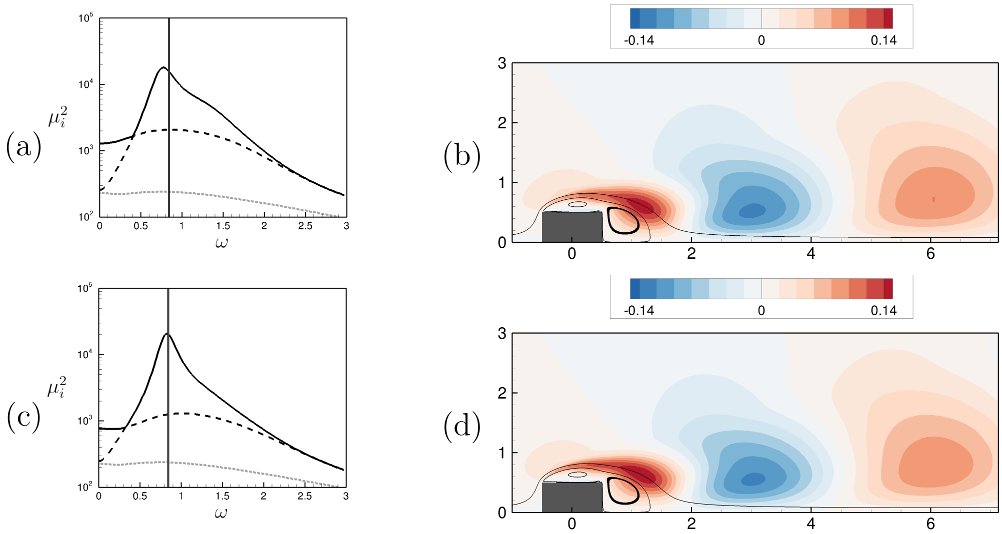

Figure 6.3: resolvent Analysis on (URANS) mean-flow: gains and leading mode at $\omega = 0.83$ (grey vertical line) for (a,b) frozen ν_t and (c,d) variable ν_t.

2.2 URANS Mean-Flow resolvent Analysis

In order to assess the effectiveness of the resolvent analysis in describing the vortex-shedding mode, we perform here two distinct ones: the frozen-eddy-viscosity, where only the momentum and continuity equations are taken in the Jacobian (where the fluctuation is taken under the form $(\mathbf{U}', P') = (\hat{\mathbf{U}}, \hat{P})e^{\mathrm{i}\omega t}$), and the variable eddy-viscosity one, where the SA model is linearized and included in the Jacobian as well (and the fluctuation is taken under the form $(\mathbf{U}', P', \tilde{N}') = (\hat{\mathbf{U}}, \hat{P}, \hat{N})e^{\mathrm{i}\omega t}$). Those analyses arise when we decompose the time-dependent variables $\mathbf{U}(t), P(t)(, \tilde{N}(t))$ into their mean-flow $\overline{\mathbf{U}}, \overline{P}, \overline{N}$ and fluctuation $\mathbf{U}', P'(, \tilde{N}')$, in a similar way we did in the last chapter. Note that for the variable eddy-viscosity case, the linearization of the model needs to be done in order to close the Jacobian matrix, which will now contain a variation $\tilde{N}'$ in the state vector. We remark that, for this data-assimilation model, we will stay consistent with the fourth chapter modeling, meaning that we will perform resolvent analysis for the recovery of the vortex-shedding mode rather than a classical stability analysis. The reason for this is that, in a way, the resolvent approach can be more general, allowing one to impose the exact frequency coming from a time-resolved probe, avoiding the possible mismatch of frequencies presented in the last chapter. Furthermore, in configurations that are dominated by only one globally unstable mechanism, those approaches lead to similar results (see Symon et al. [2018]).

We can see from figure 6.3 that the gains produced from both approaches are very close together and have a maximal value for nearly the same frequency. The maximal frequency for the frozen ν_t approach is slightly lower than the one coming from the DNS, as it was the case for the DNS-based eddy-viscosity analysis discussed in the last chapter. However, despite this small difference, we can see that the structure of the modes (both computed at $\omega_0 = 0.83$) is virtually the same. This means that, for modeling the vortex-shedding purposes, the frozen eddy-viscosity may suffice. Turning now our attention to the eddy-viscosity fluctuation, in the case where it is modeled with the linearized SA equation, we can see that its structure is not corresponding to the one coming from the

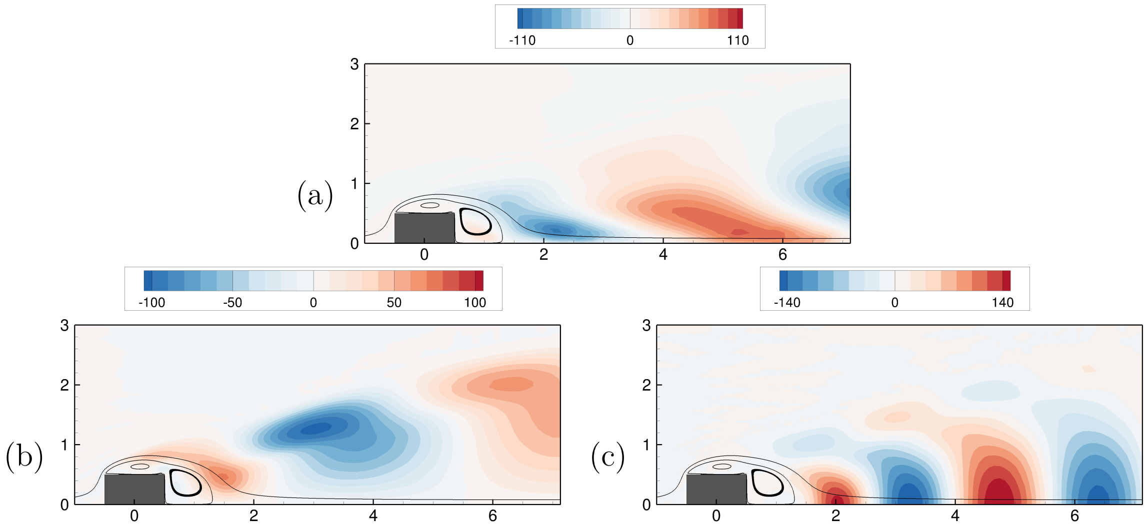

Figure 6.4: (a) resolvent mode for $\hat{N}$ (variable $\tilde{N}$ analysis). We compare it with the URANS first (b) and second (c) harmonics of $\tilde{N}$.

URANS simulations (see figure 6.4). Moreover, it is interesting to notice as well that, in the URANS computations, the second harmonics $(2\omega_0)$ of this quantity is of the same order of magnitude as the first one. This shows that the dynamics of the eddy-viscosity is more complex than the one for the velocity field only. With this in mind, we are ready to design our model for data-assimilation.

2.3 Data-Assimilation Model

In order to mimic the behavior of the URANS equations in a minimal way, the proposed model will be composed by the mean-flow moment equation, coupled with the SA model for the mean turbulent variable, together with the resolvent mode, modeling the vortex-shedding. The complete nonlinear model is given by:

$$\begin{aligned}
\overline{\mathbf{U}} \cdot \nabla \overline{\mathbf{U}} + \nabla \overline{P} - \nabla \cdot \Big((\nu + \nu_t(\overline{N})) \nabla_s \overline{\mathbf{U}} \Big) &= -2|A|^2 \mathrm{Re}\{\hat{\mathbf{U}}_0^* \cdot \nabla \hat{\mathbf{U}}_0\} && (6.7)\\
\nabla \cdot \overline{\mathbf{U}} &= 0 && (6.8)\\
\overline{\mathbf{U}} \cdot \nabla \overline{N} - \nabla \cdot (\eta(\overline{N}) \nabla \overline{N}) &= s(\overline{N}, \nabla \overline{N}, \nabla \overline{\mathbf{U}}) && (6.9)\\
R(\overline{\mathbf{U}}, \overline{N}, \omega_0) R^\dagger(\overline{\mathbf{U}}, \overline{N}, \omega_0) \hat{\mathbf{U}}_0 &= \mu_0^2 \hat{\mathbf{U}}_0,
\end{aligned}$$

where the fluctuation of the velocity field is approximated by:

$$\mathbf{U}' = A \hat{\mathbf{U}}_0 e^{i\omega_0 t} + c.c. \tag{6.10}$$

and the operator $R = R(\overline{\mathbf{U}}, \overline{N}, \omega)$ corresponds to the linearized Navier-Stokes operator around the mean-flow $\overline{\mathbf{U}}$ with the (frozen) eddy-viscosity $\overline{N}$. We notice that the non-linearities in the mean-momentum equation (and in the SA equation for $\overline{N}$) coming from the fluctuation of the eddy-viscosity are neglected since we do not have a good estimation for it, as discussed in the last paragraph. Moreover, the second-harmonics of the velocity field is also neglected, since its energy is 6.7 times lower for the URANS simulation and 8.6 times lower for the DNS. The only non-linearity taken into account in this model, other than those already existent for the RANS-SA model itself, is the advection in the momentum equation, similarly to what was proposed in the fourth chapter. Note that, although we base our model on a URANS model, the amplitude of the vortex-shedding mode is not necessarily the same one coming from the URANS run, and will be tuned using external data. This gives the model and external parameter for it to adapt to the DNS solution in a way that the URANS model itself cannot. In the next section, we will present some results on the dependency of this model's solution with respect to the parameter $|A|$, also tuning the energy of the mode. We remark that, since the model consists of a system of equations very similar to the one presented in the fourth chapter, we use the same strategy to solve it numerically. This means we solve it with a Newton method, once we rewrite the system as a set of nonlinear equations. It is worth-mentioning that, in order accelerate the resolution of the model, only a half-domain mesh was employed, where, at the center-line, symmetry conditions were used for the mean-flow and anti-symmetry conditions were used for the vortex-shedding mode.

3 Results

In this section, we study how the solution of the previously introduced model varies with $|A|$, in the same manner we did for the laminar case, where a parametric study is employed, shown in figure 6.5 (continuous black lines). We can see that, by increasing the value of $|A|$, the mean-flow related quantities represent better the actual DNS ones. The re-circulation bubble becomes shorter and the mean-flow error (with respect to the DNS mean-flow) also decreases and creates a minimum for $|A| \approx 2.18$. We can see as well that the quantities related to the fluctuation do not match the DNS ones necessarily at the same value of $|A|$. For example, if we look at the magnitude of the resolvent mode at a point close to the cylinder ($|A||\hat{\mathbf{U}}_0|$), we see that the model does not manage to react fast when $|A|$ is small. This issue was already discussed in the fourth chapter and it is related to the fact that, at this range of parameter, the bubble is still too long and the resolvent mode does not have important values close to the cylinder. However, for points far from the cylinder, this quantity is almost linear. We can see that even quantities related to the resolvent mode can hide a mean-flow dependency.

Another interesting aspect of the model would be to study how the eddy-viscosity, produced naturally by the SA model, evolve with $|A|$. We can see from figure 6.6 that this quantity becomes weaker with an increase in $|A|$, meaning that the model adjusts itself in order to produce less eddy-viscosity whenever the saturation mechanism coming from the vortex-shedding mode becomes important. Note that the effect of the eddy-viscosity on the mean-flow is similar to the nonlinear forcing term, coming from the vortex-shedding mode: an increase on both of them tends to reduce the re-circulation bubble. In our model, this interplay between them is somewhat fixed and given by the RANS-SA equation. However, this relation does not necessarily capture the actual physical behavior. Indeed, for all the mean-flow related quantities, the predicted value of $|A|$ was always much higher $|A| \approx 2.3$ than the DNS one $|A|^{DNS} = 1.58$ or the URANS one $|A|^{URANS} = 1.35$. It seems that the eddy-viscosity produced by our model is not sufficient to produce a short re-circulation bubble for lower values of $|A|$. Indeed, if we compare the eddy-viscosity in figure 6.6 (d) (for which the predicted re-circulation bubble is not too far from the URANS one) with the mean-eddy-viscosity from URANS in figure 6.2 (b), we can see that we under-predict this quantity. To investigate how a different value

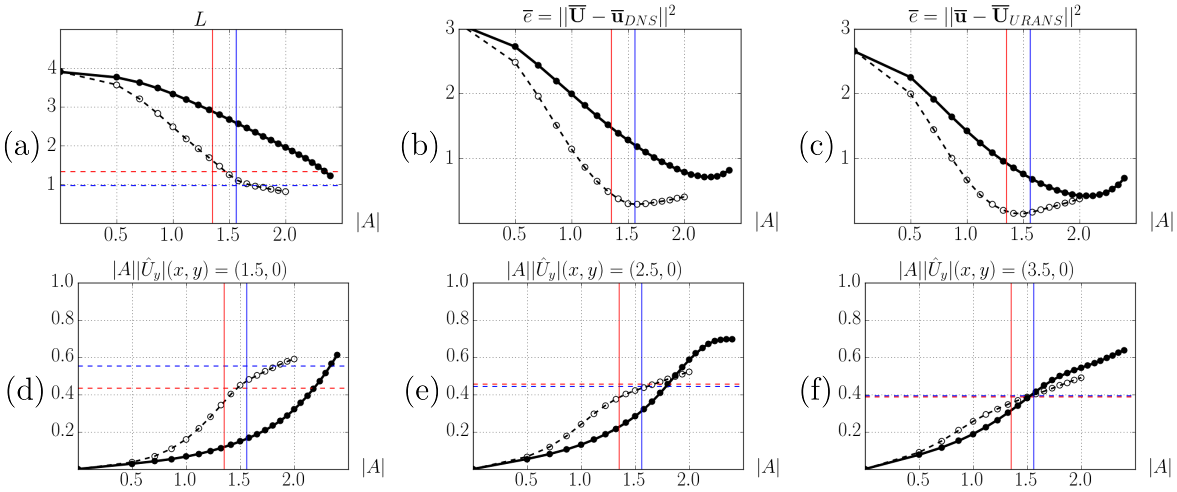

Figure 6.5: Solution of the model in function of $|A|$. Mean re-circulation bubble length (a), global error of velocity field with respect to the mean-DNS (b) and mean-URANS (c), amplitude of Fourier mode of the cross-velocity signals at $(1.5, 0)$ (c) $(2.5, 0)$ (d) and $(3.5, 0)$ (e). Black continuous curves correspond to the model given by (6.7) and dashed ones to the comparative model where the ν_t is frozen to be the RANS-SA one. DNS value of $|A| \approx 1.58$ on blue vertical line and URANS value of $|A| \approx 1.35$ on red vertical lines. Exact values of each appropriate quantities are given in dashed horizontal lines for DNS (blue) and URANS (red).

of eddy-viscosity would produce a different solution, we modify the model into a simpler one, where we no longer we solve the SA equations. Instead, we take the eddy-viscosity to be fixed with $|A|$ and equal to the RANS-SA one, which is much higher (see figure 6.6 (a)). The results for this model are also presented in figure 6.5 with a dashed line. We can see that the mean flow quantities are predicted in the proper range of $|A| \approx 1.6$, where the mean-flow error is even smaller with this simplified model. Furthermore, the resolvent mode-related quantities are also better represented at the same range of values $|A| \approx 1.5 - 1.8$. This suggests that the SA equation, taking into account only mean-flow quantities, may not be a good approximation and should also include nonlinear terms on the fluctuation. Examples of such terms are the nonlinear advection $\overline{\mathbf{U}' \cdot \nabla \tilde{N}'}$, a nonlinear diffusion $\nabla \cdot \overline{(\partial_{\tilde{N}} \eta \tilde{N}') \nabla \tilde{N}'}$ and the (second-order Taylor expanded) nonlinear source term, exemplified by $\overline{\partial_{\tilde{N}} \partial_{\tilde{N}} s (\tilde{N}')^2}$. We note that taking into account those terms is a challenging task since we do not have a proper approximation for $\tilde{N}'$ from the mean-flow resolvent analysis. Moreover, the last two terms rely on the truncation of a Taylor expansion, that may itself represent a limited description of the whole dynamics, especially if the fluctuation is large.

