



**HAL**  
open science

# Turbulent complex flows reconstruction via data assimilation in large eddy models

Pranav Chandramouli

► **To cite this version:**

Pranav Chandramouli. Turbulent complex flows reconstruction via data assimilation in large eddy models. Fluids mechanics [physics.class-ph]. Université de Rennes, 2018. English. NNT: 2018REN1S035 . tel-01951540

**HAL Id: tel-01951540**

**<https://theses.hal.science/tel-01951540>**

Submitted on 11 Dec 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1  
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 601  
*Mathématiques et Sciences et Technologies  
de l'Information et de la Communication*  
Spécialité : *Mathématiques et applications*

Par

**Pranav CHANDRAMOULI**

## **Turbulent complex flows reconstruction via data assimilation in large eddy models**

Thèse présentée et soutenue à Rennes, le 19 Octobre, 2018  
Unité de recherche : Fluminance (INRIA - IRSTEA - IRMAR), Université Rennes 1

### **Rapporteurs avant soutenance :**

Laurent Cordier, Chercheur, CNRS  
Jean-Phillippe Laval, Chercheur, CNRS

### **Composition du Jury :**

*Attention, en cas d'absence d'un des membres du Jury le jour de la soutenance, la composition du jury doit être revue pour s'assurer qu'elle est conforme et devra être répercutée sur la couverture de thèse*

Président : Roger Lewandowski, Professeur, Université de Rennes 1  
Examineurs : Chantal Staquet, Professeur, Université Grenoble Alpes  
Laurent Cordier, Chercheur, CNRS  
Jean-Phillippe Laval, Chercheur, CNRS  
Sylvain Laizet, Senior Lecturer, Imperial College London  
Dir. de thèse : Etienne Memin, Directeur de recherche, Inria  
Co-dir. de thèse : Dominique Heitz, Directeur de recherche, Irstea

*For my parents,*  
who got me here  
*and to Palak,*  
who got me through



## Acknowledgements

---

Etienne and Dominique, your guidance has ensured I traverse successfully the narrow, treacherous straits that is a PhD program. Your rigorous standards, academic familiarity and attitude towards research are traits I hope to inculcate and emulate. I would like to express my profound thanks to you for taking me in this program, and for guiding me, constantly, through it.

To Sylvain, without whom I would not be here, I am thankful for your guidance in everything numerical.

To my friends, past and present, in the corridor - 'Vincent, Benoit, Sandeep, Anca, Antoine, Manu, Valentin, and countless others', I thank you for all the special memories, fun and funny, we had over the past three years. Special mention to Long, for our mutual fifa addiction, Carlo, for our coffee, and Yacine, for our multitude of breaks. A big thank you to Huguette for taking care of all the administrative works that I would have, otherwise, been swamped with.

A final thanks to my family whose constant support and never-ending belief in me always motivated me to work harder and do something more.



# Contents

|   |           |
|---|-----------|
| <b>Introduction en français</b>                                   | <b>ix</b> |
| <b>Introduction</b>   | <b>xv</b> |
| <b>1 Fluid flow fundamentals</b>                                  | <b>1</b>  |
| 1.1 Computational fluid dynamics . . . . .                        | 2         |
| 1.1.1 A brief history of CFD . . . . .                            | 2         |
| 1.1.2 Governing equations . . . . .                               | 2         |
| 1.1.3 Discretisation techniques . . . . .                         | 3         |
| 1.1.4 Boundary conditions . . . . .                               | 5         |
| 1.1.5 Turbulence modelling . . . . .                              | 6         |
| 1.2 Experimental fluid dynamics . . . . .                         | 6         |
| 1.2.1 A brief history of EFD . . . . .                            | 6         |
| 1.2.2 Hot wire anemometry (HWA) . . . . .                         | 7         |
| 1.2.3 Laser doppler anemometry (LDA) . . . . .                    | 7         |
| 1.2.4 Particle image velocimetry (PIV) . . . . .                  | 8         |
| 1.3 Limitations . . . . .   | 10        |
| 1.4 Coupling EFD and CFD . . . . .                                | 11        |
| <b>2 Turbulence modelling</b>                                     | <b>15</b> |
| 2.1 Reynolds averaged Navier-Stokes . . . . .                     | 17        |
| 2.1.1 RANS equations . . . . .                                    | 17        |
| 2.1.2 Turbulent-viscosity models . . . . .                        | 18        |
| 2.1.3 Reynolds stress models (RSM) . . . . .                      | 21        |
| 2.2 Large eddy simulation . . . . .                               | 21        |
| 2.2.1 Filtered Navier-Stokes equations . . . . .                  | 22        |
| 2.2.2 Smagorinsky models . . . . .                                | 23        |
| 2.2.3 WALE model . . . . .  | 27        |
| 2.2.4 Models under location uncertainty . . . . .                 | 28        |
| 2.2.5 Other Approaches . . . . .                                  | 32        |
| 2.3 Hybrid approaches . . . . .                                   | 34        |
| <b>3 Large eddy simulation - applications</b>                     | <b>37</b> |
| 3.1 Flow solver - Incompact3d . . . . .                           | 37        |
| 3.2 Channel flow at $Re_\tau = 395$ . . . . .                     | 45        |
| 3.2.1 Channel flow parameters . . . . .                           | 45        |
| 3.2.2 Channel flow results . . . . .                              | 46        |
| 3.3 Wake flow around a circular cylinder at $Re = 3900$ . . . . . | 49        |
| 3.3.1 Introduction to wake flow . . . . .                         | 49        |

|          |  |            |
|----------|--|------------|
| 3.3.2    | Flow configuration and numerical methods . . . . .                     | 53         |
| 3.3.3    | Wake flow results . . . . .  | 56         |
| 3.4      | Concluding remarks . . . . .   | 76         |
| <b>4</b> | <b>Stochastic Simulations - vortex forces and Langmuir circulation</b> | <b>79</b>  |
| 4.1      | Introduction . . . . .   | 79         |
| 4.2      | Mathematical formulations . . . . .                                    | 80         |
| 4.2.1    | CL model . . . . .   | 80         |
| 4.2.2    | Large-scale drift MULU . . . . .                                       | 80         |
| 4.3      | Vortex force theory for streak formation . . . . .                     | 83         |
| 4.4      | Large-scale application . . . . .                                      | 88         |
| <b>5</b> | <b>Data assimilation techniques</b>                                    | <b>93</b>  |
| 5.1      | Mathematical Formulation . . . . .                                     | 94         |
| 5.2      | Data assimilation - a review . . . . .                                 | 94         |
| 5.3      | 4D-Var approach . . . . .  | 98         |
| 5.3.1    | 4D-Var Algorithm Variants . . . . .                                    | 100        |
| <b>6</b> | <b>Flow reconstruction - snapshot optimisation method</b>              | <b>105</b> |
| 6.1      | Introduction . . . . .   | 105        |
| 6.2      | Snapshot optimisation method . . . . .                                 | 107        |
| 6.2.1    | SO model concept . . . . .   | 107        |
| 6.2.2    | Mathematical formulation . . . . .                                     | 108        |
| 6.2.3    | Reduced order formulation . . . . .                                    | 110        |
| 6.2.4    | Averaging SO method . . . . .  | 111        |
| 6.3      | Application and results . . . . .                                      | 111        |
| 6.3.1    | Synthetic case-studies . . . . .                                       | 112        |
| 6.3.2    | Algorithm Enhancement Techniques . . . . .                             | 117        |
| 6.3.3    | Experimental case-studies . . . . .                                    | 120        |
| 6.4      | Conclusion . . . . .   | 127        |
| <b>7</b> | <b>Data assimilation - wake flow at <math>Re = 3900</math></b>         | <b>129</b> |
| 7.1      | Code Formulation . . . . .   | 129        |
| 7.1.1    | 4D-Var . . . . .   | 129        |
| 7.1.2    | 4D-Var <sub>les</sub> . . . . .  | 136        |
| 7.2      | 3D synthetic assimilation . . . . .                                    | 142        |
| 7.2.1    | Flow configuration . . . . .   | 142        |
| 7.2.2    | Observations construction . . . . .                                    | 144        |
| 7.2.3    | Initial background construction . . . . .                              | 146        |
| 7.2.4    | Results and Discussion . . . . .                                       | 146        |
| 7.3      | 2D synthetic assimilation . . . . .                                    | 152        |
| 7.3.1    | Flow configuration . . . . .   | 152        |
| 7.3.2    | Observations construction . . . . .                                    | 152        |
| 7.3.3    | Initial background construction . . . . .                              | 156        |
| 7.3.4    | Results and Discussion . . . . .                                       | 156        |
| 7.4      | Background covariance . . . . .  | 162        |
| 7.5      | Coefficient estimation . . . . .                                       | 165        |
| 7.6      | Combined case study . . . . .  | 168        |
| 7.7      | Conclusions . . . . .  | 172        |

**Conclusions and perspectives**

**175**



# Introduction en français

## Contexte

L'étude des fluides remonte à l'antiquité aux travaux d'Archimède tels que le fameux 'principe d'Archimède' sur la flottabilité. Au fil du temps, le domaine s'est diversifié dans l'étude des fluides stationnaires (ou de la statique des fluides) et des écoulements fluides (ou de la dynamique des fluides). La dynamique des fluides peut être subdivisée en l'étude de fluides newtoniens et non-newtoniens en fonction de la forme de la contrainte de frottement moléculaire. Les fluides newtoniens, qui sont les fluides étudiés dans cette thèse, ont un taux de déformation proportionnel à la contrainte appliquée et leur dynamique s'exprime mathématiquement au moyen des équations de conservation de la masse, de quantité de mouvement et d'énergie. L'étude de la mécanique des fluides a été la pierre angulaire de nombreuses avancées dans le domaine de l'aérospatial, de la prévision météorologique, des applications géophysiques, de l'automobile, du génie civil, etc. Des approches plus récentes et sophistiquées visent également à améliorer les conditions environnementales à l'intérieur de bâtiments (conception de salle blanche pour l'industrie pharmaceutique ou agroalimentaire (Arroyo et al., 2016)). Ces nombreuses applications soulignent l'importance des écoulements fluides dans de larges pans de notre vie quotidienne. Cela justifie pleinement l'étude poussée de ces écoulements.

L'analyse des écoulements fluides peut se faire de l'une des deux manières suivantes:

Expérimentalement: l'étude expérimentale des fluides (EFD) s'appuie sur des méthodes de mesures expérimentales telles que la vélocimétrie par image de particule, la vélocimétrie par suivi de particules, la vélocimétrie doppler laser, etc. Ces méthodes peuvent être intrusives comme c'est le cas pour l'anémométrie filaire (HWA) ou non intrusive (PIV, LDA, PTV). Des progrès récents ont conduit à l'élaboration de techniques de mesures 3D telles que la PIV tomographique (tomo-PIV) qui permettent d'appréhender le champ des vitesses 3D3C (trois dimensions, trois composantes) dans un volume donné. Ces techniques restent cependant encore assez limitées dans leur étendue spatiale.

Numériquement: Dynamique des fluides computationnelle (CFD) repose sur la simulation des équations mathématiques de la dynamique des fluides. Les équations de conservation décrivant la dynamique des fluides sont résolues en utilisant une grande variété de modèles numériques. La précision de ces modèles reste limitée par le coût de calcul requis pour une simulation. Avec l'émergence et le progrès constant des superordinateurs et des processeurs, des écoulements de plus en plus complexes sont simulés. Néanmoins, une grande variété d'écoulement à haute vitesse restent encore pour longtemps hors de portée de ces simulations.

Une combinaison de ces deux techniques a été utilisée efficacement, en particulier dans le domaine de la prévision météorologique où les données EFD sont utilisées afin de guider les simulations CFD. Ce couplage permet de s'affranchir ainsi partiellement des inconvénients associés à ces deux techniques. Une telle combinaison, appelée assimilation de

données (AD), est maintenant utilisée de plus en plus fréquemment dans la communauté de la mécanique des fluides avec une augmentation significative de publications au cours de la dernière décennie qui décrivent différentes techniques de couplage (Papadakis and Mémin, 2008; Gu and Oliver, 2007; Artana et al., 2012; Yang, 2014). Cette thèse étudie l'extension d'une technique d'assimilation proposée par Gronsks et al. (2013) pour des écoulements 2D et par Robinson (2015) pour des écoulements 3D. Le but de ce travail vise à étendre ces derniers travaux de Robinson (2015) à des écoulements plus complexes associés à des vitesses plus élevées en utilisant le concept de simulations aux grandes échelles de la turbulence (LES).

Les techniques de type LES ont d'abord été initiées par Smagorinsky (1963) avec la motivation de réduire le coût de calcul d'une simulation atmosphérique, en ne simulant que les grandes échelles des vitesses (les grands tourbillons) tandis que les effets des plus petites échelles sont modélisés. Smagorinsky a introduit un modèle simple qui reste très populaire dans le domaine LES. Au cours des décennies suivantes, des modèles plus complexes ont été introduits par un certain nombre d'auteurs tels que Germano et al. (1991); Nicoud and Ducros (1999), plusieurs simulation LES, ce qui rend son coût d'autant plus crucial. Dans le cadre des techniques d'assimilation de données s'avère nécessaire d'effectuer. L'objectif de cette thèse est de formuler et de valider des techniques LES pour l'assimilation de données permettant de réduire autant que possible les coûts de calcul associés.

## Aperçu des chapitres

Le chapitre 1 fournit une introduction détaillée des principes fondamentaux de la dynamique des fluides et des techniques utilisées pour la mesure des écoulements. Le chapitre 2 se concentre sur la modélisation numérique des écoulements turbulents et rassemble une abondante littérature sur les différentes méthodes / modèles existants pour la modélisation des écoulements turbulents. Le chapitre 3 montre l'application à divers écoulements de turbulents quelques modèles LES décrits dans le chapitre 2. Le chapitre 5 présente les techniques d'assimilation existantes avec la littérature associée. La méthode d'optimisation des instantanés (SO) pour la reconstruction rapide des écoulements 3D est développée et analysée dans le chapitre 6 comme un moyen de produire des conditions de base pour les études d'assimilation. La technique d'assimilation de données variationnelles (VDA) est la technique que nous avons appliquée et adaptée pour la reconstruction d'un écoulement de sillage dans le chapitre 7.