As for the reconstructed fields (mean-flow and vortex-shedding mode), the results are presented in figure 6.7 for values of $|A|$ for which the mean-flow error is minimal. We can see that the vortex-shedding modes, although they represent the vortex-emission, as in the DNS or URANS simulations, their structure can be a little "discontinuous" along some lines in the flow. This may be due to the fact that only the first harmonics is modeled and its nonlinear interactions with the second harmonics could correct locally the spatial distribution of the mode. We remark that this possible lack of the second harmonics in our model does not necessarily explain the over-prediction of the parameter $|A|$ discussed

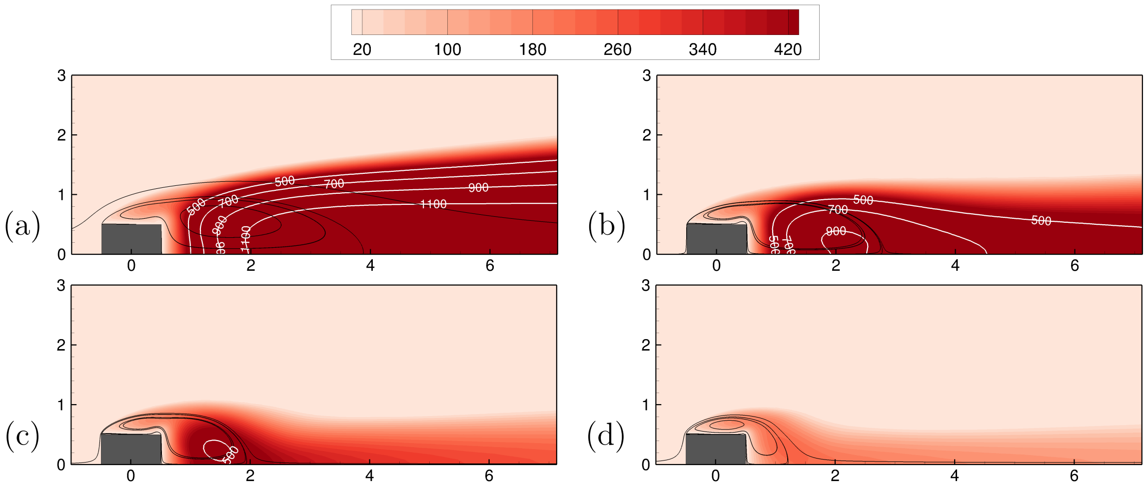

Figure 6.6: Results for reconstruction procedure: evolution of eddy-viscosity ν_t/ν for increasing values of $|A|$, $|A| = 0$ (a), $|A| = 1.4$ (b), $|A| = 2.0$ (c), $|A| = 2.4$ (d) . Iso-contours are saturated with mean-URANS eddy-viscosity color scale.

before since the ratio between the first and second harmonics energies is or order ≈ 7. This means that, if this second harmonics would have been accurately incorporated in the model, the amplitude $|A|$ could diminish by a factor of $\approx (1-1/7) \approx 0.86$, which still does not explain the factor $1 - |A|^{DNS}/|A| = 0.31$.

4 Conclusion

This chapter dealt with the resolvent-based data-assimilation for the flow around a square cylinder at Reynolds number $Re = 22000$, similarly to what was done in the fourth chapter, where the vortex-shedding mode is coupled to the mean-flow equations, using a URANS turbulence model to provide the turbulent eddy-viscosity. Although this model managed to reproduce the overall features of the flow, it was less successful than the one presented in the fourth chapter in the prediction of mean-flow quantities together with unsteady ones. This comes from the fact that the URANS model did not produce the necessary amount of eddy-viscosity such that, for example, the re-circulation bubble predicted the correct one at the right range of vortex-shedding mode amplitude. Since the fully nonlinear URANS equations actually predicted better the DNS one, we believe that the model proposed is over-simplified and should take into account more nonlinear terms. However, this is a challenging task because all he remaining terms depend on fluctuations of $\tilde{N}'$, which is poorly reconstructed from the resolvent analysis. As a possible remedy for this could be the use of a tuning source term in the SA equation, modeling all those remaining terms, similarly to what was done in the second and third chapters of this manuscript, where some other external data is considered in the assimilation.

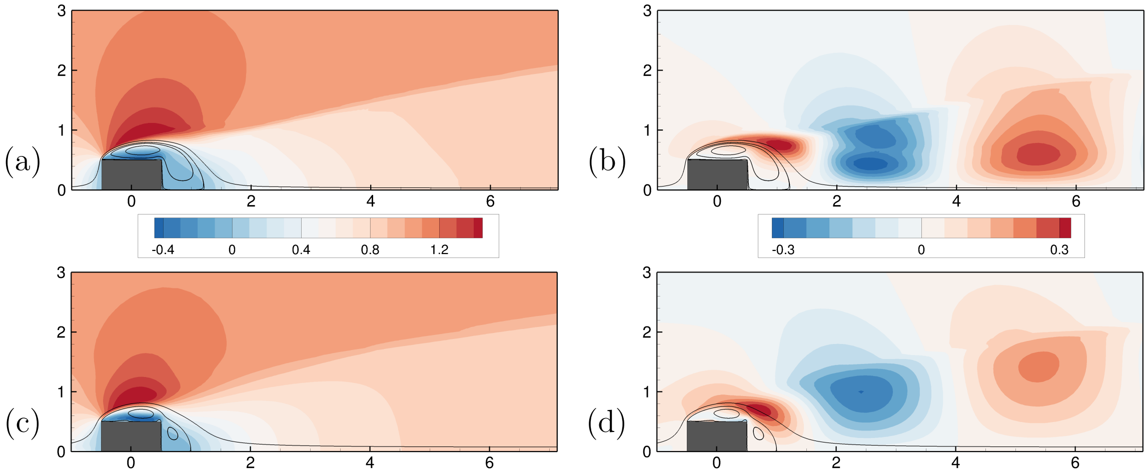

Figure 6.7: Results for reconstruction procedure: mean-flow (a) and normalized vortex-shedding mode (b) for the model given by (6.7). The same is presented for the model where the eddy-viscosity is fixed to be the RANS-SA one (c,d).

Chapter 7

Conclusions and Perspectives

In this manuscript, we investigated some modeling techniques for data-assimilation of aerodynamic flows. Those techniques can be divided into two large groups. The first one was the recovery of the time-averaged (mean) flow through minimal modifications in the RANS (Reynolds-Averaged Navier-Stokes) equations, performed by an optimization, minimizing the error committed by the solution and some mean-flow external data. This was the subject of chapters 2 and 3. The second one relied on the mean-flow analysis, mainly based on the Resolvent formalism, subject of the subsequent chapters. For the data-assimilation purposes, we were able to build a reduced order model given by the mean-flow equation coupled with the Resolvent modes that provide a representation of the fluctuation, from which we can compute the Reynolds-stress tensor, both in a transitional and turbulent scenarios. Also, in a slightly different context from data-assimilation, we provided a framework from which both periodic and broadband components could be recovered from linear analysis. In the following we will provide a quick overview of the methods and results discussed in the manuscript, together with ideas for future work for each of them separately.

1 RANS-based Data-Assimilation

From the point of view of the first one (RANS-based data-assimilation), we investigated two different models. The first (model 1) consisted by the tuning parameter on the momentum equations, which is an extension of what was done by Foures et al. [2014] to a turbulent regime. The second one (model 2) consisted by a source term in the turbulence model (Spalart-Allmaras, in the present case), correcting it. We investigated how the choice of tuning parameter and the choice of the measure operator would change the optimized procedure and the final solution. For example, we observed that, if the whole velocity field information is provided, the model 1 manages to adapt itself to represent correctly the whole velocity-field everywhere. For the model 2 however, we did not manage to capture the whole velocity-field and the overall error decreased of one order of magnitude only. This shows that this model is not fully controlable. However, when sparse velocity field is provided, the model 1 may produce some nonphysical behavior related to the sharpness of the gradients, and further penalization on the derivatives of the forcing term had to be considered to smooth it, leading to a more physical solution. This was shown not to be the case for the model 2, where the gradients were smoother and the reconstructed final solution, although did not necessarily match the provided data, could provide a better solution when very few measurements were provided. Moreover, we observed that the model 2 produced almost the same solution independently on the number of velocity measure points. We further explained this fact with an Observability Gramian analysis, showing that this model presents a much more rigid behavior, where only a few directions in the space of forcing terms are easily reachable/observable. This rigidity was also observed when, for example, we considered wall-pressure measurements, for example. In this case, we could show that, by considering enough wall measure points, the final solution tended towards the same one produced by the velocity-field information. However, for the model 1, even with complete wall-pressure information, we did not manage to outperform the model 2, showing that, in this case, wall-pressure is not enough.

With this in mind, we can propose a few directions for future studies. Firstly, the Spalart-Allmaras model showed to be rigid when we considered a corrective forcing term and this behavior could modify the final assimilated solution. For this reason, it would be interesting to explore in a future work how other turbulence models would perform. If we consider, say, the $k-\omega$ model instead of the SA, we could observe a different behavior. For example, the model could present itself being a little less rigid than the SA, opening more room for the solution to adapt to the data, if sufficient information is available. Alternatively, even if this model ends up being as rigid as the SA, it would be interesting to evaluate if it manages to produce an assimilated solution that is closer to the reference one than the one with SA. This would mean that, even if the Observability Gramian analysis still produces a large separation between the leading Singular Values and the remaining ones (meaning that the model is still poorly observable), the leading Singular Vectors may "point" towards a better direction, on which the DNS field may lie. This would favor the choice for this model instead of SA, since it appears to be more physical.

Another interesting direction would be the consideration of turbulence models that do not rely on Boussinesq hypothesis for the data-assimilation. This was already the goal of the model 1, where a free volume force was used as tuning parameter. However, the restriction of this tuning parameter within such turbulence model would perhaps limit the action of the actual forcing term to a more physical direction, improving the assimilated solution when few (or localized) measurements are provided, always retaining the (to be verified) fully controlability. This kind of model would be useful in bluff-body-like configurations that do not satisfy the Boussinesq hypothesis and for which the volume-forcing term was not capable to correctly assimilate, for example, the wall-pressure (as in the cylinder flow presented in the third chapter). Examples of such models are the Reynolds-stress tensor models (RSM, see for example Hanjalić and Launder [1972]), where the Reynolds-stress tensor is not replaced by an eddy-viscosity and will be modeled by a supplementary set of PDEs, typically one for each component of the tensor and another one for the dissipation rate. In this case, the forcing/tuning parameter would be, for example, a volume term in each of those extra equations.

2 Mean-flow analysis and Data-Assimilation

As for the Resolvent-based data-assimilation (presented in the fourth chapter), inspired on the idea that the Resolvent modes can approximate the fluctuation around the mean-flow

(McKeon and Sharma [2010], Beneddine et al. [2016]), we proposed a model composed by the mean-flow equation coupled with leading Resolvent modes through the Reynolds-stress tensor. This model represents a nonlinear set of equations, whose only unknown parameters are the energies (or amplitudes) of each of those modes, one for each modeled frequency. Since the unknowns are intimately related to the spectral (frequency) content of the fluctuation, those amplitudes can be tuned by assimilating, for example, few point-wise time-resolved probes. From an experimental point of view, this represents a gain, with respect to the RANS-based data-assimilation, in the amount of information needed to recover, not only the mean-flow but also the fluctuation around it. This procedure was first applied in three different laminar/transitional configurations: squared-section cylinder and the backward-facing step forced by two incommensurable frequencies and with (time) white-noise. Generally, what was observed in those cases is that the success of the reconstructed solution of this model, together with its robustness with respect to the choice of point-wise data, can be assured if the mean-flow Resolvent modes do manage to reconstruct the main features of the fluctuation and if all of the relevant frequencies are taken into account. This tends to be the case when the reference flow does not present strong nonlinearities. An example of flows having an important amount of nonlinearities is the two-frequencies backward-facing step where, even if the external forcing contains only two frequencies, the fluctuation contains more frequencies arising from nonlinear interactions between the two fundamental ones. The presented model did not consider those nonlinear interactions, limiting the quality of the reconstruction of the final solution. In those cases, we may employ a finer model that captures those nonlinear terms. An example of such approach is given by Symon et al. [2019a], where the fluctuation is solved in a nonlinear manner for a given (frozen) mean-flow, assimilated with similar techniques presented in the first two chapters of this manuscript. For this reason, the application of this procedure is much more challenging in our case, where the mean-flow is also an unknown of the problem. The inclusion of nonlinearities in the model of the fluctuation can thus be considered as subject of future work.

As for the turbulent case, corresponding to the square cylinder at $Re = 22000$, we considered first, in the fifth chapter, the mean-flow analysis (in a broad sense) for the recovery of both periodic mode (vortex-shedding) and the coherent broadband component (Kelvin-Helmholtz) that characterize this configuration and other flows around bluff bodies. We showed that those components could be isolated using one version of the triple decomposition, proposed by Mezić [2013]. We could recover the periodic component with the usual mean-flow stability analysis, where, in order to model the broadband component's effect on the quasi-periodic one, we considered an eddy-viscosity computed from the statistics of a DNS. We showed as well that the dependency of the broadband component (here exemplified by the Kelvin-Helmholtz instability) on the quasi-periodic component could be captured by the Resolvent operator, where the flow around which the equations were linearized was the quasi-periodic one, in a similar manner as in Floquet stability theory. Those modes showed themselves predictive when compared with DNS snapshots, where the dependency of the Kelvin-Helmholtz structures on the oscillation of the shear-layer is observed. For future work, it would be interesting to test this periodic Resolvent technique on a flow configuration where the frequency of the periodic motion is closer to the broadband structures one and the complete resolution of the modes with, for example, the Floquet-Fourier-Hill technique would be required.

In the sixth chapter, we applied the Resolvent-based data-assimilation, tested for laminar flows, in a turbulent one, consisting also in the square cylinder, same configuration as in the fifth chapter. Since this configuration is much more complex than those discussed in the fourth chapter, we proposed the use of a URANS model, from which we could use the eddy-viscosity, necessary for the prediction of the vortex-shedding mode and for the mean-flow equation saturation. The model consisted then of the mean-flow equations (momentum equations, together with the Spalart-Allmaras one) coupled through the nonlinear advection term with the resolvent mode modeling the vortex-shedding. This means that this model did not resolve the coherent broadband component, since it would leave the model much more complex. We observed that the solution of this model, although it captured the overall feature of the reference flow, did not predict well the range of amplitudes of the fluctuation of the reference flow (DNS or URANS). We could also show that the balance between this amplitude and the eddy-viscosity was not quite the one present in the URANS solution (and arguably in the DNS one). We believe the reason for that was an over-simplification of the model where, the only nonlinear term considered for the fluctuation was the nonlinear advection. The reason why the consideration of other nonlinear terms in the model is difficult is that the remaining ones are all dependent on the fluctuation of the eddy-viscosity, which is poorly represented by the mean-flow analysis when compared to the URANS one. Furthermore, the first harmonics of this quantity is of the same order of the second one, suggesting that there may be strong nonlinear interactions for this variable and a more complex dynamics for it needs to be used. One possible direction for future work could be the design of a such more refined model for this variable and its further use in the mean-flow equations through nonlinear terms. Another possible solution would be the use of of a tuning parameter such as a source term in the mean SA equation (similarly to what was done in chapters 2 and 3) that corrects this erroneous balance between the eddy-viscosity production and the vortex-shedding amplitude and its further determination with an optimization procedure. An obvious drawback of this approach would be the need for more data to be assimilated since now we tune a spatially-dependent function as well.

Appendix A

RANS - Spalart-Allmaras Numerical Resolution

1 Incompressible RANS

In this section we are interested in giving some details on the (incompressible) RANS equations, coupled with the Spalart-Allmaras model, together with its resolution in a Finite-Element code, FreeFem++ (see Hecht [2012]). They are given by:

$$\partial_t \mathbf{u} + \mathbf{u} \cdot \nabla \mathbf{u} + \nabla p - \nabla \cdot ((\nu + \nu_t) \nabla_S \mathbf{u}) = \mathbf{0} \tag{A.1a}$$

$$\nabla \cdot \mathbf{u} = 0 \tag{A.1b}$$

$$\partial_t \tilde{\nu} + \mathbf{u} \cdot \nabla \tilde{\nu} - \nabla \cdot (\eta(\tilde{\nu}) \nabla \tilde{\nu}) = s(\tilde{\nu}, \nabla \tilde{\nu}, \nabla \mathbf{u}) \tag{A.1c}$$

The last equation in A.1 is a nonlinear Advection-Diffusion-Reaction equation for the auxiliary quantity $\tilde{\nu}$, such that $\nu_t = \nu_t(\tilde{\nu})$. In the next subsection, we introduce the original Spalart-Allmaras model such that it was first proposed (Spalart and Allmaras [1994]). Then, we will comment on its modifications, in order to make it more stable and differentiable.

1.1 Original model

The dependency of ν_t with respect to $\tilde{\nu}$ is given by:

$$\nu_t = \tilde{\nu} f_{v1}, \quad f_{v1} = \frac{\chi^3}{\chi^3 + c_{v1}^3}, \quad \chi = \frac{\tilde{\nu}}{\nu}, \quad \eta = \sigma^{-1}(\nu + \tilde{\nu}) \tag{A.2}$$

Moreover, the nonlinear term $s(\tilde{\nu}, \nabla \tilde{\nu}, \nabla \mathbf{u})$ is composed by a production, cross-diffusion and destruction terms, respectively given by:

$$s = c_{b1} \tilde{S} \tilde{\nu} + \sigma^{-1} c_{b2} |\nabla \tilde{\nu}|^2 - c_{w1} f_w \left(\frac{\tilde{\nu}}{d} \right)^2 = s_P + s_C + s_D \tag{A.3}$$

each of those terms correspond to nonlinear functions on the state, which are further closed with the extra auxiliary parameters:

$$\tilde{S} = S + \frac{\tilde{\nu}}{\kappa^2 d^2} f_{v2}, \quad f_{v2} = 1 - \frac{\chi}{1 + \chi f_{v1}}, \qquad S = |\nabla \times \mathbf{u}| \tag{A.4a}$$

$$f_w = g \left[\frac{1 + c_{w3}^6}{g^6 + c_{w3}^6} \right]^{1/6}, \quad g = r + c_{w2}(r^6 - r), \qquad r = \frac{\tilde{\nu}}{\tilde{S} \kappa^2 d^2}, \tag{A.4b}$$

with the constants:

$$\sigma = 2/3, \quad c_{b1} = 0.1355, \quad c_{b2} = 0.622 \tag{A.5a}$$

$$\kappa = 0.41, \quad c_{w1} = c_{b1}/\kappa^2 + (1 + c_{b2})/\sigma, \quad c_{w2} = 0.3 \tag{A.5b}$$

$$c_{w3} = 2, \quad c_{v1} = 7.1, \tag{A.5c}$$

The original model may have some issues when one tries to derive the Jacobian operator corresponding to the nonlinear model, such as non differentiable terms or numerical problems related to the ill-posed structure of the resulting matrix. In the next subsections, we will document all the modifications made and a brief explanation of them.

1.2 Modification in S

The simplest modification made was to replace the term $S = |\nabla \times \mathbf{u}|$ by $S = \sqrt{|\nabla \times \mathbf{u}|^2 + M^2} - M$, where M is a user defined parameter (typically $M = 10^{-4}$). This new definition of S makes it differentiable, retaining the property $S > 0$ and $S = 0$ iif $\nabla \times \mathbf{u} = \mathbf{0}$.

1.3 Negative $\tilde{\nu}$ modifications

In the original model, it is possible to observe some undesirable phenomena that may not occur in a continuous framewor, but do occur in a discrete framework. One of them is the apparition of negative values of $\tilde{\nu}$ (Crivellini et al. [2013], Oliver [2008]) at the thin layer that transitions from the turbulent to the free-stream region. This phenomenon, coupled with a high-order representation may lead to highly oscillatory and unstable behavior (Persson et al. [2006], Oliver and Darmofal [2007]).

A differentiable way to overcome this problem was given by Oliver [2008], in the context of Discontinuous Galerkin Method discretization. His idea was to modify the original $\tilde{\nu}$ equation in such way that the energy of the turbulent variable $\tilde{\nu}^2/2$ to reduce when $\tilde{\nu}$ in negative. For this purpose, let us derive the equation for this quantity:

$$\tilde{\nu}\partial_t\tilde{\nu} + \tilde{\nu}\nabla\cdot(\mathbf{u}\tilde{\nu}) = \frac{1}{2}\left(\partial_t\tilde{\nu}^2 + \nabla\cdot(\mathbf{u}\tilde{\nu})\right) = \frac{\tilde{\nu}}{\sigma}\nabla\cdot(\eta\nabla\tilde{\nu}) + \tilde{\nu}s. \tag{A.6}$$

Let us define our domain Ω and let us divide it into two parts, one of them containing positive Ω^+ and negative Ω^- values of $\tilde{\nu}$. If we assume that $\partial\Omega \cap \partial\Omega^- = \emptyset$, meaning that the negative values of $\tilde{\nu}$ only happen at the interior of Ω, then:

$$\frac{1}{2}\int_{\Omega^-} \partial_t\tilde{\nu}^2 + \frac{1}{2}\int_{\partial\Omega^-} \mathbf{u}\tilde{\nu}^2 - \int_{\partial\Omega^-} \frac{\tilde{\nu}}{\sigma}\tilde{\nu}\nabla\tilde{\nu}\cdot\mathbf{n} + \int_{\Omega^-} \eta\frac{\nabla\tilde{\nu}}{\sigma}\cdot\nabla\tilde{\nu} = \int_{\Omega^-} \tilde{\nu}s \tag{A.7}$$

The second and third terms are null since, by continuity, $\tilde{\nu}|_{\partial\Omega^-} = 0$. By opening the source term s into its three terms, we have:

$$\frac{1}{2}\int_{\Omega^-} \partial_t\tilde{\nu}^2 = \int_{\Omega^-} \frac{c_{b2}\tilde{\nu} - \eta}{\sigma}\nabla\tilde{\nu}\cdot\nabla\tilde{\nu} + \int_{\Omega^-} c_{b1}\tilde{S}\tilde{\nu}^2 - \int_{\Omega^-} c_{w1}f_w\frac{\tilde{\nu}^3}{d^2} \tag{A.8}$$

Since the convective fluxes are null, the above equation represents the "energy" of $\tilde{\nu}$ in the region where it is negative in a control volume Ω^- moving with the flow. This means that, for the energy to be positive if $\tilde{\nu} < 0$, we have to ensure the right-hand-side to be negative. For that matter, we start by choosing a modification on η:

$$\eta = \begin{cases} \nu(1+\chi), & \chi \geq 0 \\ \nu(1+\chi+\chi^2/2), & \chi < 0 \end{cases} \tag{A.9}$$

ensuring its positivity and ensuring the first integral to be negative if $\tilde{\nu} < 0$. The sign of term $\tilde{\nu} s_P = c_{b1}\tilde{S}\tilde{\nu}^2$ is only dictated by the sign of $\tilde{S}$, that is positive (maybe unless $1 < \chi < 18.4$). For this reason, the following modification is used:

$$s_P = \begin{cases} c_{b1}\tilde{S}\tilde{\nu}, & \chi \geq 0 \\ c_{b1}S\tilde{\nu}g_n, & \chi < 0 \end{cases} \tag{A.10}$$

where:

$$g_n = 1 - \frac{10^3\chi^2}{1+\chi^2}. \tag{A.11}$$

We remark that this modification does not always ensure $\tilde{\nu} s_P < 0$ when $\tilde{\nu} < 0$, since g_n has the "wrong" sign if $-\sqrt{1/999} < \chi < 0$. However, this small interval is unavoidable if the differentiability of the model has to be retained, since $s_P|_{\chi=0} = 0$ and $\partial_\chi s_P|_{\chi=0} = c_{b1}S\nu \geq 0$. Finally, the last term $\tilde{\nu} s_D = c_{w1} f_w \tilde{\nu}^3/d^2$ is negative ("wrong" sign) if $f_w < 0$, which happens for $-1.185 < r < 0$. Therefore, another modification is made:

$$s_D = \begin{cases} c_{w1} f_w \frac{\tilde{\nu}}{d^2}, & \chi \geq 0 \\ -c_{w1}\frac{\tilde{\nu}}{d^2}, & \chi < 0 \end{cases} \tag{A.12}$$

Rewriting the whole nonlinear forcing term:

$$s = \begin{cases} c_{b1}\tilde{S}\tilde{\nu} & + & c_{b2}\sigma^{-1}|\nabla\tilde{\nu}|^2 & - & c_{w1} f_w \frac{\tilde{\nu}}{d^2}, & \chi \geq 0 \\ c_{b1}S\tilde{\nu}g_n & + & c_{b2}\sigma^{-1}|\nabla\tilde{\nu}|^2 & + & c_{w1}\frac{\tilde{\nu}}{d^2}, & \chi < 0 \end{cases} \tag{A.13}$$

In addition, inspired by the η modification and from the fact that negative diffusivity coefficients may lead to an ill posed behavior, we also modify the turbulent viscosity ν_t as:

$$\nu_t = f_{v1}\max(0, \tilde{\nu}) \tag{A.14}$$

This modification seems to be contrary to the general purpose of differentiability of the model, however it is not the case, since the derivative of $f_{v1}\tilde{\nu}$ approaches 0 when $\chi \to 0$.

1.4 Negative r modifications

We now discuss another issue with the original SA model, which is the possibility of having negative values of the originally positive definite quantity r, being related to a mixing length l such as $r = l^2/(\kappa^2 d^2) > 0$. This phenomenon can be better understood by rewriting this function as:

$$r = \frac{\tilde{\nu}}{\tilde{S}(\kappa^2 d^2)} = \frac{\chi}{\frac{S(\kappa^2 d^2)}{\nu} + \chi f_{v2}}. \tag{A.15}$$

We see that the function r is function of the two non dimensional groups $\chi, S(\kappa^2 d^2)/\nu$, the first being responsible for negative r, since the function f_{v2} can be negative for $1.003 <$

$\chi < 18.4$. Moreover, it has a singular behavior for $\chi f_{v2} = -S(\kappa^2 d^2)/\nu$. A . As we will see, the phenomenon of negative r (including its sigularity) may be a symptom of other anomolous behavior. Indeed, if we write the nonlinear production term, say, $s_{PD} = s_P + s_D = c_{b1}\tilde{S}\tilde{\nu} - c_{w1} f_w (\tilde{\nu}/d)^2$ as:

$$s_{PD} = \left(\frac{\nu\chi}{d}\right)^2 \left(\frac{c_{b1}}{\kappa^2 r} - c_{w1} f_w\right) \tag{A.16}$$

we can see that the production term inverts its signals for $r < 0$. Moreover the destruction term also inverts its signal if $f_w < 0$, which corresponds to the region $-1.185 < r < 0$. This phenomenon may seem a small deviation of the "normal behavior" of the model, which tends to produce solutions for which $r > 0$. However, it may not be the case in the convergence process, which in this work, is based on the Newton method. Indeed, the term $\partial s_{PD}/\partial\tilde{\nu}$ presents an oscillatory behaviour at the regions where $r < 0$ (see figure A.1). This behaviour deteriorates drastically the convergence of the Newton algorithm (Crivellini et al. [2013]).

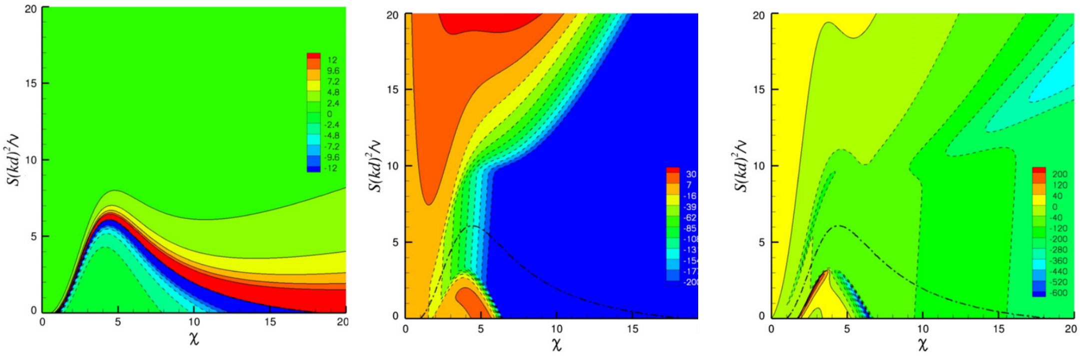

Figure A.1: original SA parameters: r, s_{PD} and its Jacobian $\partial s_{PD}/\partial\tilde{\nu}$ as function of the two non-domensional groups χ and $S(\kappa d)^2/\nu$. The dashed lines on the plots represent the set of parameters for which r is singular. Figure extracted from Crivellini et al. [2013].