## Résumé du chapitre 1 - Principes fondamentaux de l'écoulement des fluides

Une brève description des principes fondamentaux régissant la dynamique des écoulements fluides est présentée dans ce chapitre. Les concepts de la dynamique des fluides computationnelle (CFD) et de l'analyse expérimentale (EFD) des écoulements. Les équations régissant les simulations numériques et les différentes techniques numériques qui leur sont associées sont présentées. Ceci est suivi d'un bref historique des outils d'analyse expérimental. Les techniques actuelles de mesure de vitesse d'écoulements sont décrites; il s'agit en particulier de la vélocimétrie par image de particule (PIV), la vélocimétrie par suivi de particules (PTV) et la tomographie-PIV (tomoPIV). Les avantages et les limites des approches computationnelles et expérimentales sont explorés, suivis par un historique du couplage entre modèles et données, c'est-à-dire de l'assimilation de données.

## Résumé du chapitre 2 - Modélisation de la turbulence

Ce chapitre se concentre sur la présentation des concepts liés à la modélisation de la turbulence. La délimitation entre les modèles de simulation numérique directe (DNS), LES et Reynolds Averaged Navier-Stokes (RANS) est décrite. Les équations du modèle RANS, dérivées des équations de conservation et les différents types de modélisation disponibles, à savoir les modèles linéaires de viscosité turbulente, les modèles de viscosité turbulente non linéaire et les modèles du tenseur de Reynolds sont ainsi présentés. Les équations de conservation filtrées sont dérivées pour les techniques LES suivi d’une description des techniques générales de filtrage. Des modèles classiques LES tels que le modèle Smagorinsky, le modèle de Smagorinsky dynamique et le modèle de viscosité “WALE” (Wall Adaptive Local-Eddy Viscosity) sont décrits. Des approches plus récentes de la modélisation LES, telles que les modèles sous maille implicites et les modèles sous maille stochastiques, sont également explorées. Des modèles récemment introduits et intitulé “modèle sous incertitude de position” (MULU) relevant d’une description stochastique de la turbulence sont présentés et accompagnés d’une section expliquant brièvement les principes de calcul stochastique utilisés pour dériver ce modèle. Des approches hybrides de la modélisation de la turbulence sont également abordées.

## Résumé du chapitre 3 - Simulation des grandes échelles - applications

Le but de ce chapitre est d’identifier un modèle sous maille (SGS) efficace et pouvant être utilisé dans les techniques d’assimilation de données utilisées dans le reste de cette thèse. Les différents modèles SGS présentés dans le chapitre 2 sont ici appliqués à deux écoulements turbulents différents: un écoulement de canal plan et un écoulement de sillage autour d’un cylindre circulaire. Le chapitre commence par une description du solveur numérique, Incompact3d, qui est un solveur parallélisé écrit en fortran et développé par Laizet and Lamballais (2009). Ce code, qui constitue également la base de notre code d’assimilation, est utilisé pour toutes les simulations de cette thèse. Incompact3d, qui est un solveur aux différences finies défini sur un maillage cartésien, repose sur des schémas de discrétisation d’ordre six et un solveur d’équations de Poisson spectral pour la pression. Les différents schémas temporels disponibles dans Incompact3d sont détaillés. Ce code DNS a été modifié pour incorporer les différents modèles LES décrits dans cette thèse.

Une première analyse de modèle est réalisée avec un écoulement de canal à un nombre de Reynolds ( $Re$ ) défini à partir de la vitesse de frottement de 395. Cet écoulement a été choisi au raison de l’homogénéité de l’écoulement et de l’absence d’un objet solide. Les performances du modèle sont ensuite analysées sur un écoulement de sillage plus complexe autour d’un cylindre circulaire à un Reynolds de 3900. Cet écoulement est particulièrement important dans cette thèse car il est le point central de nos études d’assimilation de données. La performance du modèle est étudiée et des résultats intéressants sont détaillés. Le travail présenté dans ce chapitre a été publié dans Chandramouli et al. (2017b, 2018).

## Résumé du chapitre 4 - Simulations stochastiques - forces de vortex et circulation de Langmuir

L’effet de la turbulence dans la structuration d’un écoulement de fluide reste une question ouverte en mécanique des fluides et particulièrement dans la communauté géophysique. Dans les écoulements océaniques, ceci est bien caractérisé à travers l’interaction vague-courant via l’étude de la circulation secondaire de Langmuir. La turbulence induite par les vagues produit ce que l’on appelle une “force de vortex” qui entraîne des stries de

l'écoulement. Des stries similaires sont observées dans des écoulement pariétaux. L'origine de ces stries reste ouverte à interprétation. Cette étude propose une interprétation de l'origine des stries dans les écoulements de fluides pariétaux, à travers une vision stochastique de la turbulence. Une analogie est établie avec la structuration induite par les vagues observée dans les écoulements océaniques. L'analyse de l'inhomogénéité de la turbulence dans une simulation à haute résolution d'un écoulement pariétal montre la présence de ces forces de vortex. En modélisant l'effet de cette inhomogénéité dans des simulations à grande échelle dans le cadre stochastique, des stries sont observées. Ce chapitre est un précurseur à la réalisation d'une assimilation de données de type méthode d'ensemble à l'aide d'un modèle dynamique entièrement stochastique. L'objectif de ce chapitre est d'analyser les capacités de la formulation stochastique de MULLU. L'identification d'un modèle stochastique dérivée des lois de conservation physiques peut ouvrir la voie à une assimilation de données basée sur un ensemble en utilisant un ensemble de réalisations physiquement significatif. Cela constituerait un écart appréciable par rapport à la norme actuelle en matière d'assimilation de données d'ensemble où les réalisations sont obtenues par perturbation de la condition initiale et/ou des paramètres du modèle et sont peu contraintes par des lois physiques.

## Résumé du chapitre 5 - Techniques d'assimilation de données

Ce chapitre fournit une description approfondie des différentes techniques d'assimilation de données suivie d'une caractérisation détaillée du code '4D-Var' utilisé dans cette thèse. Les techniques d'assimilation ont évolué en partant de simples techniques d'estimations linéaires pour aller vers des techniques séquentielles et variationnelles plus complexes. Des développements récents ont également vu apparaître des techniques hybrides combinant des méthodes séquentielles et variationnelles afin de réunir les avantages de ces deux cadres. Cette thèse qui suit les travaux de Gronsksis et al. (2013) et Robinson (2015) est centrée sur les techniques variationnelle d'assimilation dites "4D-Var". Cette méthode reposant sur une formulation incrémentale de contrôle optimal a été implémenté dans 'Incompact3d' par Gronsksis pour un écoulement 2D (Gronsksis et al., 2013) et étendue dans les cas 3D par Robinson (2015). La formulation adjointe du code est utilisée pour calculer le gradient qui est minimisé en utilisant une technique de descente de gradient. Le développement du code 4D-Var et les procédures importantes qui lui sont associées sont détaillés.

## Résumé du chapitre 6 - Reconstruction d'écoulements - méthode d'optimisation de snapshots

Dans ce chapitre, une technique très rapide de reconstruction d'écoulements est proposée. Cette technique exploite des directions d'homogénéité de l'écoulement afin de recréer des champs de vitesse turbulents instantanés tridimensionnels à partir de deux séquences de mesures de vitesse 3d sur deux plans perpendiculaires. Cette méthodologie, appelée "optimisation de cliché" ou SO, peut aider à fournir des ensembles de données 3D pour la reconstruction d'écoulements à partir de mesures expérimentales. La méthode SO vise à optimiser l'erreur entre un plan d'entrée de direction homogène et des mesures instantanés de vitesse, observé sur une période de temps suffisante, sur un plan d'observation. Les observations sont effectuées sur un plan perpendiculaire au plan d'entrée les deux plans partageant un bord perpendiculaire à la direction d'homogénéité. La méthode est applicable à tous les écoulements présentant une direction d'homogénéité, tels que les écoulements de sillage symétriques, les écoulements de canal, la couche de mélange et les jets incidents

(axi-symétrique). Les résultats de la méthode sont évalués avec deux ensembles de données synthétiques, et trois ensembles de données PIV expérimentaux.

### **Résumé du chapitre 7 - Assimilation de données - écoulement de sillage à $Re = 3900$**

Les concepts présentés dans le chapitre précédent sont adaptés à un modèle LES puis appliqués à l'écoulement de sillage turbulent 3D à  $Re = 3900$ . Il s'agit de l'écoulement sur lequel le modèle LES a été validé au chapitre 3. Afin d'effectuer l'assimilation avec un modèle LES, la formulation tangente et adjointe du modèle sous maille doit être incorporée dans le code 4D-Var. Un paramètre de contrôle additionnel sur le coefficient du modèle LES a été introduit dans ce cadre. La performance du code 4D-Var volumiques LES a ensuite été évalué sur un modèle synthétique de référence. L'algorithme a été testé dans un premier temps avec des observations volumique idéales et avec des observations 2D synthétiques émulant un ensemble de données expérimentales. L'estimation du coefficient pour le modèle LES est également analysée.



# Introduction

## Context

The study of fluids dates back centuries to the works of Archimedes such as the famous ‘Archimedes’ Principle’ on buoyancy. Over the course of time the field has developed and diverged into the study of stationary fluids or fluid statics and flowing fluids or fluid dynamics. Fluid dynamics can be further sub-categorised into Newtonian and non-Newtonian fluids depending on the strain-stress relationship. Newtonian fluids, which are the fluids under study in this thesis, have a strain rate proportional to the applied stress and their dynamics are mathematically expressed through the conservation equations of mass, momentum and energy. The study of fluid mechanics has been the cornerstone for numerous advancements in the field of aerospace, weather prediction, geophysical applications, automobile industry, civil engineering, etc. More recent and sophisticated approaches have been aimed at improving indoor climate conditions, for clean-air requirements (such as in the pharmaceutical or food processing industry, see Froiloc<sup>®</sup> (Arroyo et al., 2016)), and outer-space exploration among many other things. The numerous applications indicate a permeance of fluid flows into large aspects of our day to day lives. This warrants further studies of these flows to characterise and utilise them efficiently towards solving our day-to-day problems.

The study of fluid flows can be done in one of the two following ways:

Experimentally: experimental fluid dynamics (EFD) deals with the study of fluid flows using experimental methods such as particle image velocimetry (PIV), particle tracking velocimetry (PTV), laser doppler velocimetry (LDA), etc. These methods can be intrusive (hot wire anemometry (HWA)) or non-intrusive (PIV, LDA, PTV). Recent advancements in the field of measurement techniques has led to the developed of 3D measurement techniques such as tomographic PIV (tomo-PIV) which capture the 3D3C (three dimensional, three component) velocity field in a given flow. However, such techniques are still for the moment quite limited in their spatial extent.

Numerically: computational fluid dynamics (CFD) deals with the study of fluid flows by simulating mathematical equations of fluid dynamics. The conservation equations describing various fluid flows are solved using a wide variety of dynamical models limited only by the computational cost required for a simulation and the ability to represent the fluid flow via mathematical equations. With the invention of super-computers and improved processors, more complex flows are being simulated, however, a wide variety of high speed flows still remain out of reach nowadays.

A combination of both these techniques has been used effectively especially in the field of weather forecasting where EFD data is used to guide CFD simulations thus negating partially the disadvantages of both these techniques. Such a combination, termed as Data Assimilation (DA), is now being used more and more frequently in the community of fluid mechanics with a surge of publications over the past decade delineating various techniques

(Papadakis and Mémin, 2008; Gu and Oliver, 2007; Artana et al., 2012; Yang, 2014). This thesis expands on one such DA technique proposed by Gronsksis et al. (2013) for 2D flows and extended to 3D flows by Robinson (2015). The aim of this work is to extend the works of Robinson (2015) to more complex flows at higher speeds by using the concept of large eddy simulations (LES).

LES techniques were first introduced by Smagorinsky (1963) with the motivation of reducing computational cost of a simulation by simulating only the large velocity scales or ‘large eddies’ while the smaller scales are modelled using a dynamical model. Smagorinsky introduced a simple model which remains a popular model till date for LES studies. Over the next few decades, leading from 1963, more complex models were introduced by the works of Germano et al. (1991); Nicoud and Ducros (1999); Mémin (2014) which were shown to perform well for various flow dynamics. A similar situation to Smagorinsky is currently faced by the DA community where the cost of performing a DA simulation in 3D is strictly prohibitive. The focus of this thesis is to formulate and validate LES techniques within DA with the aim of reducing computational costs thus opening avenues of DA research into complex, high speed flows.

## Preview of chapters

Chapter 1 provides a detailed introduction to fluid flow fundamentals and the measurement techniques used for fluid flow measurements. Chapter 2 focuses on numerical modelling of turbulent flows and gathers extensive literature on the various existing methods/models for modelling turbulent flows. Chapter 3 applies the LES models described in chapter 2 to various turbulent flows. Chapter 5 enumerates existing DA techniques with associated literature. The snapshot optimisation (SO) method for fast reconstruction of 3D flows is developed and analysed in chapter 6 as a means for producing background conditions for DA studies. The variational data assimilation (VDA) technique w.r.t. LES is developed and applied to wake flow around a cylinder in chapter 7.

### Summary of chapter 1 - Fluid flow fundamentals

A brief description of fluid flow fundamentals is provided in this chapter dealing with the basics of CFD and EFD. The governing equation for CFD are enumerated along with the different numerical techniques developed for solving the mathematical description of the fluid flow. Numerical schemes and methods are also described. This is followed by a brief history of EFD. The current state-of-the-art techniques for fluid flow measurements are described such as PIV, PTV, and tomoPIV. The advantages and limitations of CFD and EFD are explored, followed by a history of the coupling of EFD with CFD, i.e. data assimilation.

### Summary of chapter 2 - Turbulence modelling

This chapter focuses on the concepts of turbulence modelling. The delineation between Direct Numerical Simulation (DNS), LES, and Reynolds Averaged Navier-Stokes (RANS) models is given. RANS model equations are derived from the conservation equations and the avenues of modelling available in RANS, namely linear eddy viscosity models, non-linear eddy viscosity models, and Reynolds stress models are explored. The filtered conservation equations are derived for LES followed by a description of the general filtering techniques. Classical models for LES such as Smagorinsky model, dynamic Smagorinsky model, and Wall Adaptive Local-Eddy viscosity model are described. Newer approaches

to LES modelling such as implicit-LES and stochastic LES are also explored. The recently developed models under location uncertainty (MULU) falling under the stochastic LES umbrella of approaches are derived following a section explaining briefly the principles of stochastic calculus used to derive the MULU. Hybrid approaches to turbulence modelling are also touched upon.

### **Summary of chapter 3 - Large eddy simulation - applications**

The aim of this chapter is to identify an optimal LES sub-grid scale (SGS) model which can form the basis for DA-LES studies in further chapters. The different SGS models detailed in chapter 2 are applied to two turbulent flows: Channel Flow and Wake Flow around a Circular Cylinder. The chapter begins with a description of the flow solver, Incompact3d, which is a fortran based parallelised fluid flow solver developed by Laizet and Lamballais (2009). This code, which also forms the basis for our DA code, is used for all the simulations in this thesis. Incompact3d, which is a finite difference solver based on a cartesian mesh, works with sixth-order discretisation schemes and a spectral pressure-poisson solver. The different time-stepping schemes and flow configurations available in Incompact3d are detailed. The code which inherently performs a DNS, is modified to incorporate LES models for analysis. The different models included in Incompact3d for the purpose of this thesis are delineated.

A first model analysis is performed with channel flow at a friction velocity based Reynolds (Re) number of 395. Channel flow is chosen as the preliminary application due to homogeneity within the flow and lack of solid boundaries which make this an ideal flow for preliminary analysis. The model performances are then analysed on the more complex wake flow around a circular cylinder at a Re of 3900. This flow is of particular importance as wake flow around a cylinder is the focus for our data assimilation techniques explored in further chapters. The model performance is studied and interesting results are discussed in detail. The work presented in this chapter has been published in Chandramouli et al. (2017b, 2018).

### **Summary of chapter 4 - Stochastic Simulations - vortex forces and Langmuir circulation**

An important open question common to the fluid mechanics and oceanic communities is the effect of turbulence in the structuration of a fluid flow. In oceanic flows, this is well characterised in wave-current interactions through the study of Langmuir circulation. Wave induced turbulence produces a so called “vortex force” that results in streaking of the flow. Similar streaks are observed on wall-bounded flows of interest in fluid mechanics. The origin of such streaks remains open for interpretation. This study provides one such interpretation for the origin of streaks in wall-bounded fluid flows, through a stochastic view of turbulence. An analogy is drawn with the wave-induced structuration seen in oceanic flows. The analysis of turbulence inhomogeneity in a high-resolution simulation of wall-bounded flow, shows the presence of these vortex forces. By modelling the effect of this inhomogeneity in low-resolution large-scale simulations within the stochastic framework, streaks are observed. This chapter is to act as a precursor for performing ensemble based data assimilation using a fully stochastic turbulence model. The focus of the chapter is on analysing the capabilities of the stochastic form of the MULU. The identification of a well-posed stochastic model can pave the way for performing ensemble based data assimilation using physically meaningful ensemble of realisations. This would be a welcome deviation from the current norm in ensemble data assimilation where realisations are ob-

tained through perturbation with random vectors of the initial condition and/or model parameters which are not constrained by physical laws.

### **Summary of chapter 5 - Data assimilation techniques**

This chapter provides a fundamental description of the different DA techniques followed by a detailed characterisation of the code ‘4D-Var’ used in this thesis. DA techniques have evolved from simple linear estimation techniques to more complex sequential and VDA techniques. Recent developments have also seen some hybrid techniques combining sequential and variational methods to provide accurate results. This thesis following the works of Gronsks et al. (2013) and Robinson (2015) focuses on the 4D-Var VDA technique. This method based on incremental variational assimilation with additional control parameters is implemented on the NS solver ‘Incompact3d’ by Gronsks in 2D and expanded into 3D by Robinson. The adjoint formulation of the code is used to calculate the gradient which is minimised using a gradient-descent based optimiser. The development of the 4D-Var code and important routines with associated literature is explained in detail.

### **Summary of chapter 6 - Flow reconstruction - snapshot optimisation method**

In this chapter, a computationally efficient flow reconstruction technique is proposed, exploiting homogeneity in a given direction, to recreate three dimensional instantaneous turbulent velocity fields from snapshots of two dimension planar fields. This methodology, termed as ‘snapshot optimisation’ or SO, can help provide 3D data-sets for studies which are currently restricted by the limitations of experimental measurement techniques. The SO method aims at optimising the error between an inlet plane with a homogeneous direction and snapshots, obtained over a sufficient period of time, on the observation plane. The observations are carried out on a plane perpendicular to the inlet plane with a shared edge normal to the homogeneity direction. The method is applicable to all flows which display a direction of homogeneity such as cylinder wake flows, channel flow, mixing layer, and jet (axisymmetric). The ability of the method is assessed with two synthetic data-sets, and three experimental PIV data-sets. A good reconstruction of the large-scale structures are observed for all cases. The small-scale reconstruction ability is partially limited especially for higher-dimensional observation systems. POD based SO method and averaging SO variations of the method are shown to reduce discontinuities created due to temporal mismatch in the homogenous direction providing a smooth velocity reconstruction. The volumetric reconstruction is seen to capture large-scale structures for synthetic and experimental case-studies. The algorithm run-time is found to be in the order of a minute providing results comparable with the reference. Such a reconstruction methodology can provide important information for data assimilation in the form of initial condition, background condition, and 3D observations.

### **Summary of chapter 7 - Data assimilation - wake flow at $Re = 3900$**

The concepts developed in the previous chapter are expanded to accommodate for an LES model and then applied to 3D turbulent wake flow at  $Re = 3900$  - the flow on which the LES model was validated for in chapter 3. In order to perform VDA with an LES model, the tangent and adjoint formulation of the model need to be incorporated into the 4D-Var code. An addition control on the VDA technique is introduced through the LES model coefficient. The performance of the 4D-Var code with LES or ‘4D-Var<sub>les</sub>’ is

characterised. The algorithm is tested with ideal 3D observations as well as synthetic 2D observations emulating experimental data-sets. Coefficient estimation for the LES model is also analysed with the algorithm.



# Chapter 1

## Fluid flow fundamentals

The first question to address, when it comes to the study of fluid flows, is the why? Why do we need to study fluids, their flow, and their interactions? This question can be addressed accurately by the well-known Tacoma bridge incident of 1940. The 1940 Tacoma narrow bridge was a suspension bridge built during 1938-1940 to span across the Tacoma narrow strait of Puget Sound. Opened to the public on July 1st, 1940, the bridge famously collapsed on November 7th in the same year barely 4 months in. Figure 1.1 shows in snapshots the torsional oscillation of the bridge and its final collapse.



Figure 1.1: Tacoma Bridge Collapse - November 7th, 1940

The sudden failure of this bridge and its spectacular collapse has been recurrently referred to in major fluid mechanics and physics textbooks over the past few decades. The capture of the bridge's torsional oscillation and collapse by surrounding photographers has increased its prominent presence in major works. The bridge has been used as an example of resonance within structures due to fluid-structure interaction and as an example for the importance of the study of fluids mechanics. While the import of fluid mechanics is further established, the concept of resonance as the reason for collapse has been debunked by the work of Billah and Scanlan (1991). Upon further study using 1/50th scale models in wind-tunnels, the authors conclude that the collapse was induced due to a concept called 'aeroelastic flutter'. Aeroelastic flutter, or more specifically bridge flutter, is a phenomenon where high speeds of wind causes an unconstrained oscillatory motion in a structure which in turn induces vortices as opposed to the view where vortex sheddings induce motion in a structure, i.e. 'vortex-induced vibrations'. Billah and Scanlan (1991) provide a well-written concise view on the different phenomena associated with the Tacoma bridge collapse expressing both accurate and prevalent inaccurate view points. It is interesting to note that flows involving fluid-structure interaction, wakes, and vortices also forms the basis for this thesis work.

A comprehensive study of fluid mechanics could have well averted the situation of the Tacoma bridge collapse. This chapter addresses the concepts of how to study fluid mechanical principles and concepts. Two prominent techniques exist for fluid flow study

namely computational fluid dynamics and experimental fluid dynamics. §1.1 addresses the important points of CFD while §1.2 focuses on the methods and advancements in the field of EFD. The limitations of both methods are presented in §1.3 while a combination of the two methods is presented in §1.4.

## 1.1 Computational fluid dynamics

### 1.1.1 A brief history of CFD

Computational fluid dynamics largely revolves around solving the Navier-Stokes (NS) equations (explained in further sections). However, in 1930s, when CFD was in its incipient stage, due to a lack of computers, the equations had to be solved by hand. This was an impossible task and preliminary applications of CFD saw the reduction of the complex NS equation into simplistic 2D expressions which could be hand-solved. A major upgrade to the CFD community was the advent of computers in 1950s following which the first practical CFD application was done at the Los Alamos National Lab. The complexity of equations solved increased slowly from a 2D panel method (1968) to full potential codes (1970) to solving the Euler equations (1981) and finally the NS equations (1987). A preliminary industrial application of CFD was done by Douglas Aircraft in 1967 followed by Boeing and Nasa in the 1980s. The usefulness of CFD for automobile industry was realised mid 1990s and was subsequently adopted by Ford and GM resulting in the aerodynamically streamlined cars of the present. The field has expanded drastically with the improvement of computational resources as well as development of better mathematical models. CFD is now used as a practical tool in most fields including but not restricted to automobiles, aeronautics, aerospace, civil, energy, chemical, geophysics, climate sciences, and turbo-machinery.

### 1.1.2 Governing equations

The conservation equations for mass, and momentum form the basis of any CFD simulation. The equation for conservation of mass or the continuity equation can be expressed as,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \quad (1.1.1)$$

where  $\rho$  stands for density,  $t$  for time, and  $\mathbf{u}$  for the components of velocity. The continuity equation can be simplified under the assumption of an incompressible flow ( $\rho$  is assumed constant) as,

$$\nabla \cdot \mathbf{u} = 0. \quad (1.1.2)$$

The momentum conservation equation or the NS equation for an incompressible fluid is expressed as,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = \mathbf{g} - \frac{1}{\rho} \nabla p + \nu \Delta \mathbf{u}, \quad (1.1.3)$$

where  $\mathbf{g}$  stands for gravity force,  $p$  stands for pressure,  $\nu$  for kinematic viscosity and  $\Delta$  is the laplacian operator. For all incompressible fluid flow simulations considered in this thesis,  $\mathbf{g}$  is neglected and density is considered subsumed with the pressure term. An additional equation for conservation of energy is used for certain applications such as combustion and heat transfer flows - such flows are not considered in this thesis. The various types of fluid flows that can be expressed through modifications of the NS equations set is shown in figure 1.2.

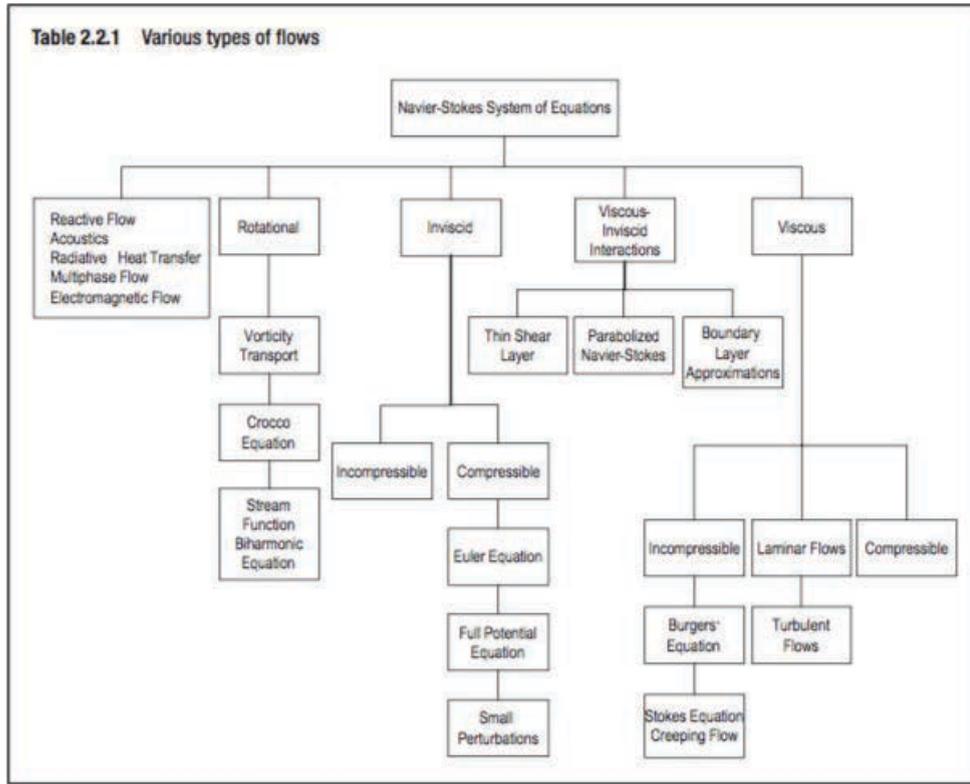


Figure 1.2: Flow types that can be depicted by the Navier-Stokes equation set. Borrowed from Chung (2010).

### 1.1.3 Discretisation techniques

The non-linearity of the term  $\mathbf{u} \cdot \nabla \mathbf{u}$  makes the analytical solution of the NS equation unachievable for all but the simplest of fluid flow problems. Thus, practical solutions of the NS equations can only be obtained, at present, by computational simulations using a discretised form of the equations. This section expands on existing numerical discretisation techniques.

#### Finite difference method

In finite difference methods (FDM), the partial differential equations (PDE) to be solved are approximated using difference equations within which the derivatives are expressed through finite differences. The expression for the derivatives of a function  $f(x)$  around a point  $x_0$  can be derived from Taylor's polynomial:

$$f(x_0 + h) = f(x_0) + \frac{f'(x_0)}{1!}h + \frac{f^{(2)}(x_0)}{2!}h^2 + \dots + R_n(x), \quad (1.1.4)$$

where  $h$  is the step size, and  $R_n(x)$  stands for the higher-order expansions beyond the order  $n$  expressed by the equation. The derivative can thus be expressed as,

$$f'(x_0) = \frac{f(x_0 + h) - f(x_0)}{h}, \quad (1.1.5)$$

where we assume the higher order terms are sufficiently small to be neglected.

Many numerical methods can be envisaged to approximate the derivatives. For a velocity field  $u$  in 1D space  $x$  with a step size of  $dx$ , the simplest approximations are,

Forward difference:

$$\frac{\partial u_i}{\partial x} = \frac{u_{i+1} - u_i}{dx} + O(dx), \quad (1.1.6)$$

Backward difference:

$$\frac{\partial u_i}{\partial x} = \frac{u_i - u_{i-1}}{dx} + O(dx), \quad (1.1.7)$$

Central difference:

$$\frac{\partial u_i}{\partial x} = \frac{u_{i+1} - u_{i-1}}{2dx} + O(dx^2), \quad (1.1.8)$$

Second order derivative:

$$\frac{\partial^2 u_i}{\partial x^2} = \frac{u_{i+1} - 2u_i + u_{i-1}}{dx^2} + O(dx^2), \quad (1.1.9)$$

where  $O$  stands for neglected higher order terms, and  $i$  is an index corresponding to the point on which the derivatives are being expressed.