To correct this behavior, we limit its value to r_{max} such that:

$$r^* = \frac{\tilde{\nu}}{\tilde{S}\kappa^2 d^2}, \qquad r(r^*) = \begin{cases} r_{max} & r^* < 0, \\ r^* & 0 \leq r^* < r_{max}, \\ r_{max} & r^* \geq r_{max}. \end{cases} \tag{A.17}$$

This function is not differentiable (nor continuous) if we see it as $r = r(r^*)$. However, what wee need to have is a differentiable $r = r(\tilde{\nu}, \mathbf{u})$ or $r = r(S(\kappa^2 d^2)/\nu, \chi)$, as evoked before.

On can from figure A.1 see that, below the singular line, the value of r^* is negative. Thus, in those regions, $r = r_{max}$. The same is true if we approach the singular line from the positive r^* region. We will have an increasing value of r^*, which also translates into $r = r_{max}$. This means that at the originally singular line, the function r is differentiable (and constant, $r = r_{max}$). Moreover, the only source of non differentiability is when $r^* = r_{max}$. At this region, the model is continuous, but not rigorously differentiable.

However, one can calculate how much is the jump in $\partial s/\partial\tilde{\nu}$ at $r^* = r_{max}$:

$$\frac{\partial s}{\partial\tilde{\nu}}\Big|_{r^*=r^+_{max}} = c_{b1}\left(\tilde{S} + \tilde{\nu}\frac{\partial\tilde{S}}{\partial\tilde{\nu}}\right) - c_{w1}\left(\frac{df_w}{dr}\underbrace{\frac{\partial r}{\partial\tilde{\nu}}_{r^*=r^+_{max}}}_{0}\left(\frac{\tilde{\nu}}{d}\right)^2 + 2f_w\frac{\tilde{\nu}}{d^2}\right) \tag{A.18}$$

$$\frac{\partial s}{\partial\tilde{\nu}}\Big|_{r^*=r^-_{max}} = c_{b1}\left(\tilde{S} + \tilde{\nu}\frac{\partial\tilde{S}}{\partial\tilde{\nu}}\right) - c_{w1}\left(\frac{df_w}{dr}\frac{\partial r}{\partial\tilde{\nu}}_{r^*=r^-_{max}}\left(\frac{\tilde{\nu}}{d}\right)^2 + 2f_w\frac{\tilde{\nu}}{d^2}\right) \tag{A.19}$$

So that the jump itself on the quantity $\partial s/\partial\tilde{\nu}$ is:

$$\frac{d^2}{\nu}\left[\frac{\partial s}{\partial\tilde{\nu}}\right] = c_{w1}\,\frac{df_w}{dr}\Big|_{r^*=r^-_{max}} r_{max}\left(1 - r_{max}\left(f_{v2} + \chi\frac{df_{v2}}{d\chi}\right)\right)\chi \tag{A.20}$$

This jump function depends directly on the factor $df_w/dr|_{r^*=r^-_{max}}$. However, for $r > 2$ the function f_w stays at a plateau, and its derivative is almost zero. This means that, for a sufficiently large value of r_{max}, the jump is almost nonexistent. In our work, this value is set to be $r_{max} = 10$, as in Crivellini et al. [2013].

With this modification, the s_{PD} function, together with its Jacobian are much more smoother and lead to a much better Jacobian matrix (see figure A.2). We remark that the production term s_P was not in our code written in function of r, as presented before. For this reason, the negative-production problem was not solved by this approach. However, it was not observed to lead any convergence problems, as documented by A.1.

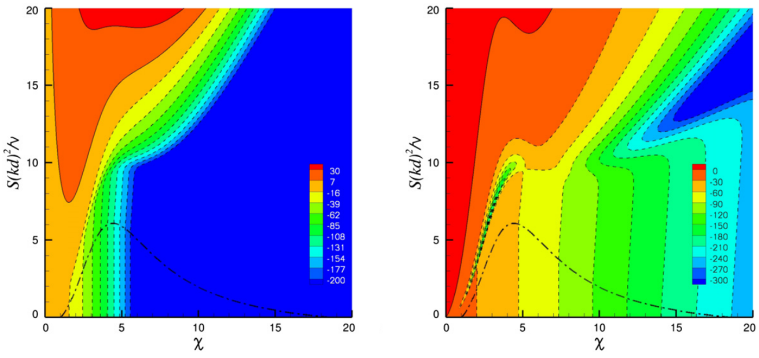

Figure A.2: r-Modified SA model: s_{PD} and its Jacobian $\partial s_{PD}/\partial\tilde{\nu}$. Figure extracted from Crivellini et al. [2013].

2 Stabilization procedure - SUPG

Contrarily to finite volume, finite differences and, more recently, discontinuous Galerkin methods, the Finite Element Method (FEM) does not provide a natural framework in which one can add some '*upwind*' effect, which stabilizes the numerical scheme for high Reynolds number flows. However, this effect can be achieved by modifying the variational formulation through a method called Streamline-Upwind Petrov-Galerkin (SUPG). More precisely, it was observed that the modification of the test function of a classical variational formulation in a precise way leads to stabilization of the scheme (Brooks and Hughes [1982]). Several forms of such modifications are proposed in the literature (see Franca et al. [1992], Franca and Frey [1992]). The main idea is to modify the test function of the FEM through the advection operator, such that the scheme becomes decentralized and stable. For the RANS equations, for simplicity, we only apply this modification of the test function on the advection operators. This leads to a simpler variational formulation than the ones from, for example, Franca et al. [1992]. However, this is extensively done in the context of unsteady simulations, without loss in precision (see, for example Bao et al. [2011], Haferssas et al. [2018]). The variational formulation employed is:

$$\begin{aligned} R([\mathbf{u},p,\tilde{\nu}],[\mathbf{v},q,\check{\nu}]) = \quad & R_{FEM}([\mathbf{u},p,\tilde{\nu}],[\mathbf{v},q,\check{\nu}]) && \text{(A.21)} \\ + \quad & \textstyle\sum_{\Omega_k} \int_{\Omega_k} \tau^{NS}_{SUPG} \mathbf{u}\cdot\nabla\mathbf{v}\,(\mathbf{u}\cdot\nabla\mathbf{u}) && \text{(A.22)} \\ + \quad & \textstyle\sum_{\Omega_k} \int_{\Omega_k} \tau^{SA}_{SUPG} \mathbf{u}\cdot\nabla\check{\nu}\,(\mathbf{u}\cdot\nabla\tilde{\nu}) && \text{(A.23)} \end{aligned}$$

where R_{FEM} is the standard, unstable, FEM discretization. The parameters τ^{NS}_{SUPG} and τ^{SA}_{SUPG} are parameters tuning the amount of artificial diffusivity and should not be taken too large, sacrificing precision, nor too low, sacrificing stability. A good compromise is given by:

$$\tau^{NS/SA}_{SUPG} = \frac{\xi_{NS/SA}(Re_h^{NS/SA})h_T}{2|\mathbf{u}|}, \tag{A.24}$$

where the parameters $\xi_{NS}(Re_h^{NS})$ and $\xi_{SA}(Re_h^{SA})$ are limiters, each function on the mesh Reynolds numbers such that they saturate for $Re_h \to +\infty$, preventing it to over-diffuse. Here we introduce two different parameters for the NS and SA equations individually since we will study their dependency on the final solution separately. Furthermore, this parameters must tend to a constant whenever $|\mathbf{u}| \to 0$, such that τ_{SUPG} tends to a finite value. Several expressions for those functions exist in the literature, all of them providing similar results. The one we will use here is given by:

$$\xi_{NS/SA} = \begin{cases} Re_h^{NS/SA}/3 & Re_h^{NS/SA} \leq 3 \\ 1 & Re_h^{NS/SA} > 3 \end{cases} \tag{A.25}$$

where the definition of the mesh Reynolds numbers is given by:

$$Re_h^{NS} = \frac{|\mathbf{u}|h_T}{2\mu}, \quad Re_h^{SA} = \frac{|\mathbf{u}|h_T}{2\eta} \tag{A.26}$$

where here we introduce two diffusivities, one for the NS equations μ and another one for the SA equations η. The reason for this is that, in what follows, we will study how their definitions can affect the solution, either by considering their laminar version (ν) or their turbulent versions (for example, $\mu = \nu + \nu_t$). Finally, the parameter h_T is the local element size. When the computational mesh contains high-aspect ratio elements, it is preferable, in order to avoid over-diffusion (see Mittal [2000]), to take it as the smallest length of the element. It was the case for the backward-facing step, treated in the second and third chapters of this manuscript. For the square cylinder case, since the mesh is essentially build with low aspect-ratio elements, this parameter is taken to be the its largest length. However, in order to evaluate how the SUPG works and how it converges in mesh refinement, it is interesting to modify this value to be equal to $\gamma \times h_t$, so that the amount of artificial eddy-viscosity will be decreased with a small value of γ. In what follows, we perform several tests using the models and numerical implementations described above.

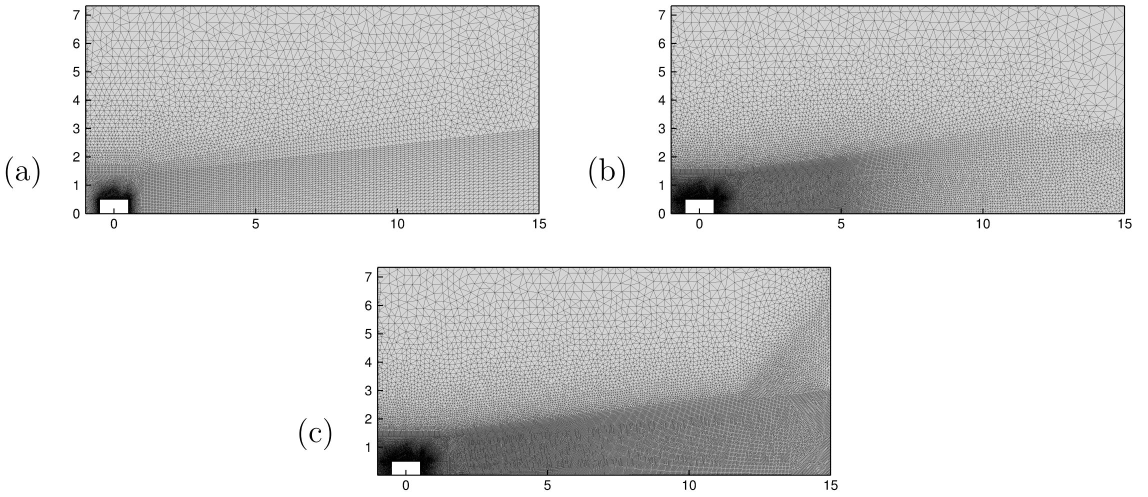

Figure A.3: Three meshes used for numerical tests. Mesh 1 (a), 2 (b), 3 (c).

3 Numerical tests and Mesh Convergence

In this section, we are going to provide some mesh-convergence on three different problems related to the RANS-SA equations, applied on the square-section cylinder, considered in the fifth chapter. The first one consists on the Newton algorithm for the convergence of the RANS-SA equations. The second one consists on the Resolvent modes, in this case the vortex shedding one. The third is the solution of the model considered in the fifth chapter for data-assimilation. All those problems will be seolved on threed different meshes, displayed on figure A.3. The first one (30000 elements) is the coarsest one, having a relatively good refinement around the cylinder. The second (75000 elements) one is more refined around the cylinder and on the wake up until $x < 5$. The third one (100000 elements) is essentially the same as the second one, however, with a finer refinement up to the outflow boundary condition. We remark that all those meshes correspond to the half-domain $y > 0$. Their use when the whole domain was needed (for example in the fifth chapter) was done by constructing a full-domain mesh constructed by uniting the original mesh with its symmetric.

3.1 RANS-SA Solution

This first problem we will consider here is the steady-flow equations, the RANS-SA ones. Those equations are solved with the Newton method, together with an increase of the Reynolds number. The strategy here is to give ourselves a few Reynolds numbers, starting with 100, for which a solution can be easilly found, and, starting from this solution, we converge a solution for $Re = 500$, then for $Re = 2000$, $Re = 5000$, $Re = 10000$ and finally $Re = 22000$.

From the point of view of the final solution, in a fist time, we investigate how the final solution is modified with the mesh and with the diffusivity taken into account in the SUPG stabilization. This diffusivity, here denoted μ for the Navier-Stokes equations and ζ for SA equation, can either be considered as the laminar one ν or the turbulent ones $\mu = \nu + \nu_t$ and $\zeta = \nu + \eta$. Those results are summarized in table A.1. We can see that the global quantities we chose to look at (namely the recirculation bubble length L, the amount of eddy-viscosity produced $||\tilde{\nu}||$ and its baricenter $\int_\Omega x\tilde{\nu}/\int_\Omega \tilde{\nu}$) do not depend much on the choice of the diffusivity. Moreover, the number of iterations needed for converging from $Re = 10000$ to $Re = 22000$ is almost independent on this choice. We can see that the choice on the mesh seems to be more important than the choice on the diffusivity. For this reason, if explicitly said otherwise, we will retain the definitions $\mu = \nu + \nu_t$ and $\zeta = \nu$.

Table A.1: RANS-SA solutions for different implementations of SUPG stabilisation. The diffusivity used in the local Reynolds number is varied between the laminar one and the turbulent one in the Navier-Stokes equation and in the Spalart-Allmaras equation separately.

Mesh (# triangles)	$Re_h = \frac{\lvert\mathbf{u}\rvert h}{2\mu}$	$Pe_h = \frac{\lvert\mathbf{u}\rvert h}{2\sigma\zeta}$	L	$\lvert\lvert\tilde{\nu}\rvert\rvert$	$\int_\Omega x\tilde{\nu}/\int_\Omega \tilde{\nu}$
1 (29384)	$\mu = \nu$	$\zeta = \nu$	3.804	0.211	8.500
	$\mu = \nu + \nu_t$	$\zeta = \nu$	3.800	0.213	8.497
	$\mu = \nu + \nu_t$	$\zeta = \nu + \eta$	3.809	0.214	8.505
3 (99922)	$\mu = \nu$	$\zeta = \nu$	3.835	0.220	8.503
	$\mu = \nu + \nu_t$	$\zeta = \nu$	3.832	0.221	8.501
	$\mu = \nu + \nu_t$	$\zeta = \nu + \eta$	3.837	0.222	8.507

We now study how the effect of the extra diffusivity given by the version of SUPG implemented as function of the mesh. To do so, we mask the mesh-size h that appears in the computation of the SUPG functions by the factors γ_h^{NS} (in NS equations) and γ_h^{SA} (in SA equation). By doing so, the SUPG functions will "see" the element sizes smaller than they are, minimizing their effect and thus the numerical diffusion. We can see from table A.2 that, by taking smaller values of those γ parameters, the solution converges towards a longer bubble. It is interesting to note that even the coarser mesh manages to recover the same solution for the finer one, requiring only a sufficient small value for the γ parameters. Moreover, the number of iterations that the Newton method takes, say for the mesh 1 (coarser one), is similar to the one for the mesh 3 (finer one). This shows that, even if we decrease the numerical diffusivity, the convergence is not altered. We note that this is not true for the mesh 2 (intermediate one), where the discretization towards the outflow border is very coarse. This is related to some oscillations of the variable $\tilde{\nu}$ at the outflow region where it transitions from the farfield value to the turbulent one inside the bubble. Apparently, this is caused by a negative $\tilde{\nu}$ that persists even with the corrections mentioned in previous paragraphs. This effect is not fully understood and will be avoided by considering meshes sufficiently refined at that region and more refined towards the outflow.