Similarly any order of derivatives can be derived from Taylor's approximation using any specific number of points. These formulae can also be expanded easily to 2D and 3D. For example, the laplacian in 2D for an equi-spaced mesh can be expressed as,

$$\Delta u_{ij} = \frac{u_{i+1,j} + u_{i-1,j} + u_{i,j+1} + u_{i,j-1} - 4u_{i,j}}{dx^2} + O(dx^4, dy^4), \quad (1.1.10)$$

where  $i$  and  $j$  are indices in 2D for the  $x$  and  $y$  directions respectively.

The accuracy of a finite difference scheme is assessed in three ways. The scheme is consistent if the approximation tends to the partial differential equation as the grid size is reduced. The scheme is considered stable if its error remains bounded. Finally, the results of a convergent scheme approach the results of the original equation as the grid size approaches zero. The advantage of FDM lies in its intuitive nature and ease of formulation. However, implementation becomes difficult with complex geometries and moving boundaries.

## Finite volume method

The finite volume method (FVM) provides an alternative to FDM for spatial discretisation where the integral form of the equation set is reduced to a set of ordinary differential equations (ODEs). FVM is used commonly in the field of CFD due to the conservative nature of the resultant equations as well as their applicability to unstructured, irregular meshes without any coordinate transformation.

Given a conservation equation of  $Q$  over a volume  $V$ ,

$$\frac{\partial Q}{\partial t} + \nabla \cdot F = Pr, \quad (1.1.11)$$

with  $F$  denoting the flux of  $Q$  per unit volume per unit time and  $Pr$  the production rate of  $Q$ , its integral form can then be written as,

$$\frac{d}{dt} \int_{V(t)} Q dV + \oint_{S(t)} n \cdot F dS = \int_{V(t)} (Pr) dV. \quad (1.1.12)$$

To solve this integral form, the volume  $V$  is broken into multiple control volumes which collate to form the total volume  $V$ . For each control volume, certain approximations are applied to calculate the value of  $Q$  and thus the flux  $F$  at the boundary of each volume along with the production  $Pr$  within the volume. A time marching scheme is then applied to advance the solution of  $Q$  in time.

## Finite element method

Finite element method (FEM) was first introduced in the field of structural mechanics and was later adopted for solving fluid flow problems especially when complex geometry was involved. The ability of the method to decompose the domain into smaller elements or finite elements of flexible geometry highlights the main advantage of the FEM. In addition, in FEM, the weak form of the governing equation is approximated unlike FDM/FVM where we work with the strong form of the governing equations. This allows the ability to provide higher order of approximations as well as flexibility in the implementation of boundary conditions. The weak form is obtained by the inner product of the residual of the governing equations and chosen test functions/basis functions. Such an equation represents the orthogonal projection of the residual error onto the test functional subspace integrated over the domain. Equating this to zero minimises the error functional thus providing the best approximation of the solution to the governing equation. While FEM are effective for structural problems, in the field of CFD involving a large number of grid points, the use of simpler schemes with a lower order approximation such as FDM/FVM is preferred. FEM within CFD is generally applied in the case of complex geometries mainly with industrial applications. For more information about different discretisation techniques with examples, please refer to Chung (2010); Pulliam and Zingg (2014).

### 1.1.4 Boundary conditions

Boundary conditions play an important role in CFD. They define the behaviour of the fluid at the boundaries outside of which no information of the fluid flow is known. Thus these conditions need to be defined accurately in space and for transient cases, an additional condition defining the flow at the initial time step is required. An error in the boundary condition can lead to errors in simulation as well as cause issues in convergence and stability of the simulation. In general, the boundary condition can take one of three forms: for a Dirichlet condition, a value for the function is defined at the boundary, while for a Neumann condition, the normal derivative of the function at the boundary is defined and a combination of the two is defined as the Robin boundary condition. Common boundary conditions for CFD (for velocity fields) are:

- **Inlet condition** defines the value of the fluid flow velocity at the inlet boundary. Wake flows, channel flows, and pipe flows are a few example which use the inlet boundary condition.
- **Outlet condition** defines the value of the fluid flow velocity at the outlet boundary when this velocity field is known. Channel flows, and pipe flows are a few example which use the outlet boundary condition.
- **No slip condition** is used to define the velocity at solid boundaries such as the wall of a channel flow. The normal component of the velocity is set to zero while the tangential component is set to the wall speed. The boundary on the cylinder, the walls of a channel or pipe are examples where the no slip condition is implemented.
- **Constant pressure condition** defines the value of pressure usually at inlet or outlet boundaries especially when the velocity field is unknown. Buoyancy driven flows, free surface flows, and flows around buildings are examples which can employ the constant pressure condition.

- **Periodic condition** are applied in flows where a repeated flow pattern is observed along a given direction. This condition is applied for the theoretical Taylor-Green vortex simulations, as well as in channel flow for inlet/outlet boundaries and for wake flows in the spanwise direction of the cylinder.

### 1.1.5 Turbulence modelling

To accurately simulate a fluid flow, all the length and time scales of the fluid need to be resolved explicitly by the discretisation mechanism employed. However, for the case of turbulent flows, the range of length scales increases exponentially with the Reynolds number and is, in general, computationally prohibitive to resolve all of these scales in a simulation. Thus, an alternative approach is required to model the smaller scales of turbulent motion which are not captured by the discretisation method. Such an approach is referred to as turbulence modelling. A simulation resolving all the scales of motion is referred to as a direct numerical simulation (DNS) and is the most accurate form of simulation possible - DNS results are used as a reference to assess model performance. Depending on the range of scales resolved by the simulations, and equivalently, the range of scales modelled by a model, the classification of the turbulence model can be done. Large eddy simulations resolve most of the turbulent scales while modelling the smallest scales removed from the flow by a filtering mechanism. In the Reynolds-Averaged Navier-Stokes equations, which is the oldest approach to turbulence modelling, the time-averaged velocity field is obtained by simulating the time-averaged governing equations where the contribution of the fluctuation components are expressed through additional terms known as Reynolds stresses which need to be modelled. A more detailed study of turbulence modelling is presented in chapter 2 elaborating on LES and RANS models as well as other methods available in literature.

## 1.2 Experimental fluid dynamics

### 1.2.1 A brief history of EFD

A succinct and well-written history of the major contributors and their contributions to fluid dynamics, as we know it, is provided by Anderson (2010). Here, I summarise in brief the contribution related to the field of experimental developments for fluid flow measurement. The first diagram of flow streamlines around an object dates back 500 years to Leonardo Da Vinci. He also stated vaguely the continuity equation for low-speed flows as well as stating the 'wind-tunnel principle' - the relative flow over a stationary body mounted in a wind-tunnel is the same as the relative flow over the same body moving through a stationary fluid. Following this, experiments on fluids were conducted on fluid statics during the 17-18th century. An important advancement to the field of EFD is the invention of the Pitot tube by Henri Pitot in the year 1732. The Pitot tube is capable of measuring the flow velocity at a point within the fluid and is still a commonplace equipment in any fluids laboratory. The study of turbulent flows is attributed to the famous experiments of Osborne Reynolds who studied the flow of fluids in a pipe and the effect of varying parameters on this flow by releasing a dye in the stream. His study on the flow variation with parameters led to the famous non-dimensional Reynolds number used to segregate a flow into laminar or turbulent. Another prominent example in the field of flow measurement is the experiments of Prandtl in 1904, where he studied the boundary layer as well as flow separation characteristics behind solid bodies in a flow using a water channel. Since the experimental attempts of Reynolds and Prandtl, technical advances in

the associated fields of optics, computers, electronics, image and signal processing, among others, has revolutionised measurement techniques for EFD. A few established techniques are discussed in successive sections.

### 1.2.2 Hot wire anemometry (HWA)

HWA was first introduced in the early 20th century before being commercialised mid-century and is one of the fundamental tools used for measuring point-wise velocity fields especially in turbulent flows. HWA calculates velocity based on the amount of cooling induced by the flow on a heated sensor maintained at a constant temperature by a feedback-control loop. The voltage drop across the sensor required to maintain equilibrium is used to measure the instantaneous velocity of the flow. As the velocity over the sensor changes, the corresponding convective heat transfer coefficient varies leading to a different voltage drop necessary to maintain equilibrium. Figure 1.3 depicts the electrical circuit diagram for HWA. The heat transfer from the sensor is not just a function of the flow velocity but also its geometry, orientation, the flow temperature, and concentration. Thus, it is possible to measure with HWA other fluid parameters such as concentration (Harion et al., 1996), temperature (Ndoye et al., 2010), as well as the individual components of the velocity field. Newer designs of probes are capable of measuring simultaneously all three velocity components. In addition, HWA's capability for high data-gathering rate of the order of several hundred kHz as well as the very small size of the sensor which minimises disturbance of the measured flow are distinct advantages of HWA over other intrusive flow measurement techniques. However, a stringent need for calibration, as well as the possibility of contamination/breakage due to fragility of the sensor limit application of HWA to all flows.

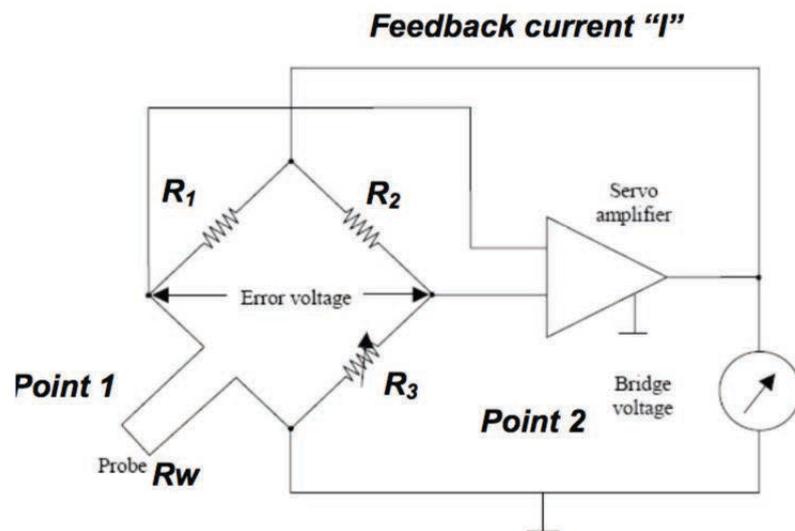


Figure 1.3: Circuit diagram for hot-wire anemometer.

### 1.2.3 Laser doppler anemometry (LDA)

The Doppler effect refers to the change in frequency or wavelength of a given wave signal when the relative velocity between the source and the observer varies. This shift is exploited in LDA to measure the velocity of a flow by illuminating it with a laser beam

which are reflected/scattered by particles seeded into the flow. The frequency difference between the original incident beam and the scattered beam, the Doppler shift, is linearly proportional to the particle velocity. The non-intrusive nature of LDA along with the high spatial and temporal resolution of the method and its ability to perform multi-component measurements are intrinsic advantages of LDA. However, its application is limited to areas with optical access, and single-point measurements. Seeding particles and seeding concentration are important factors which affect the performance of LDA. Depending on the optical arrangement of the Laser and the receiver, LDA can be classified as forward-scattering, side-scattering, or back-scattering LDA with the last being the most preferred option in modern set-ups. One such commercial set-up for an LDA experiment is shown in figure 1.4 - the fringe mode of LDA measurements is employed here by splitting the laser into two beams using a Bragg cell which intersect in the measurement zone. The frequency shift between the two intersection laser beams, induced due to their different incident angles, is used to calculate the Doppler frequency which is directly proportional to velocity.

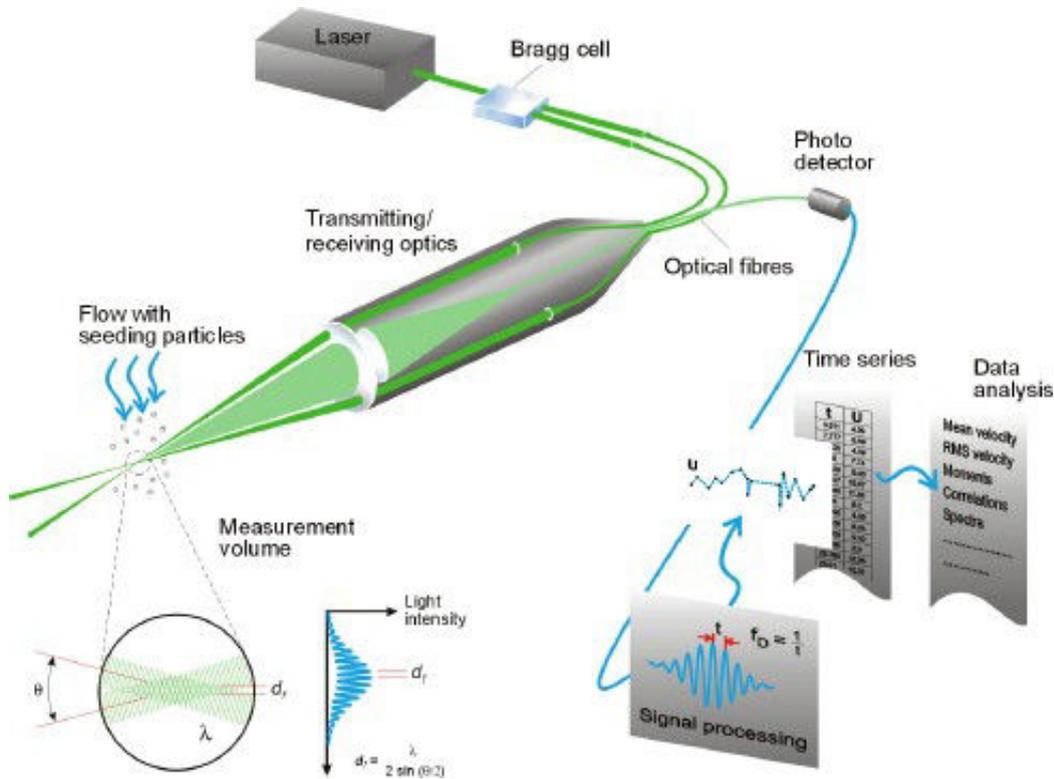


Figure 1.4: Laser Doppler Anemometer measurement set-up. Obtained from <sup>a</sup>

#### 1.2.4 Particle image velocimetry (PIV)

Both HWA and LDA are point-wise flow measurement techniques capable of measuring velocity with good temporal resolution at a given point in a flow. However, there is a strong need for full field (both 2D and 3D) velocity measurements - this was the motivation behind the development of PIV techniques. The development of PIV was done by Adrian (1984) under the name pulsed light velocimetry. This was followed by variants of the technique

<sup>a</sup><https://www.dantecdynamics.com/measurement-principles-of-lda>

proposed by Lourenco and Krothapalli (1986); Willert and Gharib (1991); Westerweel (1994) under different names although all these methods now fall in the PIV category. Since its development, PIV has become the dominant method for fluid flow measurement and remains till date the most popular method for turbulent flow measurement. The dominance of PIV can be attributed to the robustness of the method which arises from the simplicity of concept behind PIV. Adrian and Westerweel, two pioneers in the field of PIV, explain that the ability of PIV to directly measure displacement and time increment, which combined define the velocity field, provides robustness to the method. Adding to this, the ability of PIV to provide fields of velocity measurements and using particles as markers (omnipresence of particles and their ability to scatter more light compared to dye or molecules) are important characteristics. The art of particle imaging and processing can occupy entire textbooks, but for the purpose of this thesis, only a precise introduction to PIV principles and its extension to 3D flows (Tomographic PIV or recently 4D-PTV) is relevant and forms the core of this section.

Figure 1.5 is a schematic of basic planar PIV systems consisting of a double-pulsed laser, optics system, particle seeding, camera system, image processing software, and associated hardware for storage. A double-pulsed laser, generally a solid-state Nd:YAG laser, is used to produce two laser pulses to illuminate the flow which is seeded using the particle seeder with flow tracer particles, such as oil droplets in air or titanium dioxide in fluids. Newer methods have been developed which use helium filled soap bubbles as tracer particles (Schneiders et al., 2016) whose large size allow lasers of lower intensity to be used in the experiments. The reflected light is then captured using a camera-lens setup, digitised and then processed by an image processing software to improve quality, de-noising, etc. The next and an essential feature of PIV technique is the measurement of particle displacement between two consecutive images. This step, referred to as image or displacement interrogation, calculates from the image data particle displacements and thus flow velocity, using interrogation windows. A general method is the correlation technique where the image is split into interrogation windows of custom pixel size over which particle displacement correlation is performed using appropriate algorithms to identify the flow velocity for the window. This represents a large-scale view point of the flow and for highly turbulent flows, it fails to capture the small scale turbulent structures. A thorough description of PIV techniques can be found in any of the books dedicated to the subject.

Arroyo and Greated (1991) installed two cameras in their PIV set-up making it capable of measuring not just the two in-plane velocity components but also the out-of-plane component using the concept of stereoscopic imaging. Such a method, named stereoscopic PIV (SPIV), has certain advantages as it eliminates perspective error of monoscopic imaging, provides the full velocity vector field, and requires only an additional camera and corresponding calibration. Of particular interest to experimentalists is the ability to measure the full velocity vector over a volumetric domain. Different PIV techniques are capable of achieving this: Swept-Beam PIV (Gray et al., 1991) sweeps the planar light sheet over the volumetric domain, however, the measured field is not instantaneous. Photogrammetric PTV (Kent et al., 1993) uses multiple cameras to determine the 3D position of particles over two or more consecutive snapshots but is limited to large-sized tracer particles and associated flows. Holographic PIV (Fabry, 1998; Soria and Atkinson, 2008) uses principles of holographic recording and reconstruction to obtain volumetric measurements using SPIV like correlation algorithms but the method is limited to very-small volumes due to high-resolution requirements, Tomographic PIV (Elsinga et al., 2006) uses photogrammetric PTV like set of cameras (minimum 4 cameras) to record double-frame images and mathematically reconstruct the volume by projecting the image pixel gray-levels back into

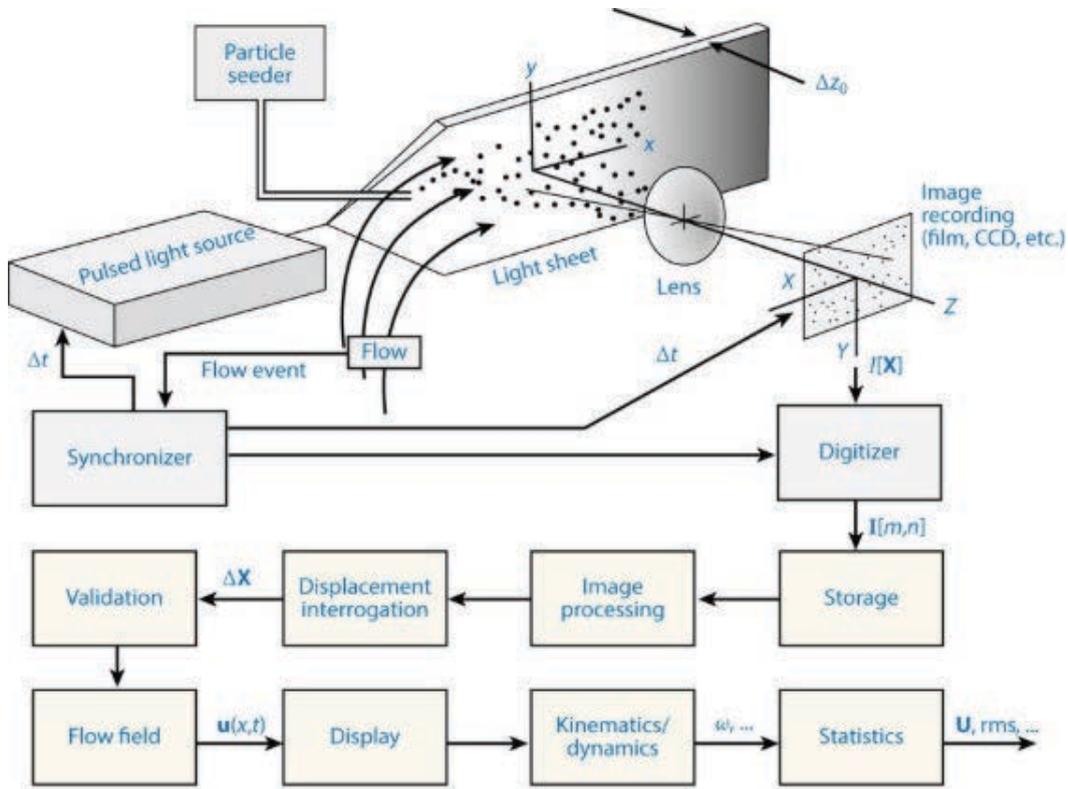


Figure 1.5: 2D Planar PIV system schematic representation (Westerweel et al., 2013).

the fluid and combing projections as and when they intersect. This is followed by a three-dimensional correlation analysis to interrogate the volume. Tomo-PIV is limited due to high computational costs and limited volumetric size but has been increasingly applied to measure the three components of flow velocity within a 3D volume (3D-3C) producing rewarding results (Hain et al., 2008; Kim and Chung, 2004; Weinkauff et al., 2013). The use of Helium-Filled Soap Bubbles (HFSB) as tracer particles has enabled a considerable increase in the measured volume size to a cubic-meter. The large size of these particles leads to increased light scattering and sparse seeding enabling large-scale measurements. Other innovative 3D measurement and reconstruction techniques include Tomo-PTV (Schneiders and Scarano, 2016), Shake The Box (Schanz et al., 2016), Holographic PTV, Vortex-in-cell (VIC) (Schneiders et al., 2014), etc.

### 1.3 Limitations

Both CFD and EFD have been extensively used to study fluid flows over the past century. However, there are associated drawbacks with each method which affect the accuracy of flow representation. CFD simulates a numerical approximation of the mathematical equations describing the flow. When the exact governing equations are known, an accurate approximation is in principle possible. However, for many cases the exact equations are unknown necessitating the introduction of a model introducing model errors. Even when exact equations are available, models are still used to reduce computational cost (turbulence modelling), which introduces model errors. Discretisation errors are introduced by the discretisation methods employed (FDM/FEM/FVM) in the form of truncation and

round-up errors. These can be minimised by using more complex schemes but at the cost of increased computational time. The use of iterative methods could also lead to numerical errors if sufficient iterations are not performed for convergence. Errors are also introduced due to inaccurate descriptions of the boundary conditions and initial conditions which are complex for realistic turbulent flows. Even for simple turbulent flows such as channel flow or wake flow around a cylinder, an accurate, realistic turbulent inlet/initial condition is difficult to envisage. Errors increase due to approximations or simplifications used to describe these conditions. Clearly, all these factors need to be taken into account when simulating any fluid flow. A compromise is generally required between computational cost and accuracy.

EFD methods are capable of capturing realistic flows which are, as of now, out of reach of CFD simulations. However, there are major limitations when it comes to experimental measurement techniques. Point-wise measurement techniques are capable of measuring accurate time-resolved fields but are severely limited in spatial extent. Even for 2D measurement techniques such as PIV, the measurement size is limited due to optical restrictions. The lack of sophisticated techniques to measure fluid properties over a large volumetric domain remains a major limitation of EFD techniques. Existing 3D measurement techniques are limited either to small volumetric domains or to non-instantaneous profiles (as obtained from swept-beam PIV) due to prohibitive costs for working the laser. The spatial resolution of measurements for PIV, SPIV and Tomo-PIV are all limited to large-scales especially for turbulent flows. Apart from the spatial limitations, the inability of EFD methods to simultaneously measure multiple properties of the fluid (velocity, pressure, temperature, density, etc.) is a major drawback.

## 1.4 Coupling EFD and CFD

An analysis of the limitation and abilities of CFD and EFD indicates a certain amount of complementarity between them. CFD is limited by the accuracy of its inlet and initial conditions, while EFD is capable of measuring an accurate initial and inlet conditions. EFD is limited in spatial extent while a large domain can be easily simulated using CFD. EFD is capable of measuring accurate but sparse and selective flow field properties while CFD is capable of measuring an approximated but complete flow-field properties over a large domain. The exploitation of this complementarity is a classic tradition in the field of fluid mechanics where experimental measurements are used to validate CFD results. Recently, this complementary has attracted more attention for possible coupling of EFD and CFD by using a dynamical model guided by partial or sparse experimental observations. Such a coupling, referred to as Data Assimilation, within the context of fluid mechanics is the primary focus of this thesis. A brief introduction to DA is provided here while a technical discussion can be found in chapter 5.

DA, as a field, has been driven by the works of researchers from the atmospheric and oceanographic sciences as well as contribution from geosciences. The need for DA was realised well before computers and numerical solvers were invented, in the year 1913 by Lewis Fry Richardson. He was working in the field of weather prediction and aimed at representing the physical principles governing the atmosphere as a set of mathematical equations which could be solved numerically. Richardson solved these equations manually with the available sparse observation set to disastrous results (Lynch, 2008). He realised that his erroneous results were due to an incorrect initial condition and concluded that a cure would be to initialise the numerical method with more realistic and complete set of observations, i.e. data assimilation. The sensitivity of atmospheric systems to the

initial condition is an important backdrop to the origination of DA techniques within the field of atmospheric sciences. With the invention of computers, a deeper understanding of physical principles governing the system, and more stable algorithms, the field of DA has come forward with leaps and bounds.

The current version of DA procedures can be broadly categorised into two main categories. Methods which are derived from stochastic filtering principles fall under the category of sequential data assimilation (SDA) approaches. The particle filter method (Gordon et al., 1993), the Ensemble Kalman Filter (EnKF) (Evensen, 1994), or a combination of both methods (Papadakis et al., 2010) are prominent examples of this method. These methods are generally based on the principle of Bayesian minimum variance estimation. They are termed as ‘sequential’ due to the constant forward propagation of the system state. As and when observations are encountered, they are assimilated into the system to correct the predicted state.

The second category of DA procedures are referred to as variational data assimilation (VDA) approaches and these originate from concepts of optimal control theory and variation calculus thus deriving the name (Lions, 1971). These methods aim at estimating the optimal trajectory, starting from a background condition, which minimises a cost function leading to lowest error between system and observations. The works of Bergthórsson and Döös (1955); Cressman (1959) on Optimal Interpolation methods were stepping stones to the VDA methods. With VDA, the approaches can be declassified into 3D variational and 4D variational methods depending on the spatial dimensions of the simulation and the inclusion or not of a temporal window for the system’s dynamical evolution. A first application of such methods was done by Sasaki (1958) who further extended them to 4D analysis in Sasaki (1970). Since then, variants of the VDA approach has been used for DA with Le Dimet and Talagrand (1986); Gronsksis et al. (2013) being some prominent examples.

While SDA and VDA methods are inherently based on different principles, an equivalence between the two can be drawn for specific cases as shown by Li and Navon (2001); Yang (2014). In adherence to the essence of DA towards combining complementary methods, the work of Yang et al. (2015) combines the ensemble Kalman filter approach with the 4D-Var method to provide a hybrid methodology. This is built on the publication of Hamill and Snyder (2000) who combined an ensemble based background covariance with 3D variational methods and Lorenc (2003) who expanded it to 4D. A different set of hybrid techniques which are based on SDA approaches incorporating variational formalisms were developed and analysed by Hunt et al. (2004, 2007); Fertig et al. (2007).

The various methodologies developed over the last decade for studying fluid flows in the fields of EFD, CFD and DA are graphically shown in figure 1.6. A stagnation can be seen in the development of methodologies which are restricted to a single field of focus namely EFD or CFD. 3D-PIV for EFD and stochastic NS/LES for CFD are the only major developments in the respective fields over the past decade. The data assimilation domain, however, has clearly attracted major research interest with increasing number of publications. Even within experimental methodologies, a shift towards combinatory techniques can be seen with techniques such as VIC (Schneiders et al., 2014), VIC+ (Schneiders et al., 2016), FlowFit Gesemann et al. (2016), etc.

In this present work, the VDA approach developed by Gronsksis et al. (2013) for 2D DNS of cylinder wake flow and extended to 3D DNS of wake flows by Robinson (2015) (termed henceforth as 4D-Var) is expanded to account for an LES turbulence model. Such an approach should be capable of performing VDA on flows of higher Reynolds numbers

which were previously limited by high computational costs. This 4D-Var<sub>les</sub> approach is used to optimise a three dimensional three component velocity field for wake flow at  $Re = 3900$  while assimilating observations. In the same vein as the preceding contributions, the reconstruction is performed by controlling the initial and inflow conditions. However, an additional control is also introduced for the LES on the turbulence model parameter.