3.2 Resolvent mode analysis

The Resolvent analysis (or the Stability analysis) is a much more different test case from the RANS-SA solution, since it has a connection with the Jacobian matrix and not with the residual itself and the SUPG stabilization may reflect differently for this problem. To assess this difference, we consider only the vortex-shedding mode, corresponding to the Resolvent mode on the RANS-SA solution, at frozen $\tilde{\nu}$, at $\omega = 0.83$ and investigate how

Table A.2: RANS-SA solutions for different implementations of SUPG stabilisation. The mesh element size is masked with a parameter γ_h^{NS} and γ_h^{SA} in the SUPG, tuning the diffusivity of the method.

Mesh (# triangles)	γ_h^{NS}	γ_h^{SA}	L	$\lvert\lvert\tilde{\nu}\rvert\rvert$	$\int_\Omega x\tilde{\nu}/\int_\Omega \tilde{\nu}$	# iterations
1 (29384)	1	1	3.800	0.213	8.497	5
	0.2	1	3.876	0.225	8.496	8
	0.05	1	3.891	0.227	8.495	8
	0.05	0.05	3.898	0.231	8.516	12
2 (74563)	1	1	3.832	0.221	8.502	4
	0.2	1	3.886	0.228	8.505	8
	0.05	1	3.896	0.230	8.505	19
	0.05	0.05	3.900	0.231	8.514	10
3 (99922)	1	1	3.832	0.221	8.501	7
	0.2	1	3.886	0.228	8.503	7
	0.05	1	3.896	0.230	8.504	7
	0.05	0.05	3.900	0.231	8.513	10

it converges. This results are summarized in table A.3, where the spatial distribution of the energy of the mode is evaluated. We can see that whenever the mesh is coarse and the SUPG used is the full version $\gamma_h^{NS} = 1$, the energy tends to be more localized at the vicinity of the cylinder where if the mesh is refined and/or the parameter γ_h^{NS} is decreased, the energy tends to be more distributed at the outflow region of the domain. In particular, even for the coarsest mesh, we manage already to converge, if the parameter γ_h^{NS} is sufficiently small.

Table A.3: Mesh convergence for SVD modes for some SUPG implementations and mesh refinements. The energy distribution of the Resolvent mode (given by $\kappa_{\Omega_i} = \int_{\Omega_i} |\hat{\mathbf{u}}|^2)$ is studied in function of the SUPG implementation, for $\Omega_1 = (-0.75, 0.75) \times (-0.75, 0.75)$, $\Omega_2 = (-0.75, 5) \times (-0.75, 0.75) \backslash \Omega_1$ and $\Omega_3 = \Omega \backslash (\Omega_1 \cup \Omega_2)$. Results presented for $\zeta = \nu + \nu_t$

	$\gamma_h^{NS} = 1/0.2/0.05$		
Mesh (# triangles)	κ_{Ω_1}	κ_{Ω_2}	κ_{Ω_3}
1 (29384)	2.8/2.0/1.9 %	33.0/27.2/26.8 %	64.2/70.8/71.2 %
2 (74563)	1.6/1.6/1.5 %	26.5/24.4/24.2 %	71.8/74.0/74.3 %
3 (99922)	1.8/1.7/1.7 %	26.2/24.8/24.6 %	72.0/73.5/73.7 %

3.3 Data-Assimilation model resolution

The final test for our model would be the resolution of the model proposed in the last chapter for data-assimilation, where not only the mean-flow equation is solved with the Resolvent vortex-shedding mode. This is thus some sort of combination between the two last mesh-convergence studies. To perform this test, we solve the model for several values of $|A|$ and evaluate the model's solution in function of it. This is performed for the three meshes and for three parameters $\gamma_h^{NS} = 0.05/0.2/1$. Those results are plotted in figure

A.4. We can see from those results that indeed the model's solution converges either with a mesh refinement or with a decrease of γ_h^{NS}, as observed before. Furthermore, the mesh chosen for this kind of model, presented in the fourth chapter, is the coarser one, since no observed quantity seemed to depend on the discretization.

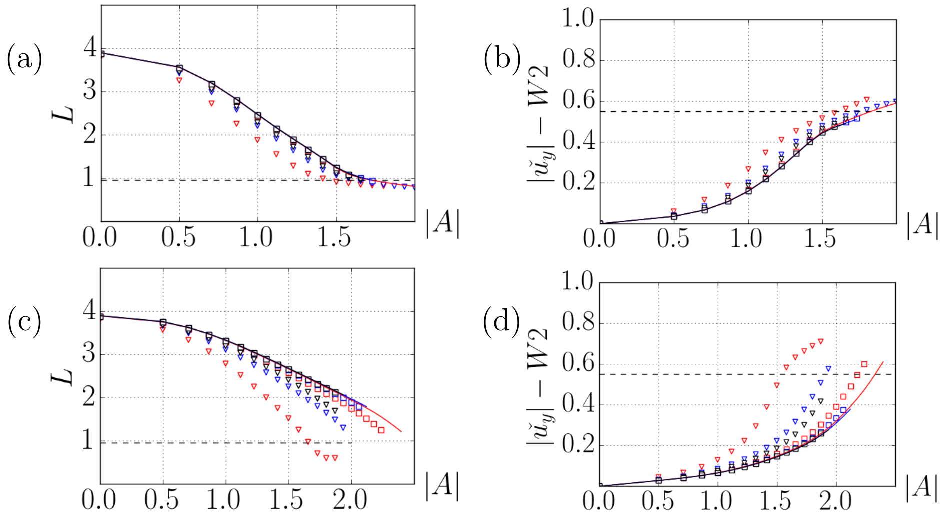

Figure A.4: Convergence of the model for several meshes and SUPG implementations ((a,b) frozen ν_t and (c,d) variable ν_t). The triangles correspond to $\gamma_h^{NS} = 1$, the squares to $\gamma_h^{NS} = 0.2$ and the solid lines to $\gamma_h^{NS} = 0.05$. The red curves correspond to mesh 1, the blue one to mesh 2 and the black curve to mesh 3.

Appendix B

BFGS Algorithm and its use with a Finite-Element Solver

1 BFGS and Line search Methods

With the information of the gradient $\nabla_{\tilde{\mathbf{f}}}\mathcal{J}$ of our cost functional $\mathcal{J}$ with respect to the vector of control variables (here denoted $\tilde{\mathbf{f}}$), several optimization algorithms can be used to find the optimal conditions. They all rely on the building of a sequence of forcing terms $\tilde{\mathbf{f}}^n$ such that $\tilde{\mathbf{f}}^n \to \tilde{\mathbf{f}}^*$, where $\tilde{\mathbf{f}}^*$ is the optimal forcing term, for which, ideally, $\nabla_{\tilde{\mathbf{f}}}\mathcal{J}|_{\tilde{\mathbf{f}}^*} = \mathbf{0}$. A simple gradient-based method (employed, for example, in Foures et al. [2014], Symon et al. [2017], etc) creates this sequence such that $\tilde{\mathbf{f}}^{n+1} = \tilde{\mathbf{f}}^n - \varepsilon_n \nabla_{\tilde{\mathbf{f}}}\mathcal{J}$, where the parameter ε_n is determined, for example, with a line-search algorithm. However, this is a first-order method since it relies solely on the information of the gradient. A well known example of second-order method is the Newton method for optimization, that constructs this sequence of forcing terms such that $\tilde{\mathbf{f}}^{n+1} - \tilde{\mathbf{f}}^n = -\varepsilon_n H_n^{-1} \nabla_{\mathbf{f}}\mathcal{J}|_{\mathbf{f}_n}$, where $H_n = \nabla_{\mathbf{f}}\nabla_{\mathbf{f}}\mathcal{J}$ is the Hessian matrix at iteration step n. This method is generally not so easy to apply since one does not, in general, have the Hessian matrix. However, some methods try to approximate it, retaining a good memory requirement (see Fletcher and Powel [1963], Broyden [1983], Fletcher [1970], Nocedal and Wright [2006], Bonnans et al. [2006]). In this context, we propose to use the limited-memory Broyden–Fletcher–Goldfarb–Shanno (l-BFGS) method. This method approximates the Hessian matrix from previous iterations computations in the following way (with $H_0 = I$):

$$H_{n+1} = H_n + \frac{y_n y_n^T}{y_n^T s_n} - \frac{H_n s_n s_n^T H_n}{s_n^T H_n s_n} \tag{B.1}$$

where $y_n = \nabla_{\mathbf{f}}\mathcal{J}|_{\mathbf{f}_{n+1}} - \nabla_{\mathbf{f}}\mathcal{J}|_{\mathbf{f}_n}$ is the difference of the gradient between two successive iterations and $s_n = \mathbf{f}_{n+1} - \mathbf{f}_n = \varepsilon_n \mathbf{p}_n$ ($\mathbf{p}_n$ being the descent direction, $\mathbf{p}_n = H_n^{-1}$) the difference of control parameters. Since at every successive iteration this matrix is incremented by two sets of line/column, we can use the Woodbury identity to invert it:

$$H_{n+1}^{-1} = \left(I - \frac{s_n y_n^T}{y_n^T s_n}\right) H_n^{-1} \left(I - \frac{y_n s_n^T}{y_n^T s_n}\right) + \frac{s_n s_n^T}{y_n^T s_n} \tag{B.2}$$

Where, in order to respect some memory requirement, the storage of previous vectors s_n, y_n is limited to a certain number k of previous iterations. Together with the estimation of the Hessian matrix, the method is coupled with a line search algorithm. This algorithm finds the parameter ε_n such that the following (Wolfe) conditions are satisfied:

$$\begin{aligned} \mathcal{J}(\tilde{\mathbf{f}}_n + \varepsilon_n \mathbf{p}_n) \leq J(\tilde{\mathbf{f}}_n) + c_1 \varepsilon_n \mathbf{p}_n^T \nabla_{\tilde{\mathbf{f}}} J|_{\tilde{\mathbf{f}}_n} \\ -\mathbf{p}_n^T \nabla_{\tilde{\mathbf{f}}} \mathcal{J}|_{\tilde{\mathbf{f}}_n + \varepsilon_n \mathbf{p}_n} \leq -c_2 \shortmid_n^T \nabla_{\tilde{\mathbf{f}}} \mathcal{J}|_{\tilde{\mathbf{f}}_n} \end{aligned} \tag{B.3}$$

for $0 < c_1 < c_2 < 1$. This roughly summarizes how the BFGS algorithm works. In the next section, we will discuss how to use it with a Partial Differential Equations (PDE) solver, in particular with a Finite-Element solver.

2 BFGS with Finite-Element Solver

2.1 A toy model for EDP solver - Laplace equation

In what follows, we will be interested in applying the optimization procedure on a very simple PDE, namely the Laplace equation (with homogeneous Cauchy boundary conditions):

$$\begin{cases} -\Delta u = f & \mathbf{x} \in \Omega \\ u = 0 & \mathbf{x} \in \partial\Omega \end{cases}, \quad \Omega = (0,1) \times (0,1) \tag{B.4}$$

We remark however that the presentation of the discretization of those equation, together with the optimization procedure are general and valid as well for nonlinear model, such as the ones considered in this manuscript. The discrete version of the continuous problem (2.1) is given here by the Finite-Element Method (FEM). This method arises from the variational formulation associated with the strong equations (2.1). To derive it, we multiply it by an arbitrary test function, integrate by parts an apply the boundary conditions, leading to:

$$a(u,v) = \int_\Omega \nabla u \cdot \nabla v \; d\Omega = \int_\Omega f v \; d\Omega = l(v), \quad \forall v \tag{B.5}$$

In FEM, we give ourselves a set of base functions $\{\phi_i\}$ that will be used to span our solution and the forcing term. We can say that:

$$u(\mathbf{x}) = \sum_i U_i \phi_i(\mathbf{x}), \quad f(\mathbf{x}) = \sum_i F_i \phi_i(\mathbf{x}) \tag{B.6}$$

where U_i, F_i are the coefficients that represent numerically the function $u(\mathbf{x})$ and $f(\mathbf{x})$. Since the variational formulation has to be true for any v, we may choose $v(\mathbf{x}) = \phi_j(\mathbf{x}), \forall j$. Each such choice of test function will lead to a line of the linear system:

$$\mathbb{K}U = \mathbb{M}F \tag{B.7}$$

where $U = \{U_i\}$, $F = \{F_i\}$, $\mathbb{K}_{i,j} = a(\phi_i, \phi_j)$ is the so-called rigidity matrix and $\mathbb{M}_{i,j} = \int_\Omega \phi_i \phi_j d\Omega$ is the so-called mass matrix. We remark that the inner product associated with the continuous problem is given in its discrete version by the mass matrix:

$$(u,v)_\Omega = \int_\Omega uv d\Omega = U^T \mathbb{M} V \tag{B.8}$$

Once the discretization of the continuous equations are set, we proceed to the associated optimization problem.

2.2 Optimization Problem

The optimization problem will consist of finding the optimal forcing term f such that the cost functional:

$$\mathcal{J} = \frac{1}{2}\int_\Omega (u - u_r)^2 d\Omega \tag{B.9}$$

is minimal, where the relation between u and f is given by the equation (2.1). Here the reference state u_r will consist on the solution of this Laplace equation imposing $f_r = \tanh(10(x - 0.5)) * \sin(\pi x) * \sin(\pi y)$, an arbitrary function. This problem can be viewed in two different manners: a continuous one and a discrete one.

Continuous Formalism: The continuous version is essentially the one presented in the main body of this manuscript in chapters 2 and 3 and consists of the Lagrangian multiplier formalism posed for the continuous equations, where the adjoint system and the gradient are set. The Lagrangian of this system is given by:

$$\mathcal{L} = \mathcal{J} + (u^\dagger, -\Delta u - f)_\Omega \tag{B.10}$$

By deriving this functional with respect to the state, adjoint and forcing term, we have:

$$\begin{cases} -\Delta u^\dagger = -(u - u_r) & \mathbf{x} \in \Omega \\ u^\dagger = 0 & \mathbf{x} \in \partial\Omega \end{cases}, \quad \nabla_f \mathcal{J} = -u^\dagger \tag{B.11}$$

The discrete version of those equations, necessary for numerical implementation, is given by:

$$\mathbb{K}U^\dagger = -\mathbb{M}(U - U_r), \quad \mathbf{G}_c = -U^\dagger \tag{B.12}$$

where $\mathbf{G}_c$ is the discrete gradient in this framework.