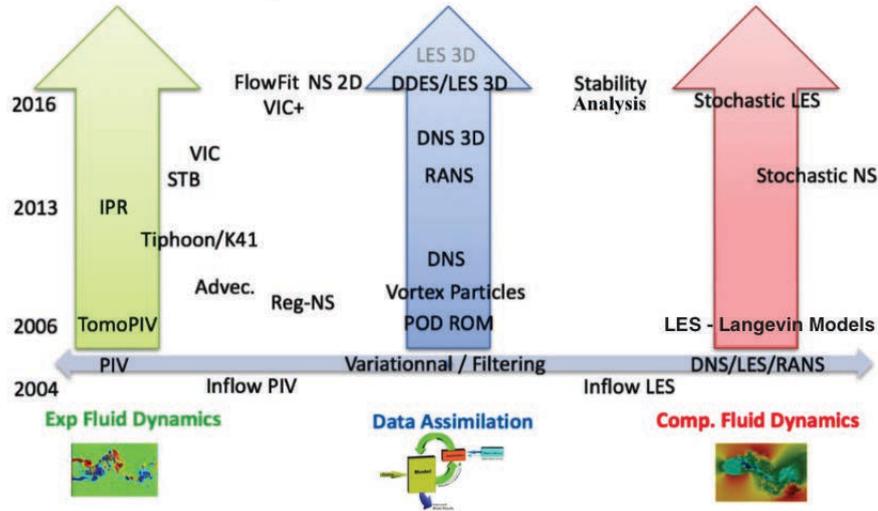


Figure 1.6: Different methodologies developed for fluid flow study over the past decade categorised into EFD, CFD and DA (Heitz, 2017)



## Chapter 2

# Turbulence modelling

The concepts of turbulence modelling are an essential part of CFD simulations as they enable simulations of highly turbulent flows which would otherwise be computationally too expensive to simulate. This section provides a detailed overview on the numerical modelling aspects of turbulence modelling with only concise mathematical representations. For a detailed description and derivation of turbulence principles, please refer to Rebollo and Lewandowski (2014); Lesieur (2008). Two different approaches have been developed for turbulence modelling - the statistical approach built on the theories of Osborne Reynolds, Prandtl, Kolmogorov, etc. and the more recently developed deterministic approach based on coherent structures and low-dimension dynamical systems. The deterministic approach is still in its infancy and is not presented in this thesis.

The classification of turbulence models is best represented by figure 2.1. Kolmogorov's theory on the energy cascade laid the foundation for the development of many classical models which are still in use. He postulated that for very high Reynolds number flows, the small scales of turbulence are statistically isotropic unlike the large scales. He additionally postulated that the statistics of these small scales of turbulence can be represented through the kinematic viscosity  $\nu$  and the rate of energy dissipation  $\epsilon$ . Figure 2.1 represents this energy cascade where injection of energy occurs in the large scale eddies which then cascades towards the smaller scales where it is finally dissipated. Based on the size range of eddies resolved, turbulent simulations can be classified as direct numerical simulations, large eddy simulations, or Reynolds averaged Navier-Stokes equations in decreasing order of resolved turbulent scales.

DNS resolves all the scales of turbulent motions from the characteristic length scale eddies down to the Kolmogorov scale. Thus, in this case, no extraneous model is required as all scales are accounted for. DNS is the simplest approach towards turbulence modelling (despite the lack of a model, DNS is considered here under turbulence modelling to simply the categorisation) and is the most accurate method for turbulent simulations provided appropriate boundary and initial conditions are provided describing the flow accurately. DNS is capable of providing maximum information regarding the flow as all scales are resolved and thus the level of detail of a DNS is unrivalled by any other approach. However, such detail comes at the price of restrictive computational cost. Consider a flow with length scales ranging from the integral length scale  $L$  to the Kolmogorov length scale  $\eta$ . As the Reynolds number is increased, the flow becomes more turbulent and to maintain equilibrium with energy production rate, smaller scales are necessarily generated until viscosity overtakes inertial forces leading to dissipation. Thus, these length scales are dependent only on the kinetic energy production rate or the dissipation rate ( $\epsilon$ ) (for a flow in equilibrium) and viscosity ( $\nu$ ). Using dimensional analysis, we can then infer that the

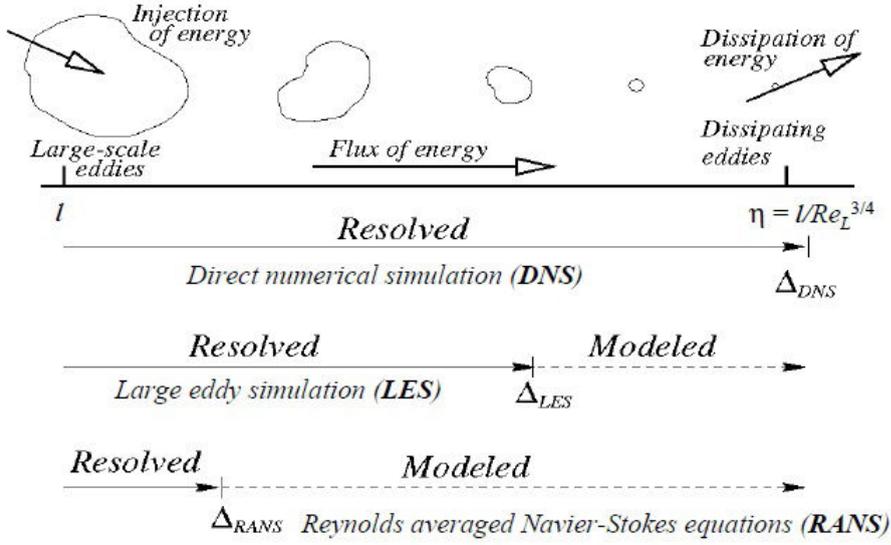


Figure 2.1: Classification of turbulence models based on cost and accuracy.

smallest scale that needs to be resolved is:

$$\eta = \left( \frac{\nu^3}{\epsilon} \right)^{\frac{1}{4}} \quad (2.0.1)$$

Approximating the dissipation rate as  $\frac{U^3}{L}$ , where  $U$  is the characteristic velocity scale and  $L$  is the integral length scale, the relation between the smallest and largest length scale in a flow can be given as a function of the integral length scale based Reynolds number ( $Re_L$ ) as:

$$\frac{L}{\eta} \sim \left( \frac{UL}{\nu} \right)^{\frac{3}{4}} = Re_L^{\frac{3}{4}} \quad (2.0.2)$$

Thus, to represent all the scales of motion, one needs a mesh increment ( $h$ ) smaller than  $\eta$  with  $Nh \geq L$ , where  $N$  is the number of mesh-points in one direction. For a 3D flow, this becomes:

$$N \geq Re_L^{\frac{9}{4}} \quad (2.0.3)$$

Similarly, to capture all time-scales in a 3D turbulent flow, one needs a time-step ( $dt$ ) smaller than  $t_\eta$  with  $N_t dt \geq t_L$  where  $N_t$  is the number of time steps, giving:

$$N_t \geq \frac{T}{t_\eta} = \frac{LU}{U\eta} = Re_L^{\frac{3}{4}} \quad (2.0.4)$$

An estimate of computation cost can be obtained as

$$N \times N_t \sim Re_L^{\frac{9}{4}} \times Re_L^{\frac{3}{4}} = Re_L^3 \quad (2.0.5)$$

Clearly, it is not feasible to perform DNS for high Reynolds numbers with current computational resources. DNS simulations are thus restricted to simplistic flows at low

Reynolds number. To simulate higher Re flows, it is necessary to move towards turbulence modelling concepts such as LES or RANS which are explored now in detail starting with the RANS models.

## 2.1 Reynolds averaged Navier-Stokes

The RANS models are not the focus of this thesis work but they are an important cog in the wheel of turbulence modelling alongside LES and are thus explored here in detail. Before delving in to the RANS models, the mathematical framework for the models is derived.

### 2.1.1 RANS equations

The RANS equations are named after Reynolds Osborne who first derived them in 1894 by decomposing the velocity field into its mean ( $\langle \mathbf{u} \rangle$ ) and fluctuating ( $\mathbf{u}'$ ) component. Within this section,  $\langle \rangle$  stands for mean quantities, and  $\prime$  stands for fluctuating quantities. Termed as Reynolds decomposition, such a decomposition when applied to the incompressible continuity equation gives:

$$\nabla \cdot \mathbf{u} = \nabla \cdot (\langle \mathbf{u} \rangle + \mathbf{u}') = 0, \quad (2.1.1)$$

given that  $\nabla \cdot \langle \mathbf{u} \rangle = 0$ , we get:

$$\nabla \cdot (\mathbf{u}') = 0. \quad (2.1.2)$$

Similarly, taking the mean of the NS equation (hence, Reynolds averaged NS or RANS), we get:

$$\frac{\partial \langle u_i \rangle}{\partial t} + \langle \mathbf{u} \rangle \cdot \nabla \langle u_i \rangle = -\frac{1}{\rho} \frac{\partial \langle p \rangle}{\partial x_i} + \nu \Delta \langle u_i \rangle - \frac{\partial \langle u'_j u'_i \rangle}{\partial x_j}, \quad (2.1.3)$$

where  $\langle p \rangle$  stands for the mean pressure field.

The RANS equations are similar to NS equation except for the additional term  $\frac{\partial \langle u'_j u'_i \rangle}{\partial x_j}$ . This term arises due to the averaging procedure applied on the non-linearity term:

$$\langle u_i u_j \rangle = \langle (\langle u_i \rangle + u'_i)(\langle u_j \rangle + u'_j) \rangle = \langle u_i \rangle \langle u_j \rangle + \langle u'_j u'_i \rangle \quad (2.1.4)$$

The RANS equation can be rewritten as,

$$\rho \left( \frac{\partial \langle u_i \rangle}{\partial t} + \langle \mathbf{u} \rangle \cdot \nabla \langle u_i \rangle \right) = \frac{\partial}{\partial x_i} \left[ -\langle p \rangle \delta_{ij} + \mu \left( \frac{\partial \langle u_i \rangle}{\partial x_j} + \frac{\partial \langle u_j \rangle}{\partial x_i} \right) - \rho \langle u'_j u'_i \rangle \right], \quad (2.1.5)$$

where  $\delta_{ij}$  is the Kronecker delta function. The three terms within the square brackets represent the isotropic stress from the pressure field, the viscous stress, and the apparent stress (also called Reynolds stresses) arising from the fluctuation field. It is now common-practice to refer to  $\langle u_i u_j \rangle$  as Reynolds stresses for simplicity instead of  $\rho \langle u'_j u'_i \rangle$ . Eq. (2.1.5) for the three velocity components combined with the mean continuity equation gives four independent equations for more than four unknowns - we have the three components of velocity, pressure and the Reynolds stresses creating a closure problem. The concept of RANS modelling revolves around the characterisation of these Reynolds stresses through appropriate models and/or equations, thus closing the set of equations.

The Reynolds stresses can be determined through a turbulence model based on the turbulent-viscosity hypothesis (first attributed to Boussinesq and hence referred to also as Boussinesq viscosity assumption (Boussinesq, 1877)), or by solving a set of transport equation, or using probability distribution function (PDF) methods. In the following, turbulent-viscosity based models are explored in detail while the other two types are described briefly. This chapter is not an exhaustive study of all the turbulence models but a concise description of the framework and some established models in each category.

### 2.1.2 Turbulent-viscosity models

The turbulent or eddy viscosity hypothesis by Boussinesq (Boussinesq, 1877) postulates that the momentum transfer due to the small scales can be modelled using an eddy viscosity. This is based on the assumption that turbulence dissipation has exactly the same form as the molecular dissipation. With this analogy, the Reynolds stress tensor can be modelled as:

$$\langle u'_i u'_j \rangle = \frac{2}{3} k \delta_{ij} - \nu_t \left( \frac{\partial \langle u_i \rangle}{\partial x_j} + \frac{\partial \langle u_j \rangle}{\partial x_i} \right), \quad (2.1.6)$$

where  $k$  is the turbulent kinetic energy,  $\frac{2}{3} k \delta_{ij}$  stands for isotropic part of the Reynolds stresses, and  $\nu_t$  is the so called turbulence or eddy viscosity. Correspondingly, the deviatoric anisotropic part of the stresses ( $A_{ij}$ ) can be expressed as:

$$A_{ij} = -2\nu_t \langle S_{ij} \rangle, \quad (2.1.7)$$

where the mean rate-of-strain  $\langle S_{ij} \rangle$  is given as,

$$\langle S_{ij} \rangle = \frac{1}{2} \left( \frac{\partial \langle u_i \rangle}{\partial x_j} + \frac{\partial \langle u_j \rangle}{\partial x_i} \right). \quad (2.1.8)$$

Thus to close these set of equations, only the eddy viscosity  $\nu_t$  needs to be defined through an appropriate model.

For simple turbulent shear flows, where the rate of change of mean velocity gradients and turbulence characteristics is smooth, such a hypothesis is assumed to be reasonable but for more complex flows such as axisymmetric contraction, swirling flows, and non-circular ducts, the hypothesis is shown to fail (Pope, 2000). Despite this shortcoming, the models based on this hypothesis remain one of the most popular set used for turbulence modelling and have been shown to provide good results for the mean quantity while the accuracy of the second order moments are questionable.

#### Algebraic models

The first model is the uniform-turbulent viscosity model given as,

$$\nu_t(x, y) = \frac{U(x)\delta(x)}{R_t} \quad (2.1.9)$$

where  $U$  and  $\delta$  are the characteristic velocity and length scale,  $R_t$  is a flow dependent constant, and  $x, y$  define the spatial location. The eddy viscosity is assumed to vary only in the direction of flow. A pre-defined direction of flow, as well as the characteristic velocity and length scales are compulsory inputs for this model. Such inputs can be obtained only for the simplest of flows limiting the application of the uniform-turbulent viscosity model.

The mixing-length model is the second and final algebraic model which, for a thin shear layer, is defined as,

$$\nu_t = l_m^2 \left| \frac{\partial \langle \mathbf{u} \rangle}{\partial y} \right|, \quad (2.1.10)$$

where  $y$  is the transverse height, and  $l_m$  stands for the mixing length which in the words of Prandtl can be described as, ‘the mixing length may be considered as the diameter of the masses of fluid moving as a whole in each individual case; or again, as the distance traversed by a mass of this type before it becomes blended in with neighbouring masses...’ (Prandtl, 1925). The mixing-length model can be easily applied to a multitude of turbulent flows provided the mixing length is known a-priori. For many important fluid flows,  $l_m$  has already been studied and the specifications are well-established making this the simplest model with a moderate range of applicability. The mixing-length hypothesis has been used to derive more complex models such as the Smagorinsky model for LES (Smagorinsky, 1963), and Baldwin-Lomax’s model (Baldwin and Lomax, 1978).

### One equation models

Based on the independent suggestions of Kolmogorov (1942) and Prandtl (1945), the velocity scale was specified as,

$$U = ck^{\frac{1}{2}}, \quad (2.1.11)$$

giving an eddy viscosity, in combination with the mixing-length hypothesis, as,

$$\nu_t = U\delta = ck^{\frac{1}{2}}l_m, \quad (2.1.12)$$

where  $k$  is the turbulent kinetic energy (tke), and  $c$  is a constant. Thus to specify the eddy viscosity, one needs to specify  $k$ . In the one-equation model of Prandtl, this is done by solving a model transport equation for  $k$  given as,

$$\frac{\partial k}{\partial t} + \langle \mathbf{u} \rangle \cdot \nabla k = -\nabla \cdot \mathbf{T}'_k + \mathcal{P}_k - \epsilon, \quad (2.1.13)$$

where  $\mathbf{T}'$ ,  $\mathcal{P}_k$ , and  $\epsilon$  stand for tke flux, tke production and tke dissipation respectively. The flux is given as,

$$T'_i = \frac{1}{2} \langle u_i u_j u_j \rangle + \frac{\langle u_i p' \rangle}{\rho} - 2\nu \langle u_j S'_{ij} \rangle, \quad (2.1.14)$$

where  $S'_{ij}$  is the strain rate tensor and is defined in further sections. The flux is modelled using a gradient diffusion hypothesis as,

$$\mathbf{T}' = -\frac{\nu_t}{\sigma_k} \nabla k, \quad (2.1.15)$$

where  $\sigma_k = 1.0$  is the turbulent Prandtl number. The production  $\mathcal{P}_k$  can be defined from known variables using Boussinesq’s viscosity hypothesis as,

$$\mathcal{P}_k = -\overline{\rho u'_i u'_j} \frac{\partial \bar{u}_i}{\partial x_j} \quad (2.1.16)$$

$$= \rho \nu_t \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right) \frac{\partial \bar{u}_i}{\partial x_j} \quad (2.1.17)$$

and the dissipation  $\epsilon$  can be modelled as,

$$\epsilon = \frac{C_D k^{\frac{3}{2}}}{l_m} \quad (2.1.18)$$

where  $C_D$  is a model constant.

An alternative one-equation model was suggested by Spalart and Allmaras (1992) specially for aerodynamic flows, where-in a model transport equation is solved directly for an eddy-viscosity like variable called the Spalart-Allmaras variable. The model transport equation for the variable is quite complex and is not reproduced in this next. For the mathematical equations, please refer to the original paper (Spalart and Allmaras, 1992). The Spalart-Allmaras model is shown to work well for aerodynamic flows but it is limited in use as a generic model for all flows.

### Two equation models

Adding a further level of intricacy to the one equation models, the two equation models solve transport equation for two variables of which one is generally  $k$ . The  $k - \epsilon$  model, credited to Jones and Launder (1972), defines the eddy viscosity as,

$$\nu_t = C_\mu \frac{k^2}{\epsilon}, \quad (2.1.19)$$

where  $C_\mu$  is one of five model constants. This model solves model transport equations for the (eq. (2.1.13)) and dissipation ( $\epsilon$ ) with the empirical transport equation for  $\epsilon$  written as,

$$\frac{\partial \epsilon}{\partial t} + \langle \mathbf{u} \rangle \cdot \nabla \epsilon = -\nabla \cdot \left( \frac{\nu_t}{\sigma_\epsilon} \nabla \epsilon \right) + C_{\epsilon 1} \frac{P_\epsilon}{k} - C_{\epsilon 2} \frac{\epsilon^2}{k}. \quad (2.1.20)$$

where  $P_\epsilon$  is the production term and the model constants are given by Launder and Sharma (1974) as,

$$C_\mu = 0.09, \quad C_{\epsilon 1} = 1.44, \quad C_{\epsilon 2} = 1.92, \quad \sigma_k = 1.0, \quad \sigma_\epsilon = 1.3. \quad (2.1.21)$$

The  $k - \epsilon$  model is considered as one of the simplest-complete turbulence model and applicable to a wide variety of turbulent flows. Since its conception, numerous variants of the model have been implemented and studied. Yang and Shih (1993) developed a variant for wall-bounded turbulent flows, while Shih et al. (1998) extended the model to a cubic stress-strain relationship, i.e. non-linear, applicable to shear flows. Humphrey and Pourahmadi (1983) used algebraic approximation to better estimate  $C_\mu$  applied to curved channel flow. A commonly used variant is the Re-normalisation Group (RNG)  $k - \epsilon$  model, developed by Yakhot et al. (1992), with a revised production term and correspondingly adjusted constants which has been applied with mixed success to rotational flows but is now the favoured model for indoor air simulations.

The  $k - \omega$  model, developed by Wilcox (1988), is another established RANS models found in most commercial codes alongside the  $k - \epsilon$  model. Two transport equations are solved for the tke  $k$  and the specific dissipation  $\omega$ . The transport equation for  $\omega$  is written as,

$$\frac{\partial \omega}{\partial t} + \langle \mathbf{u} \rangle \cdot \nabla \omega = -\nabla \cdot \left( \frac{\nu_t}{\sigma_\omega} \nabla \omega \right) + C_{\omega 1} \frac{P_\omega}{k} - C_{\omega 2} \omega^2. \quad (2.1.22)$$

where  $P_\omega$  is the production term, and  $C_{\omega 1}$ ,  $C_{\omega 2}$ , and  $\sigma_\omega$  are model constants.

This equation can be mathematically shown to be identical to the transport equation for  $\epsilon$  under the assumption of homogenous turbulence and fixed constants. However, for inhomogeneous turbulence the equality fails (Pope, 2000). Wilcox (1993) performed a comparative study of the different two-equation models and concluded that the  $k - \omega$  model improves on the  $k - \epsilon$  model within the viscous near-wall region and in treatment of stream-wise pressure-gradients for boundary-layer flows but fails at the free-stream boundary. To

off-set this error, Menter (1994) developed the baseline  $k - \omega$  model which blends the  $k - \omega$  model near surfaces with the  $k - \epsilon$  model in free-stream flow. Another version, also proposed by Menter, involves a limiting function on the eddy viscosity with the model blend and this is now recognised as the shear stress transport (SST)  $k - \omega$  model (Menter, 1993).

### 2.1.3 Reynolds stress models (RSM)

The validity of all the models described previously are strictly dependent on the validity of the Boussinesq turbulent viscosity hypothesis. During the conception of the models, the reasonability of the hypothesis was assumed for simple shear flows. However, Schmitt (2007) evaluates numerically the validity of the hypothesis using numerical and experimental data-sets and concludes that the validity of the hypothesis is questionable even for simple shear flows. This clearly is a major drawback. An alternative approach to address this, advocated by Chou (1945); Rotta (1951), led to the development of the class of models referred to as the Reynolds Stress Models.

RSM provide an additional level of turbulence closure (second order closure) by solving transport equations for the Reynolds stresses and an additional variable (generally dissipation  $\epsilon$ ) thus removing the dependency on the turbulent-viscosity hypothesis. The transport equations for the stresses are given as,

$$\frac{\partial \langle u'_i u'_j \rangle}{\partial t} + \langle \mathbf{u} \rangle \cdot \nabla \langle u'_i u'_j \rangle + \frac{\partial}{\partial x_k} \mathbf{T}_{kij} = \mathcal{P}_{ij} + \mathcal{R}_{ij} - \epsilon_{ij}. \quad (2.1.23)$$

where  $\mathbf{T}_{kij}$  stands for the Reynolds-stress flux,  $\mathcal{P}_{ij}$  the production term for the Reynolds-stresses,  $\mathcal{R}_{ij}$  the pressure-rate-of-strain tensor, and  $\epsilon_{ij}$  the dissipation tensor which is different from dissipation  $\epsilon$ . The material derivative and the production term are in closed form while a closure model is required for the remaining terms. An isotropy based model for dissipation is easily constructed but the other two terms need to be explicitly modelled. A first attempt at modelling the velocity-pressure-gradient tensor which combines  $\mathbf{T}_{kij}$  and  $\mathcal{R}_{ij}$  was the Rotta model (Rotta, 1951). Other variants include the LRR model by Launder et al. (1975), and the SSG model by Speziale et al. (1991) among others. Each RANS model possesses certain advantages over other models but it is debatable on which model is the most accurate as their performance is flow dependent. A comparative performance analysis of the RANS models for different flows can be found in Martinuzzi and Pollard (1989); Pollard and Martinuzzi (1989); Harrison and Bogard (2008).

An independent approach to turbulence modelling is based on stochastic Lagrangian models and are derived from stochastic differential equations. Such methods, classified as ‘probability distribution function or PDF’ methods, are mentioned here for sake of completeness and are not discussed further. For more information, please refer to Pope (2000); Lesieur (2008).

## 2.2 Large eddy simulation

Large eddy simulations provide a computationally economical option to DNS for flows involving turbulence structures of minute spatial and temporal scales. LES modelling of turbulence is an established methodology in the fluid dynamics domain and is used to simulate a wide variety of flows. LES has been successfully applied to varied flows including canonical flows (Moser et al., 1999), atmospheric flows (Chemel et al., 2010), wake flows (Chandramouli et al., 2018), mixing layer (Lesieur et al., 1988), and reactive flows (Minier, 2015) to study flow statistics, turbulence structures and coherence, scalar transport and

diffusion (Kempf et al., 2005), particle entrainment (Chemel and Staquet, 2007), and flow interaction (for example in two-phase flows, wake flows, etc.) among others.

The mathematical formulation for the filtered NS equations and the corresponding closure problem are presented first followed by a brief description of established models and newer approaches, based on stochastic processes, are also highlighted.

### 2.2.1 Filtered Navier-Stokes equations

RANS models' distinctive advantage in modelling turbulent flows through reduced computational cost is off-set by the loss in accuracy of the simulation. This, among other reasons, led to researchers looking for an alternate approach to modelling turbulence. Similar to DA, the pioneering works on LES was performed by meteorological scientists motivated by the restrictive cost of simulations for weather prediction. The primary concept of LES is to resolve the large scales of turbulence numerically while the small (considered universal) scales are modelled by appropriate models. LES is capable of simulating unsteady, transient flow structures which is limited in RANS models, while also capable of reproducing turbulent structures more accurately compared to any RANS model. The principle behind LES lies in the splitting of variables into large resolvable scales (denoted by  $(\bar{\quad})$ ) and smaller universal scales (denoted by  $(\prime)$ ), also called sub-grid scales (SGS), using appropriate filtering options (it is to be noted that the fluctuation quantities of RANS equations and the filtered NS equations are theoretically different but for case of simplicity are both referred to as  $(\prime)$  - henceforth only LES models are considered and hence  $(\prime)$  refers to the SGS quantities unless specified otherwise). The filtered form of the continuity equation is,

$$\frac{\partial(\bar{u}_i)}{\partial x_i} = 0. \quad (2.2.1)$$

From the mass-conservation equation, it can be easily seen that for an incompressible flow both the resolved field and the SGS field are solenoidal provided the derivative and filtering operation commute.

$$\nabla \cdot (\bar{\mathbf{u}} + \mathbf{u}') = 0 \implies \nabla \cdot \mathbf{u}' = 0 \quad \text{if} \quad \overline{\nabla \cdot \mathbf{u}} = \nabla \cdot \bar{\mathbf{u}} = 0. \quad (2.2.2)$$

The filtered NS equation is given as,

$$\frac{\partial(\bar{u}_i)}{\partial t} + \frac{\partial(\bar{u}_i \bar{u}_j)}{\partial x_j} = -\frac{1}{\rho} \frac{\partial \bar{p}}{\partial x_i} + 2\nu \frac{\partial}{\partial x_j} \bar{S}_{ij} - \frac{\partial \tau_{ij}^r}{\partial x_j}, \quad (2.2.3)$$

where  $\bar{p}$  is the filtered pressure,  $\bar{S}_{ij}$  is the filtered rate-of-strain tensor and  $\tau_{ij}^r$  indicates the residual stress tensor which is similar to the Reynolds stress tensor and triggers the closure problem requiring modelling. This additional term arises due to filtering of the non-linear advection term which is filtered as,

$$\overline{u_i u_j} = \bar{u}_i \bar{u}_j + \tau_{ij}^r + \frac{2}{3} k_r \delta_{ij}. \quad (2.2.4)$$

The SGS stress component has been split into the isotropic  $\frac{2}{3} k_r \delta_{ij}$ , where  $k_r$  stands for the residual kinetic energy and  $\delta_{ij}$  stands for the kronecker delta function, and the anisotropic 'deviatoric'  $\tau_{ij}^r$  parts respectively. Such a decomposition is not unique in its nature and many other decompositions of the SGS terms have been proposed by Leonard (1975); Germano (1986) among others. Both these decompositions will be described and used in the section dealing with dynamical models.

Table 2.1: Filter function and corresponding transfer functions for one-dimension LES filters (Pope, 2000)

| Name           | Filter Function $G(r)$   | Transfer Function $\hat{G}(\kappa)$                             |
|----------------|--|---|
| General        | $G(r)$   | $\int_{-\infty}^{\infty} e^{i\kappa r} G(r) dr$                 |
| Box            | $\frac{1}{\Delta} H(\frac{1}{2}\Delta -  r )$  | $\frac{\sin(\frac{1}{2}\kappa\Delta)}{\frac{1}{2}\kappa\Delta}$ |
| Gaussian       | $\left(\frac{6}{\pi\Delta^2}\right)^{\frac{1}{2}} \exp\left(-\frac{6r^2}{\Delta^2}\right)$ | $\exp\left(-\frac{\kappa^2\Delta^2}{24}\right)$                 |
| Sharp Spectral | $\frac{\sin(\pi r/Delta)}{\pi r}$  | $H(\kappa_c -  \kappa ), \kappa_c \equiv \pi/\Delta$            |
| Cauchy         | $\frac{a}{\pi\Delta[(r/\Delta)^2+a^2]}, a = \frac{\pi i}{24}$                              | $\exp(-a\Delta \kappa )$  |

Before exploring the wide variety of LES SGS models, a brief interlude is taken focusing on the filtering approaches used in LES. The filtering operation is given as (Leonard, 1975),

$$\bar{\mathbf{u}}(\mathbf{x}, t) = \int G(\mathbf{r}, \mathbf{x}) \mathbf{u}(\mathbf{x} - \mathbf{r}, t) d\mathbf{r}, \quad (2.2.5)$$

where  $G$  is the filtering function employed over the entire domain. The filter width in the physical domain is denoted as  $\Delta$  and in the spectral domain as  $\pi/\Delta$ . The residual or SGS field is then defined as,

$$\mathbf{u}'(\mathbf{x}, t) = \mathbf{u}(\mathbf{x}, t) - \bar{\mathbf{u}}(\mathbf{x}, t). \quad (2.2.6)$$

A few common filter functions  $G$  in 1D are given in table 2.1. Applying such filter functions to a velocity field is termed as explicit filtering. However, it is quite common in LES studies to simply discretise the domain into a mesh which performs an implicit filtering of the velocity as all scales smaller than the smallest mesh size are implicitly filtered out. A comparison of accuracy and effectiveness of explicit filtering vs implicit filtering can be found in Lund and Kaltenbach (1995); Winckelmans et al. (2001); Lund (2003); Gullbrand and Chow (2003). The specifics about the common filter functions and their performance can be found in Pope (2000). Complex filter functions for non-uniform filter-widths were proposed by van der Ven (1995) with the special property of being able to commute with differentiation up to any order.