Discrete Formalism: The discrete formalism arises when we treat the optimization problem as the minimization of the cost functional:

$$\mathcal{J}(U) = \frac{1}{2}(U - U_r)^T \mathbb{M}(U - U_r) \tag{B.13}$$

whose inputs are the coordinates of the vector U and its relation with F is given by the discrete system $\mathbb{K}U = \mathbb{M}F$. If we apply again the Lagrangian multiplier formalism, given by the Lagrangian:

$$\mathcal{L} = \mathcal{J} + (U^\dagger)^T(\mathbb{K}U - \mathbb{M}F) \tag{B.14}$$

If we regard this functional as being dependent on each component of the state U_i, the adjoint $U_i^\dagger$ and the forcing term F_i, we can derive it with respect to each of those components, leading to:

$$\mathbb{K}U^\dagger = -\mathbb{M}(U - U_r), \quad \mathbf{G}_d = -\mathbb{M}U^\dagger = \mathbb{M}\mathbf{G}_c \tag{B.15}$$

We notice that the discrete gradient $\mathbf{G}_d$ differs from the continuous one by the factor of the mass-matrix. This can be understood by noticing that the discrete gradient should quantify the derivative of the cost functional with respect to the coefficients of the forcing vector F, whereas the continuous one informs how the cost functional varies with respect to an arbitrary variation of the forcing term f. Indeed, if we give ourselves a perturbation in the forcing term $\delta f = \sum_i \delta F_i \phi_i$, a variation of the cost functional $\delta\mathcal{J}$ will be given by:

$$\delta\mathcal{J} = \mathbf{G}_c^T \mathbb{M}\delta F = \mathbf{G}_d^T \delta F = (\nabla_f \mathcal{J}, \delta f)_\Omega \tag{B.16}$$

In what follows, we will discuss the implications of the choise between those approaches on the numerical results.

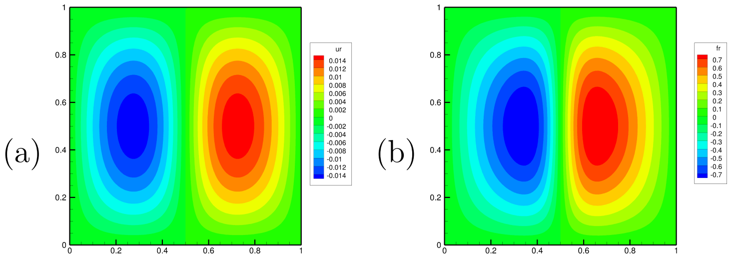

Figure B.1: Reference values for u_r (a) and f_r (b).

2.3 Optimization Results

The goal of this section is to, through the optimization procedure, find the pair (u_r, f_r), displayed in figure B.1. At first, we will naively apply the continuous and discrete formalism by providing BFGS either $(\mathcal{J}, \mathbf{G}_c)$ or $(\mathcal{J}, \mathbf{G}_d)$.

Continuous Formalism: By applying this formalism we were not able to converge to any solution. This was the case not only in this simple Laplace equation model, but also in any other PDE-constrained optimization problem considered in this manuscript. In fact, the algorithm was not able to exit the very first global iteration, even if the cost functional clearly decreased in the line-search procedure. Typically, the line-search method kept decreasing the parameter ε until the point it became small enough for the optimization procedure to stop. We are lead to believe that the main reason for that is a norm-inconsistency between our PDE-based model and the internal BFGS routines. Indeed, as mentioned in section 1, we can see that all the norms and inner products performed by the BFGS algorithm are based on the euclidean norm $U^T V = \sum_i U_i V_i$, which does not make any physical or numerical meaning, if we regard U, V as finite-element representations of their original functions. For this reason, we find necessary to test the discrete formalism.

Discrete Formalism: Contrarily to the continuous formalism, the discrete one always produces an optimal condition. The optimal solutions u^* (figure B.2 (b,e)) correspond very well to the reference one. However, we can see that, in function of the mesh, typically if the mesh is structured or not, the forcing term may become very noisy. To understand this, we need to take a closer look to how the gradient information $\mathbf{G}_d$, provided by our FEM solver, will be re-fed to it in the form of a new forcing term F. Indeed, for the sake of clarity, if we place ourselves at the first global iteration, for which $F_1 - F_0 = \varepsilon_0 \mathbf{G}_{d,0}$, we can see that the solution of the next iteration will be given by the system:

$$\mathbb{K}U_1 = \mathbb{M}(F_0 + \varepsilon_0 \mathbf{G}_{d,0}) = \mathbb{M}F_0 - \varepsilon_0 \mathbb{M}\mathbb{M}U_0^\dagger \tag{B.17}$$

we can see that on the right-hand-side of this equation we have a term with a double mass-matrix product, acting on a FEM representation of the adjoint function $U_0^\dagger$. This procedure produces a forcing term that is inconsistent with the FEM. If the mesh we consider is somewhat structured and uniform (as it is the case of the mesh 1), this inconsistency may produce some negligible effects, where the mass-matrix itself will be very well structured and will not establish locally any preferential direction. However, if the mesh is unstructured (mesh 2), this effect is quite visible. Furthermore, in those cases, the number of iterations for the convergence of the optimization algorithm may increase drastically. Indeed, from figure B.3 we can see that, in order for the algorithm to correct

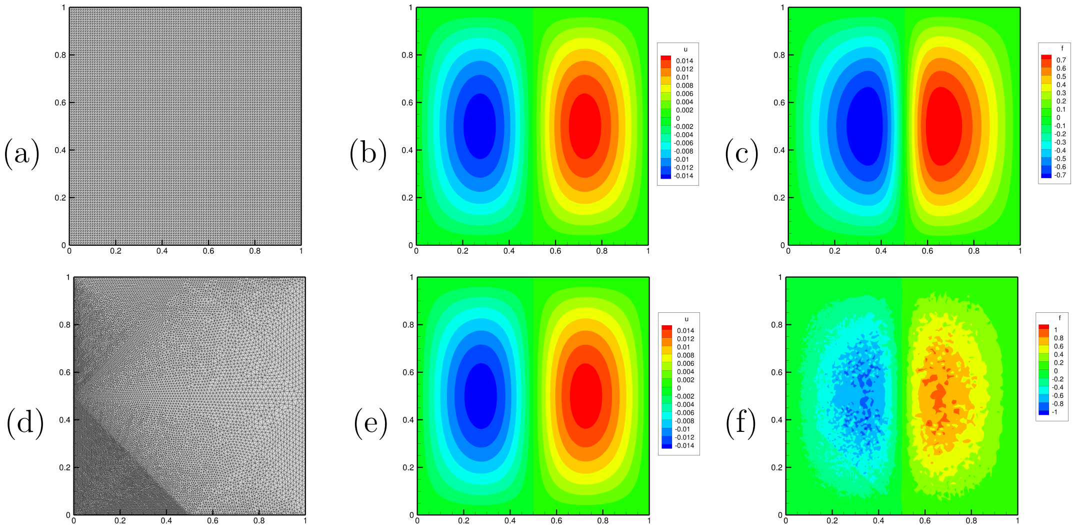

Figure B.2: Discrete Formalism Optimization: mesh (a,d), optimal solution u^* (b,e) and optimal forcing term f^*.

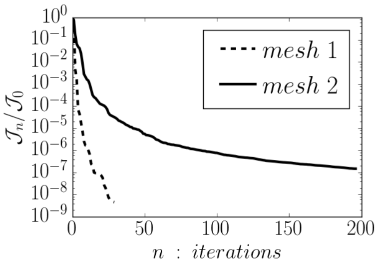

Figure B.3: Discrete Formalism Optimization: Convergence of the cost functional $\mathcal{J}_n/\mathcal{J}_0$.

those effects of the forcing term (that may extend to the solution), it takes generally 200 iterations for a level of convergence of 10^{-7}, while for the mesh 1, it takes 10-20 iterations. This poses a real problem for the applications presented in this manuscript since we cannot afford to have structured and uniform meshes.

Proposed Solution: Cholesky Decomposition It is interesting to note that, at the same time where the Discrete Formalism produced the best results so far, it produces some inconsistencies in the FEM formulation. On the other hand, the Continuous formalism did not produce any result, but, if we look closely, we find that the inconsistency described does not take place, since the gradient definition does not contain the mass-matrix. To make this Continuous Formalism applicable on a BFGS algorithm, we have to perform a change the variables, on which the Euclidean norm (performed inside BFGS) will make sense both physically and numerically. Generally, this change of variables reads $\tilde{F} = \mathbb{L}^T F$. The Euclidean inner-product of two arbitrary vector is given the by:

$$(\tilde{F}_1)^T \tilde{F}_2 = F_1^T \mathbb{L}\mathbb{L}^T F_2 = F_1^T \mathbb{M} F_2 = (f_1, f_2)_\Omega \tag{B.18}$$

where the last two terms correspond to the imposition that this quantity represents the energy inner product of the two FEM functions f_1, f_2. We can see that the matrix $\mathbb{L}$ has to be given by the Cholesky decomposition of the mass-matrix: $\mathbb{M} = \mathbb{L}\mathbb{L}^T$. For this

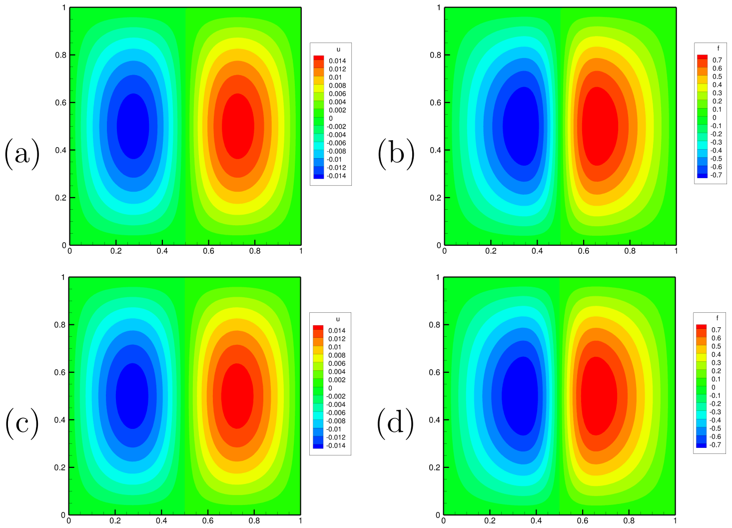

Figure B.4: Cholesky decomposition change of variables: final solution (a,c) and forcing term (b,d) for meshes 1 (a,b) and 2 (c,d).

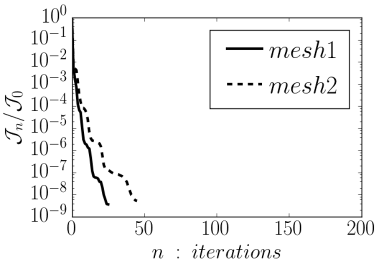

Figure B.5: Discrete Formalism Optimization: Convergence of the cost functional $\mathcal{J}_n/\mathcal{J}_0$.

reason, the continuous gradient, written now in terms of this new variable becomes:

$$\tilde{\mathbf{G}}_c = \nabla_{\tilde{F}}\mathcal{J} = \mathbb{L}^T\mathbf{G}_c \tag{B.19}$$

We note that, with this changes of variables, not only all the norms performed within BFGS will correspond to the physical (energy norm) ones, but also the FEM formulation will be consistent, since the forcing term on the system 2.3 will be:

$$\mathbb{K}U_1 = \mathbb{L}(\tilde{F}_0 + \varepsilon_0\tilde{\mathbf{G}}_{d,0}) = \mathbb{M}F_0 - \varepsilon_0\mathbb{M}U_0^{\dagger} \tag{B.20}$$

Indeed, with this kind of approach, we obtain better results (figure B.4). The are mesh-independent and manage to reconstruct both the forcing term and the solution itself. Furthermore, the convergence of the cost functional B.5 is also (almost) mesh-independent.

3 Remarks on Navies-Stokes equations: Incompressibility of $\tilde{\mathbf{f}}_{\mathbf{x}}$

We notice as well that, for the model 1, where the tuning parameter is the volume force $\tilde{\mathbf{f}}_{\mathbf{x}}$, the incompressibility of this forcing term, discussed in the second chapter, is not obvious since the Hessian matrix may modify the spatial structure of the gradients, leading to an increment in $\tilde{\mathbf{f}}_{\mathbf{x}}$ that is not solenoidal. However, this is not the case as we will see here. To be able to prove this, we need to show that the application of H_{n+1}^{-1} onto any divergence-free vector z (here, the gradient) is also divergence-free. To do so, we use induction, where we suppose it is true at some iteration n, i.e., $\mathbb{D}H_n^{-1}z = 0$ (true for $n = 0$ since $H_0 = I$), where $\mathbb{D}$ is the discrete-version of the divergence operator. If we suppose that $\mathbb{D}s_n = 0$, we have that:

$$\begin{aligned}\mathbb{D}H_{n+1}^{-1}z &= \mathbb{D}\left(I - \frac{s_n y_n^T}{y_n^T s_n}\right) H_n^{-1}\left(I - \frac{y_n s_n^T}{y_n^T s_n}\right) z + \mathbb{D}\frac{s_n s_n^T}{y_n^T s_n} z \\ &= \mathbb{D}H_n^{-1}\left(I - \frac{y_n s_n^T}{y_n^T s_n}\right) z \\ &= \mathbb{D}H_n^{-1}z - \mathbb{D}H_n^{-1}\frac{y_n s_n^T}{y_n^T s_n} z = 0\end{aligned} \tag{B.21}$$

which is zero since both terms on the last equation are applications of $\mathbb{D}H_n^{-1}$ onto divergence-free vectors.

We notice that the fact we used a linear transformation coming from the Cholesky decomposition does not invalidate this argument since the vector z can be considered as being already transformed by it and the matrix $\mathbb{D}$ modified to take into account for the Cholesky matrix $\mathbb{L}$ incorporated in z.

4 Adjoint Discretization in FEM

In this section, we briefly discuss the adjoint problem for a (sufficiently) general nonlinear Partial Differential Equation (PDE) and its further discretization in FEM. The main goal here is to prove that the transpose of the Jacobian matrix provides indeed possible discretization of the adjoint problem. For our purpose here, a nonlinear scalar advection-diffusion equation under the form:

$$\mathbf{a} \cdot \nabla u - \nabla \cdot (\nu(u)\nabla u) = f, \mathbf{x} \in \Omega \tag{B.22}$$

is sufficient. Here f is a forcing term similar to those used as tuning parameters in the body of this work. The boundary conditions associated with this problem are typically of Dirichlet or Neumann types:

$$\Gamma_D : u = g(\mathbf{x}), \quad \Gamma_N : \nu(u)\mathbf{n} \cdot \nabla u = 0 \tag{B.23}$$

The Finite-Element discretization of such equation considers its weak-form, given by:

$$r(u, v) = \int_\Omega (\mathbf{a} \cdot \nabla u v + \nu(u)\nabla u \cdot \nabla v - f v)\, d\Omega = 0, \quad \forall v \tag{B.24}$$

This (semi-linear) form provides a nonlinear residual that needs to be zero for all test functions v so that u is a weak solution of our problem. In practice, it is achieved with a Newton method, which converges towards a solution by successive inversions of the Jacobian of the system, which corresponds to the linearization of this residual:

$$\begin{aligned}a(\delta u, v) &= \int_\Omega (\mathbf{a} \cdot \nabla(\delta u)v + \nu(u)\nabla(\delta u) \cdot \nabla v + \partial_u \nu(u)(\delta u)\nabla u \cdot \nabla v)\, d\Omega \\ &= -r(u, v), \forall v\end{aligned} \tag{B.25}$$

where $a(\delta u, v)$ is a bilinear form that leads to the Jacobian matrix upon discretization. Here, for simplicity, we suppose that the solution guess u already satisfies the Dirichlet boundary condition, meaning that the correction δu has to satisfy $\delta u = 0$ at Γ_D. This imposition of the Dirichlet boundary conditions is done through, for example, elimination of columns and lines of the complete matrix corresponding to the degrees of freedom on which we impose $u = 0$. Another way, which is default in FreeFEM++ and used throughout this work, is the penalization of those degrees of freedom in the diagonal of the matrix, freezing them.

We proceed now to the adjoint problem. This problem appears whenever we desire to optimize some functional $\mathcal{J}$ under the PDE constraint. We consider the Lagrangian:

$$\mathcal{L} = \mathcal{J} + \left(u^\dagger, \mathbf{a} \cdot \nabla u - \nabla \cdot (\nu(u)\nabla u) - f\right)_\Omega \tag{B.26}$$

The variation of this Lagrangian with respect to u, and the further integration by parts of the viscous terms lead to:

$$\begin{aligned}\left(\frac{\partial \mathcal{L}}{\partial u}, \delta u\right)_\Omega &= \left(\frac{\partial \mathcal{J}}{\partial u}, \delta u\right)_\Omega + \left(u^\dagger, \mathbf{a} \cdot \nabla \delta u - \nabla \cdot (\nu(u)\nabla(\delta u) + \partial_u \nu(u)(\delta u)\nabla u)\right)_\Omega \\ &= \left(\frac{\partial \mathcal{J}}{\partial u}, \delta u\right)_\Omega + a^\dagger(u^\dagger, \delta u) - \left(u^\dagger, \nu(u)\mathbf{n} \cdot \nabla(\delta u) + \partial_u \nu(u)(\delta u)\mathbf{n} \cdot \nabla u\right)_{\partial\Omega}\end{aligned} \tag{B.27}$$

where, at this point we can identify the bilinear form $a^\dagger(u^\dagger, \delta u) = a(\delta u, u^\dagger)$ in the formulation of the adjoint equations. Since this form is the same as the one leading to the Jacobian, only with changed arguments, we can see that the bilinear form $a^\dagger(u^\dagger, \delta u)$ will be represented numerically by the transpose of the Jacobian matrix. However, to be able to fully justify the use of the transposed matrix as the adjoint one, we need to prove two things: (a) that the last boundary term appearing in (B.27) is zero and (b) that the adjoint variable should satisfy a Dirichlet boundary condition at Γ_D. This last point presents itself numerically in the imposition of the boundary condition by line/columns elimination or penalization, and we need to ensure that the positions of the adjoint matrix where this procedure was done are the same as the ones in the Jacobian matrix. We remark that, if an homogeneous Dirichlet condition is verified at Γ_D, the aforementioned boundary term is only non null at Γ_N, where, by application of the linearized Neumann boundary conditions for δu, i.e., $\nu(u)\mathbf{n} \cdot \nabla(\delta u) + \partial_u \nu(u)(\delta u)\mathbf{n} \cdot \nabla u = 0$, we prove it is also null. This means that both points (a) and (b) can be assured by showing that $u^\dagger = 0$ at Γ_D. To do so, we continue the integration by parts in the adjoint derivation:

$$\begin{aligned}\left(\frac{\partial \mathcal{L}}{\partial u}, \delta u\right)_\Omega &= \left(-\nabla \cdot (u^\dagger \mathbf{a}) + (\partial_u \nu(u))\nabla u^\dagger \cdot \nabla u - \nabla \cdot (\nu(u)\nabla u^\dagger) + \partial \mathcal{J}/\partial u, \delta u\right)_\Omega \\ &\quad - \left(u^\dagger, \nu(u)\mathbf{n} \cdot \nabla(\delta u) + (\partial_u \nu)\mathbf{n} \cdot \nabla u\right)_{\partial\Omega} \\ &\quad + \left(\nu \mathbf{n} \cdot \nabla u^\dagger + \mathbf{a} \cdot \mathbf{n} u^\dagger, \delta u\right)_{\partial\Omega}\end{aligned} \tag{B.28}$$

This determines the volume (strong) equation for the adjoint problem as:

$$-\nabla \cdot (u^\dagger \mathbf{a}) + (\partial_u \nu(u))\nabla u^\dagger \cdot \nabla u - \nabla \cdot (\nu(u)\nabla u^\dagger) = -\partial \mathcal{J}/\partial u \tag{B.29}$$

and the boundary conditions. Indeed, at Γ_N, by applying the linearized Neumann boundary condition, we have that, to the boundary terms to be zero, $\mathbf{n} \cdot (\mathbf{a} u^\dagger + \nu(u)\nabla u^\dagger) = 0$.

At Γ_D, since $\delta u = 0$, to the remaining boundary terms to be zero, we have that $u^\dagger + 0$. We have proven this way that the transpose of the Jacobian matrix provides indeed a FEM discretization of the adjoint equations.

Bibliography

Allmaras, S. R., Johnson, F. T., and Spalart, P. R. (2012). Modifications and clarifications for the implementation of the spalart-allmaras turbulence model. *Seventh International Conference on Computational Fluid Dynamics (ICCFD7) and Big Island and Hawaii and 9-13 July.*

Arbabi, H. and Mezić, I. (2017). Study of dynamics in post-transient flows using koopman mode decomposition. *Phys. Rev. Fluids*, 2(12):124402.

Bao, Y., Zhou, D., Huang, C., Wu, Q., and Chen, X.-q. (2011). Numerical prediction of aerodynamic characteristics of prismatic cylinder by finite element method with spalart–allmaras turbulence model. *Comput. Struc.*, 89(3-4):325–338.

Barkley, D. (2006). Linear analysis of the cylinder wake mean flow. *Europhys. Lett.*, 75:750–756.

Barkley, D., Gomes, M. G. M., and Henderson, R. D. (2002). Three-dimensional instability in flow over a backward-facing step. *J. Fluid Mech.*, 473:167–190.

Beneddine, S., Sipp, D., Arnault, A., Dandois, J., and Lesshafft, L. (2016). Conditions for validity of mean flow stability analysis. *J. Fluid Mech.*, 798:485–504.

Beneddine, S., Yegavian, R., Sipp, D., and Leclaire, B. (2017). Unsteady flow dynamics reconstruction from mean flow and point sensors: an experimental study. *J. Fluid Mech.*, 824:174–201.

Bewley, T. R. and Protas, B. (2004). Skin friction and pressure: the "footprints" of turbulence. *Physica D*, 196(1-2):28–44.

Blackburn, H. M., Barkley, D., and Sherwin, S. J. (2008). Convective instability and transient growth in flow over a backward-facing step. *J. Fluid Mech.*, 603:271–304.

Bonnans, J., Gilbert, J. C., Lemaréchal, C., and Sagastizábal, C. A. (2006). *Numerical optimization: theoretical and practical aspects.* Springer Science & Business Media.

Bosch, G. and Rodi, W. (1998). Simulation of vortex shedding past a square cylinder with different turbulence models. *Int. J. Numer. Meth. Fl.*, 28(4):601–616.

Brandt, L., Sipp, D., Pralits, J. O., and Marquet, O. (2011). Effect of base-flow variation in noise amplifiers: the flat-plate boundary layer. *J. Fluid Mech.*, 687:503–528.

Brooks, A. and Hughes, T. (1982). Streamline upwind/petrov-galerkin formulations for convection dominated flows with particular emphasis on the incompressible navier-stokes equations. *Comput. Method. Appl. M.*, 32(1-3):199–259.

Broyden, C. G. (1983). Quasi-newton methods and their application to function minimization. *Maths. Comput.*, 21:368.

Brun, C., Aubrun, S., Goossens, T., and Ravier, P. (2008). Coherent structures and their frequency signature in the separated shear layer on the sides of a square cylinder. *Flow Turbul. Combust.*, 81(1-2):97–114.

Camarri, S., Trip, R., and Fransson, J. H. (2017). Investigation of passive control of the wake past a thick plate by stability and sensitivity analysis of experimental data. *J. Fluid Mech.*, 828:753–778.

Chomaz, J.-M. (2005). Global instabilities in spatially developing flows: Non-normality and nonlinearity. *Annu. Rev. Fluid Mech.*, 37:357–392.

Colburn, C., Cessna, J., and Bewley, T. (2011). State estimation in wall-bounded flow systems. part 3. the ensemble kalman filter. *J. Fluid Mech.*, 682:289–303.

Cossu, C., Pujals, G., and Depardon, S. (2009). Optimal transient growth and very large–scale structures in turbulent boundary layers. *J. Fluid Mech.*, 619:79–94.

Crivellini, A., D'Alessandro, V., and Bassi, F. (2013). A spalart–allmaras turbulence model implementation in a discontinuous galerkin solver for incompressible flow. *J. Comp. Phys.*, 241:388–415.

Crouch, J., Garbaruk, A., Magidov, D., and Travin, A. (2009). Origin of transonic buffet on aerofoils. *J. Fluid Mech.*, 628:357–369.

D'Adamo, J., Papadakis, N., Memin, E., and Artana, G. (2007). Variational assimilation of pod low-order dynamical systems. *J. Turbul.*, 8:1–22.

Dandois, J., Garnier, E., and Sagaut, P. (2007). Numerical simulation of active separation control by a synthetic jet. *J. Fluid Mech.*, 574:25–58.

Dandois, J., Mary, I., and Brion, V. (2018). Large-eddy simulation of laminar transonic buffet. *J. Fluid Mech.*, 850:156–178.

Deck, S. (2005). Zonal-detached-eddy simulation of the flow around a high-lift configuration. *AIAA J.*, 43 (11):2372–2384.

Deconinck, B. and Kutz, J. N. (2006). Computing spectra of linear operators using the floquet–fourier–hill method. *J. Comp. Phys.*, 219(1):296–321.

Duraisamy, K., Iaccarino, G., and Xiao, H. (2019). Turbulence modeling in the age of data. *Annu. Rev. Fluid Mech.*, 51:357–377.

Ehrenstein, U. and Gallaire, F. (2005). On two-dimensional temporal modes in spatially evolving open flows: the flat-plate boundary layer. *J. Fluid Mech.*, 536:209–218.

Ehrenstein, U. and Gallaire, F. (2008). Two-dimensional global low-frequency oscillations in a separating boundary-layer flow. *J. Fluid Mech.*, 614:315–327.

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5):10143–10162.

Evensen, G. (2009). *Data Assimilation: The Ensemble Kalman Filter.* Springer-Verlag, 2nd edition.

Fletcher, R. (1970). A new approach to variable metric algorithms. *Comp. J.*, 13:317.

Fletcher, R. and Powel, M. J. D. (1963). A rapidly convergent descent methods for minimization. *Comp. J.*, 6:163.

Foures, D. P. G., Dovetta, N., Sipp, D., and Schmid, P. J. (2014). A data-assimilation method for reynolds-averaged navierstokes-driven mean ow reconstruction. *J. Fluid Mech.*, 759:404–431.

Franca, L. P. and Frey, S. L. (1992). Stabilized finite element methods: Ii. the incompressible navier-stokes equations. *Comput. Method. Appl. M.*, 99(2-3):209–233.

Franca, L. P., Frey, S. L., and Hughes, T. J. (1992). Stabilized finite element methods: I. application to the advective-diffusive model. *Comput. Method. Appl. M.*, 95(2):253–276.

Gronskis, A., Heitz, D., and Memin, E. (2013). Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *J. Comp. Phys.*, 242:480–497.

Gronskis A., Heitz D., M. E. (2013). Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *J. Comp. Phys.*, 242:480–497.

Haferssas, R., Jolivet, P., and Rubino, S. (2018). Efficient and scalable discretization of the navier–stokes equations with lps modeling. *Comput. Method. Appl. M.*, 333:371–394.

Hanjalić, K. and Launder, B. (1972). A reynolds stress model of turbulence and its application to thin shear flows. *J. Fluid Mech.*, 52(4):609–638.

He, C., Liu, Y., Gan, L., and Lesshafft, L. (2019a). Data assimilation and resolvent analysis of turbulent flow behind a wall-proximity rib. *Physics of Fluids*, 31(2):025118.

He, C., Liu, Y., Gan, L., and Lesshafft, L. (2019b). Data assimilation and resolvent analysis of turbulent flow behind a wall-proximity rib. *Phys. Fluids*, 31 and (025118).

Hecht, F. (2012). New development in freefem++. *J. Num. Math.*, 20(3-4):251–266.

Heitz, D. and Mémin, E. (2010). Variational fluid flow measurements from image sequences: synopsis and perspectives. *Exp. Fluids*, 48:369–393.

Hervé, A., Sipp, D., and Schmid, P. J. (2012). A physics-based approach to flow control using system identication. *J Fluid Mech.*, 702:26–58.

Huerre, P. and Monkewitz, P. A. (1985). Absolute and convective instabilities in free shear layers. *J. Fluid Mech.*, 159:151–168.

Hussain, A. F. (1983). Coherent structures—reality and myth. *Phys. Fluids*, 26(10):2816–2850.

Hussain, A. F. (1986). Coherent structures and turbulence. *J. Fluid Mech.*, 173:303–356.

Iglesias, M. A., Law, K. J., and Stuart, A. M. (2013). Ensemble kalman methods for inverse problems. *Inverse Problems*, 29(4):045001.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *J. Basic Eng.*, 82(1):35–45.

Kato, H. and Obayashi, S. (2013). Approach for uncertainty of turbulence modeling based on data assimilation technique. *Comput. Fluids*, 85:2–7.

Kato, H., Yoshizawa, A., Ueno, G., and Obayashi, S. (2015). A data assimilation methodology for reconstructing turbulent flows around aircraft. *J. Comp. Phys.*, 283:559–581.

Lazarus, A. and Thomas, O. (2010). A harmonic-based method for computing the stability of periodic solutions of dynamical systems. *Comptes Rendus Mécanique*, 338(9):510–517.

Le Dimet, F. X. and Talagrand, O. (1986). Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspect. *Tellus 38A*, pages 97–110.

Leclercq, C., Demourant, F., Poussot-Vassal, C., and Sipp, D. (2019). Linear iterative method for closed-loop control of quasiperiodic flows. *J. Fluid Mech.*, 868:26–65.

Li, Z., Zhang, H., Bailey, S. C. C., Hoagg, J. B., and Martin, A. (2017). A data-driven adaptive reynolds-averaged navier–stokes $\kappa - \omega$ model for turbulent flow. *J. Comp. Phys.*, 345:111–131.

Lions, J. L. (1971). *Optimal Control of Systems Governed by Partial Differential Equations (Grundlehren der Mathematischen Wissenschaften)*, volume 170. Springer Berlin.

Liu, C., Xiao, Q., and Wang, B. (2008). An ensemble-based four-dimensional variational data assimilation scheme. part i: Technical formulation and preliminary test. *Mon. Weather Rev.*, 136:3363–3373.

Lorenc, A. C. (1986). Analysis methods for numerical weather prediction. *Quart. J. R. Met. Soc.*, 112:1177–1194.

Lund, T. S. (1998). Generation of turbulent inflow data for spatially-developing boundary layer simulations. *J. Comp. Phys.*, 140:233–258.

Lyn, D., Einav, S., Rodi, W., and Park, J.-H. (1995). A laser-doppler velocimetry study of ensemble-averaged characteristics of the turbulent near wake of a square cylinder. *J. Fluid Mech.*, 304:285–319.

Mantič-Lugo, V., Arratia, C., and Gallaire, F. (2014). Self-consistent mean flow description of the nonlinear saturation of the vortex shedding in the cylinder wake. *Phys. Rev. Lett.*, 113(8):084501.

Mantic-Lugo, V., Arratia, C., and Gallaire, F. (2015). A self-consistent model for the saturation dynamics of the vortex shedding around the mean flow in the unstable cylinder wake. *Phys. Fluids*, 27.

Mantic-Lugo, V. and Gallaire, F. (2016). Self-consistent model for the saturation mechanism of the response to harmonic forcing in the backward-facing step flow. *J. Fluid Mech.*, 793:777–797.

Marquet, O. and Sipp, D. (2010). Global sustained perturbations in a backward-facing step flow. *IUTAM Bookseries*, 18.

McKeon, B. J. and Sharma, A. S. (2010). A critical-layer framework for turbulent pipe flow. *J. Fluid Mech.*, 658:336–382.

Meldi, M. (2018). Augmented prediction of turbulent flows via sequential estimators. *Flow Turbul. Combust.*, 101(2):389–412.

Meliga, P., Cadot, O., and Serre, E. (2016). Experimental and theoretical sensitivity analysis of turbulent flow past a square cylinder. *Flow Turbul. Combust.*, 97(4):987–1015.

Meliga, P., Pujals, G., and Serre, E. (2012). Sensitivity of 2-d turbulent flow past a d-shaped cylinder using global stability. *Phys. Fluids*, 24(6):061701.

Mettot, C., Renac, F., and Sipp, D. (2014a). Computation of eigenvalue sensitivity to base flow modifications in a discrete framework: Application to open-loop control. *J. Comp. Phys.*, 269:234–258.

Mettot, C., Sipp, D., and Bézard, H. (2014b). Quasi-laminar stability and sensitivity analyses for turbulent flows: prediction of low-frequency unsteadiness and passive control. *Phys. Fluids*, 26(4):061701.

Mezić, I. (2013). Analysis of fluid flows via spectral properties of the koopman operator. *Annu. Rev. Fluid Mech.*, 45:357–378.

Minguez, M., Brun, C., Pasquetti, R., and Serre, E. (2011). Experimental and high-order les analysis of the flow in near-wall region of a square cylinder. *International Journal of Heat and Fluid Flow*, 32(3):558–566.

Mittal, S. (2000). On the performance of high aspect ratio elements for incompressible flows. *Comput. Method. Appl. M.*, 188(1-3):269–287.

Mons, V., Chassaing, J. C., Gomez, T., and Sagaut, P. (2014). Is isotropic turbulence decay governed by asymptotic behavior of large scales? an eddy-damped quasi-normal markovian-based data assimilation study. *Phys. Fluids*, 115105.

Mons, V., Chassaing, J.-C., Gomez, T., and Sagaut, P. (2016). Reconstruction of unsteady viscous flows using data assimilation schemes. *J. Comp. Phys.*, 316:255–280.

Mons, V., Chassaing, J.-C., and Sagaut, P. (2017a). Optimal sensor placement for variational data assimilation of unsteady flows past a rotationally oscillating cylinder. *J. Fluid Mech.*, 823:230–277.

Mons, V., Margheri, L., Chassaing, J. C., and Sagaut, P. (2017b). Data assimilation-based reconstruction of urban pollutant release characteristics. *J. Wind Eng. Ind. Aerod.*, 169:232–250.

Morra, P., Semeraro, O., Henningson, D. S., and Cossu, C. (2019). On the relevance of reynolds stresses in resolvent analyses of turbulent wall-bounded flows. *arXiv preprint arXiv:1901.04356.*

Nakayama, A. and Vengadesan, S. (2002). On the influence of numerical schemes and subgrid–stress models on large eddy simulation of turbulent flow past a square cylinder. *Int. J. Numer. Meth. Fl.*, 38(3):227–253.

Nocedal, J. and Wright, S. (2006). *Numerical optimization.* Springer Science & Business Media.

Oliver, T. and Darmofal, D. (2007). An unsteady adaptation algorithm for discontinuous galerkin discretizations of the rans equations. In *18th AIAA Computational Fluid Dynamics Conference*, page 3940.

Oliver, T. A. (2008). *A High-Order and Adaptive and Discontinuous Galerkin Finite Element Method for the Reynolds-Averaged Navier-Stokes Equations.* PhD thesis, Massachusetts Institute of Technology.

Palkin, E., Mullyadzhanov, R., Hadžiabdić, M., and Hanjalić, K. (2016). Scrutinizing urans in shedding flows: the case of cylinder in cross-flow in the subcritical regime. *Flow Turbul. Combust.*, 97(4):1017–1046.

Parish, E. J. and Duraisamy, K. (2017). A paradigm for data-driven predictive modeling using field inversion and machine learning. *J. Comp. Phys.*, 305:758–774.

Persson, T., Liefvendahl, M., Bensow, R., and Fureby, C. (2006). Numerical investigation of the flow over an axisymmetric hill using les, des, and rans. *J. Turbul.*, (7):N4.

Powell, M. (1994). *A direct search optimization method that models the objective and constraint functions by linear interpolation.* Springer Netherlands, 1st edition.

Pujals, G., García-Villalba, M., Cossu, C., and Depardon, S. (2009). A note on optimal transient growth in turbulent channel flows. *Phys. Fluids*, 21(1):015109.

Raffel, M., Willert, C. E., Scarano, F., Kähler, C. J., Wereley, S. T., and Kompenhans, J. (2018). *Particle image velocimetry: a practical guide.* Springer.

Reynolds, W. and Hussain, A. (1972). The mechanics of an organized wave in turbulent shear flow. part 3. theoretical models and comparisons with experiments. *J. Fluid Mech.*, 54(2):263–288.

Rodi, W. (1997). Comparison of les and rans calculations of the flow around bluff bodies. *J. Wind Eng. Ind. Aerod.*, 69:55–75.

Schmidt, O. T., Towne, A., Rigas, G., Colonius, T., and Brès, G. A. (2018). Spectral analysis of jet turbulence. *J. Fluid Mech.*, 855:953–982.

Shaabani-Ardali, L., Sipp, D., and Lesshafft, L. (2017). Time-delayed feedback technique for suppressing instabilities in time-periodic flow. *Phys. Rev. Fluids*, 2(11):113904.

Singh, A. P. and Duraisamy, K. (2016). Using field inversion to quantify functional errors in turbulence closures. *Phys. Fluids*, 28.

Sipp, D., , Marquet, O., Meliga, P., and Barbagallo, A. (2010). Dynamics and control of global instabilities in open-flows: A linearized approach. *Appl. Mech. Rev.*, 63(3):030801.

Sipp, D. and Lebedev, A. (2007a). Global stability of base and mean flows: a general approach and its applications to cylinder and open cavity flows. *J. Fluid Mech.*, 593:333–358.

Sipp, D. and Lebedev, A. (2007b). Global stability of base and mean flows: a general approach and its applications to cylinder and open cavity flows. *J. Fluid Mech.*, 593:333–358.

Sohankar, A., Norberg, C., and Davidson, L. (1998). Low-reynolds-number flow around a square cylinder at incidence: study of blockage, onset of vortex shedding and outlet boundary condition. *Int. J. Numer. Methods Fluids*, 26:39–56.

Spalart, P. and Allmaras, S. (1992). A one-equation turbulence model for aerodynamic flows. In *30th aerospace sciences meeting and exhibit*, page 439.

Spalart, P. R. and Allmaras, S. R. (1994). A one-equation turbulence model for aerodynamic flows. *La Rech. Aérospatiale*, 1:5–21.

Speziale, C. G. (1991). Analytical methods for the development of reynolds-stress closures in turbulence. *Annu. Rev. Fluid Mech.*, 23(1):107–157.

Suzuki, T. (2012). Reduced-order kalman-filtered hybrid simulation combining particle tracking velocimetry and direct numerical simulation. *J. Fluid Mech.*, 709:249–288.

Symon, S., Dovetta, N., McKeon, B. J., Sipp, D., and Schmid, P. J. (2017). Data assimilation of mean velocity from 2d piv measurements of flow over an idealized airfoil. *Exp. Fluids*, 61:1–17.

Symon, S., Rosenberg, K., Dawson, S. T., and McKeon, B. J. (2018). Non-normality and classification of amplification mechanisms in stability and resolvent analysis. *Phys. Rev. Fluids*, 3(5):053902.

Symon, S., Sipp, D., and McKeon, B. J. (2019a). A tale of two airfoils: resolvent-based modelling of an oscillator vs. an amplifier from an experimental mean. *arXiv preprint arXiv:1904.10131.*

Symon, S., Sipp, D., Schmid, P. J., and McKeon, B. J. (2019b). Mean and unsteady flow reconstruction using data-assimilation and resolvent analysis. *AIAA Journal*, pages 1–14.

Talagrand, O. and Courtier, P. (1987). Variational assimilation of meteorological observations with the adjoint vorticity equation. i: Theory. *Q. J. Roy. Meteor. Soc.*, 113(478):1311–1328.

Towne, A., Schmidt, O. T., and Colonius, T. (2018). Spectral proper orthogonal decomposition and its relationship to dynamic mode decomposition and resolvent analysis. *J. Fluid Mech.*, 847:821–867.

Trias, F., Gorobets, A., and Oliva, A. (2015). Turbulent flow around a square cylinder at reynolds number 22,000: A dns study. *Comput. Fluids*, 123:87–98.

Turton, S. E., Tuckerman, L. S., and Barkley, D. (2015a). Prediction of frequencies in thermosolutal convection from mean flows. *Phys. Rev. E.*, 91 (4), 043009.

Turton, S. E., Tuckerman, L. S., and Barkley, D. (2015b). Prediction of frequencies in thermosolutal convection from mean flows. *Physical Review E*, 91(4):043009.

Verstappen, R. and Veldman, A. (1997). Fourth-order dns of flow past a square cylinder: First results. In *Direct and Large-Eddy Simulation II*, pages 381–384. Springer.

Yang, Y., Robinson, C., Heitz, D., and Mémin, E. (2015). Enhanced ensemble-based 4dvar scheme for data assimilation. *Comput. Fluids*, 115:201–210.

Yegavian, R., Leclaire, B., Champagnat, F., and Marquet, O. (2015). Performance assessment of piv super-resolution with adjoint-based data assimilation. *11th International symposium on particle image velocimetry, Santa Barbara, USA*, pages 14–16.

université
PARIS-SACLAY

ÉCOLE DOCTORALE
Sciences mécaniques et énergétiques, matériaux et géosciences (SMEMAG)

Titre : Stratégies de modélisation pour la reconstruction d'écoulements aérodynamiques à partir de mesures partielles

Mots clés : turbulence, assimilation de données, modélisation

Résumé : Dans un premier temps, nous nous intéresserons à la récupération du champ moyen à partir d'informations partielles ou éparses, allant des sondes de vitesse ponctuelles à la pression ou frottement de paroi. Pour le réaliser, on considère les équations de *Reynolds-Averaged Navier-Stokes* (RANS), complétées par un modèle, ici le Spalart-Allmaras. Ce type de modélisation a été conçu pour quelques configurations d'écoulement de référence et peut manquer de généralité, ce qui conduit à des prédictions erronées, surtout lorsqu'il y a recirculation. Nous modifions ce modèle avec un paramètre de contrôle tel que la solution modifiée corresponde le mieux aux données de champ moyen mentionnées ci-dessus. La configuration considérée est une marche descendante à $Re = 28275$, avec des données réelles provenant d'une DNS.
Ensuite, nous nous intéressons à l'analyse linéaire de champ moyen et à son utilisation pour prédire la fluctuation non linéaire instationnaire. En particulier, nous concevons un modèle d'ordre réduit, composé de l'équation du champ moyen couplé aux modes de résolvent, qui prédit la fluctuation pour chaque fréquence existante. Les énergies de ces modes sont utilisées comme paramètres à régler par la procédure d'assimilation des données, qui nécessite généralement (très) peu de donnée, typiquement des signaux résolus en temps issus de sondes ponctuelles. Cette technique sera appliquée dans des écoulements transitoires tels que celui autour d'un cylindre à section carrée, un cas de référence pour les écoulements oscillateurs, et une marche descendante, un écoulement type d'amplificateur de bruit.
Nous considérons ensuite un cas turbulent correspondant à l'écoulement autour d'un cylindre à section carrée à $Re = 22000$, ayant à la fois des caractéristiques d'oscillateur (émission périodique de vortices) et d'amplificateur de bruit (représenté par les structures Kelvin-Helmholtz). L'analyse classique de stabilité de champ moyen est utilisée pour récupérer le mode d'emission de vortex et une technique de résolvent, basée sur les équations linéarisées autour de la composante périodique, est utilisée pour récupérer la dépendance des modes Kelvin-Helmholtz avec l'emission de vortex.

Title : Modeling Strategies for Aerodynamic Flow Reconstruction from partial measurements

Keywords : turbulence, data-assimilation, modeling

Abstract : In a first moment we will be interested in the recovery of the mean-flow quantities from partial or sparse information, ranging from point-wise velocity probes to wall-pressure and friction. This will be achieved by considering the Reynolds-Averaged Navier-Stokes (RANS) equations, completed with a model, here the Spalart-Allmaras. This kind of modeling has been conceived for some benchmark flow configurations and may lack generality, leading to erroneous predictions, especially when re-circulation is present. We modify this model with a tuning parameter such that its solution matches the best the aforementioned mean-flow data. The configuration considered was a Backward-Facing Step at $Re = 28275$, with actual data stemming from a DNS.
Then, we turn our attention to linear mean-flow analysis and its use to predict the nonlinear unsteady fluctuation. In particular, we design a reduced-order model, composed by the mean-flow equation coupled with the resolvent modes, predicting the fluctuation for each existing frequency. The energies of those modes are used as tuning parameters for the data-assimilation procedure, that will take as input typically (very) few point-wise time-resolved information. This technique will be applied in transitional flows such as the one around a squared-section cylinder, a benchmark case for oscillator flows, and a backward-facing-step, a typical noise-amplifier flow.
We then consider a turbulent case corresponding to the flow around a squared-section cylinder at $Re = 22000$, having both oscillator (periodic vortex-shedding) and noise-amplifier-like characteristics (represented by the Kelvin-Helmholtz structures). Classical mean-flow stability analysis is used to recover the the vortex-shedding mode and a resolvent technique, based on the linearized equations around the periodic component, is used to recover the dependency of the Kelvin-Helmholtz modes with the vortex-shedding.

Université Paris-Saclay
Espace Technologique / Immeuble Discovery
Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France