### 2.2.2 Smagorinsky models

A first approach to LES was undertaken by Smagorinsky (Smagorinsky, 1963) in which he proposed a model for simulating atmospheric flows. Smagorinsky's model and its variants has since been established as the favoured model for LES studies. However, in its basic version, the model has various limitations and the stringent efforts of numerous researchers has led to various modifications of the original model. In this section, the original formulation of Smagorinsky is provided followed by a description of some important modifications of the model with a focus on the dynamical version of the model.

### Classical version

Smagorinsky's original model was based on the linear eddy-viscosity model and the mixing-length hypothesis which also formed the basis for the turbulent-viscosity based RANS models. To reiterate, the linear eddy-viscosity model for the residual stress tensor is given by:

$$\tau_{ij}^r = -2\nu_t \bar{S}_{ij}, \quad \text{where} \quad \bar{S}_{ij} = \frac{1}{2} \left( \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \right), \quad (2.2.7)$$

the equation relates the residual (sub-grid scale) stress with the filtered strain rate ( $\bar{S}_{ij}$ ). The eddy viscosity ( $\nu_t$ ) can be modelled by an analogy with the mixing length hypothesis giving:

$$\nu_t = (C_s \Delta)^2 |\bar{S}|, \quad (2.2.8)$$

where  $C_s$  is the Smagorinsky coefficient which is assumed to be constant in space and time,  $\Delta$  is the filter width and  $|\bar{S}|$  is the filtered rate of strain norm given as  $|\bar{S}| = (2\bar{S}_{ij}\bar{S}_{ij})^{\frac{1}{2}}$  with an implicit summation of  $i$  and  $j$ . The Smagorinsky coefficient is flow dependent and fixing this coefficient is one of the main difficulties/limitations of the classical version of the model. A first attempt at estimating this coefficient was done by Lilly (1967) by equating energy production ( $\Pi$ ) with the dissipation.

$$\Pi = \epsilon_{RS} = - \langle \tau_{ij}^r \bar{S}_{ij} \rangle, \quad (2.2.9)$$

where  $\epsilon_{RS}$  stands for the dissipation at the resolved scales. Defining  $\tau_{ij}^r$  using Smagorinsky's model gives,

$$\epsilon_{RS} = 2(C_s * \Delta)^2 \langle |\bar{S}| \bar{S}_{ij} \bar{S}_{ij} \rangle, \quad (2.2.10)$$

$$\epsilon_{RS} \approx (C_s * \Delta)^2 2^{\frac{3}{2}} \langle (\bar{S}_{ij} \bar{S}_{ij})^{\frac{3}{2}} \rangle. \quad (2.2.11)$$

For an isotropic homogenous turbulence flow,  $\langle \bar{S}_{ij} \bar{S}_{ij} \rangle$  can be expressed as (Sagaut, 2006),

$$\langle 2\bar{S}_{ij} \bar{S}_{ij} \rangle = \int_0^{\frac{\pi}{\Delta}} 2\kappa^2 E(\kappa) d\kappa, \quad (2.2.12)$$

where  $\kappa$  is the wavenumber, and  $E(\kappa)$  is the energy spectrum. By using the universal form of the energy spectrum  $E(\kappa) = c_K \epsilon^{\frac{2}{3}} \kappa^{-\frac{5}{3}}$ , we get

$$\langle 2\bar{S}_{ij} \bar{S}_{ij} \rangle = \int_0^{\frac{\pi}{\Delta}} 2\kappa^2 c_K \epsilon^{\frac{2}{3}} \kappa^{-\frac{5}{3}} d\kappa, \quad (2.2.13)$$

$$= 2c_K \epsilon^{\frac{2}{3}} \int_0^{\frac{\pi}{\Delta}} \kappa^{\frac{1}{3}} d\kappa, \quad (2.2.14)$$

$$= c_K \epsilon^{\frac{2}{3}} \frac{3}{2} \left( \frac{\pi}{\Delta} \right)^{\frac{4}{3}}, \quad (2.2.15)$$

where  $c_K$  is the Kolmogorov constant. By equating the two dissipations ( $\epsilon_{RS} = \epsilon$ ), we obtain,

$$\epsilon \approx C_s^2 \Delta^2 \left( c_K \epsilon^{\frac{2}{3}} \frac{3}{4} \left( \frac{\pi}{\Delta} \right)^{\frac{4}{3}} \right)^{\frac{3}{2}} \quad (2.2.16)$$

$$C_s = \left( \frac{3c_K}{2} \right)^{\frac{-3}{4}} \pi^{-1} \quad (2.2.17)$$

For a  $c_K$  of 1.6, as given in literature, a coefficient value of 0.165 is obtained which is used to simulate channel flow. Such a value works well for isotropic, high Reynolds number flows but for non-isotropic flows, which most flows in reality are, a numerical approach is necessary to quantify the coefficient. For simple turbulent flows the coefficient has been numerically studied and an estimate has been provided. For example, Moin and Kim (1982) suggests a value of 0.065 for channel flow while Ouvrard et al. (2010) suggests 0.1 for wake flow around a cylinder.

While this classical version of the Smagorinsky model is one of the simplest models for Large eddy simulations, it has numerous drawbacks. The model assumes the existence of equilibrium between the kinetic energy flux across scale and the large scales of turbulence - this does not hold true for many cases. The Smagorinsky coefficient is assumed to be spatially and temporally constant - this is considered to be the most serious drawback of the model and is redressed in the dynamic Smagorinsky model. In addition, Smagorinsky model predicts unphysical non-zero eddy viscosity at the boundary walls which leads to re-laminarisation of the flow in numerous cases (Nicoud and Ducros, 1999). This near wall inaccuracy led to the development of Wall-Adapting Local Eddy-Viscosity (WALE) model by Nicoud and Ducros (1999). Both these models are commonly used in LES studies. Rebollo and Lewandowski (2014) study the stability and convergence of the steady Smagorinsky model using finite element approximation and conclude that the model displays a lack of optimality due to over-diffusive effects warranting wall-modelling.

### Dynamical version

Lilly suggests that the Smagorinsky coefficient should be constant and derives a value for this coefficient (Lilly, 1967). However, the study of Meyers and Sagaut (2006), and Germano et al. (1991) suggest that a constant value does not efficiently capture turbulence characteristics especially in regions around boundaries such as walls. Meyers and Sagaut (2006) state that “the Smagorinsky coefficient strongly depends on the local ratio of the filter width  $\Delta$  to the Kolmogorov scale  $\eta$ ”

In order to resolve this basic flaw within the Smagorinsky model, attempts were made to include damping functions such as Van-Driest Damping (Van Driest, 1956), or to use renormalisation group theory (Yakhot and Orszag, 1986). However, the work by Germano et al. (1991) provided a non ad-hoc manner for calculating the Smagorinsky coefficient which varies with both space and time. This work by Germano et al. (1991) is referred to as the Dynamic Smagorinsky model due to the dynamic nature of its varying coefficient. The mathematical framework for the model is given below.

The model locally estimates the Smagorinsky coefficient based on the smallest resolved turbulent scales in the flow. Two filters, namely grid ( $\bar{\cdot}$ ) and test ( $\tilde{\cdot}$ ), are used to identify the smallest resolved scales within the flow - the grid filter is the regular filter used for any LES model to define the resolved and sub grid scales. Consider the grid filtered Navier Stokes equation:

$$\frac{\partial \bar{u}_i}{\partial t} + \frac{\partial}{\partial x_j} (\bar{u}_i \bar{u}_j) = -\frac{\partial \bar{p}}{\partial x_i} - \frac{\partial \tau_{ij}^r}{\partial x_j} + \frac{1}{Re} \frac{\partial^2 \bar{u}_i}{\partial x_j \partial x_j}. \quad (2.2.18)$$

Applying to this grid filtered equation, the test filter (indicated by  $\tilde{\cdot}$ ) gives

$$\frac{\partial \tilde{u}_i}{\partial t} + \frac{\partial}{\partial x_j} (\tilde{u}_i \tilde{u}_j) = -\frac{\partial \tilde{p}}{\partial x_i} - \frac{\partial \tilde{\mathcal{T}}_{ij}^r}{\partial x_j} + \frac{1}{Re} \frac{\partial^2 \tilde{u}_i}{\partial x_j \partial x_j}. \quad (2.2.19)$$

The sub-grid scale stress contribution for the two cases are:

$$\tau_{ij}^r = \bar{u}_i \bar{u}_j - \bar{u}_i \bar{u}_j, \quad (2.2.20)$$

$$\mathcal{T}_{ij}^r = \widetilde{\overline{u_i u_j}} - \widetilde{u_i} \widetilde{u_j}. \quad (2.2.21)$$

Germano defined a residual turbulent stress ( $\mathcal{L}_{ij}$ ) which is the contribution of the smallest resolved scales of motion to the Reynolds stress.

$$\mathcal{L}_{ij} = \widetilde{\overline{u_i u_j}} - \widetilde{u_i} \widetilde{u_j}. \quad (2.2.22)$$

The three stress variables each representing the contribution to the Reynolds stress of various scales of turbulence can be algebraically related by the Germano identity (Germano et al., 1991),

$$\mathcal{L}_{ij} = \mathcal{T}_{ij}^r - \widetilde{\tau}_{ij}^r. \quad (2.2.23)$$

The two sub-grid scale stress contribution can be parametrised using a SGS model. Considering the Smagorinsky model gives,

$$\mathcal{L}_{ij} - \delta_{ij} \frac{\mathcal{L}_{kk}}{3} = -2C_s M_{ij}, \quad (2.2.24)$$

where,

$$M_{ij} = \widetilde{\Delta}^2 |\widetilde{\mathcal{S}}| \widetilde{\mathcal{S}}_{ij} - \widetilde{\Delta}^2 (|\widetilde{\mathcal{S}}| \widetilde{\mathcal{S}}_{ij}), \quad (2.2.25)$$

with  $\widetilde{\Delta}$  and  $\bar{\Delta}$  denoting the test and grid filter widths respectively.

Thus we get an expression for  $\mathcal{L}_{ij}$  in terms of the Smagorinsky coefficient.

$$\mathcal{L}_{ij} \bar{\mathcal{S}}_{ij} = -2C_s (\widetilde{\Delta}^2 |\widetilde{\mathcal{S}}| \widetilde{\mathcal{S}}_{ij} \bar{\mathcal{S}}_{ij} - \widetilde{\Delta}^2 (|\widetilde{\mathcal{S}}| \widetilde{\mathcal{S}}_{ij}) \bar{\mathcal{S}}_{ij}). \quad (2.2.26)$$

From the above relation, the Smagorinsky coefficient can be dynamically calculated as,

$$C_s^2 = -\frac{1}{2} \frac{\mathcal{L}_{ij} \bar{\mathcal{S}}_{ij}}{M_{kl} \bar{\mathcal{S}}_{kl}}. \quad (2.2.27)$$

An alternate least squares based equation for the coefficient was proposed by Lilly (1992) as,

$$C_s^2 = -\frac{1}{2} \frac{\mathcal{L}_{ij} M_{ij}}{M_{kl} M_{kl}}. \quad (2.2.28)$$

Both these procedure provides a dynamically updated coefficient at each time step for each point within the domain. The dynamic Smagorinsky model successfully overcomes the issue associated with the use of constant value for the coefficient - it accurately predicts zero SGS stress at walls. In addition, it portrays the correct asymptotic behaviour in the near wall region while it is capable of taking into account back-scatter as well.

However, for certain situations the denominator can equate to zero giving unphysical value for the Smagorinsky coefficient - it was indeed observed that a large number of negative coefficients were also obtained using this method giving rise to problematic anti-diffusion or energy injection. In order to avoid this, an averaging procedure in statistically homogeneous planes was suggested by Germano - this procedure, while easily implemented for channel flow and other homogenous flows, is not a viable solution for complex geometries. Novel approaches to stabilise the model have been proposed such as the dynamic localisation model by Ghosal et al. (1995) and the Lagrangian dynamic model by Meneveau et al. (1996). However, it is important to note that using such procedures leads to a non-trivial increase in computational cost and must be taken into account (You and Moin, 2006).

### 2.2.3 WALE model

One of the unsaid requirements of the dynamic Smagorinsky model is the need to perform averaging over statistically homogenous planes. Ignoring this averaging procedure leads to zero or negative Smagorinsky coefficient leading to numerical instabilities. The dynamical nature of the model further increases complexity and computational effort needed for performing LES. Additionally, all Smagorinsky based models assume a dependance of eddy viscosity on just the irrotational strain rate ( $\bar{S}_{ij}$ ) - this strain rate operator is only of order  $O(1)$  at the wall while the physical behaviour should be asymptotic. The operator is expected to go to zero at the wall, however the second invariant of  $\bar{S}$  is  $-(1/2)\bar{S}_{ij}\bar{S}_{ij}$  which is a modified form of the operator used in Smagorinsky. Moreover, Wray and Hunt (1990) suggest that the eddy viscosity should show a dependence not only on the irrotational strain but also its rotational component. In the words of Nicoud and Ducros (1999), "turbulent energy is concentrated in the streams and energy dissipation in eddies and convergence zones".

The presence of these difficulties, or thereby the lack of physical behaviour, in Smagorinsky based models led to the development of the WALE Model by Nicoud and Ducros (1999). Maintaining a similar framework to the Smagorinsky model, the WALE model uses a different operator compared to the strain rate used in Smagorinsky and dynamic Smagorinsky - the operator was expected to satisfy invariance to coordinate transformation, easily solvable in any computational grid, be a function of irrotational and rotational strain rate, and decay to zero at any boundary or wall. The mathematical formulation for the WALE model is presented below.

Consider the generic form of any SGS model:

$$\nu_t = C_m \Delta^2 \overline{OP}(x, t) \quad (2.2.29)$$

where  $\overline{OP}(x, t)$  signifies the operator in space and time such as the strain rate ( $\bar{S}_{ij}$ ) operator for Smagorinsky and dynamic Smagorinsky models,  $C_m$  stands for the model coefficient while  $\Delta$  is the characteristic filter width.

For the WALE model, Nicoud and Ducros suggest an operator based on the traceless symmetric part of the square of the velocity gradient tensor ( $\varsigma_{ij}^d$ ) and the corresponding anti-symmetric ( $\bar{\Omega}_{ij}$ ) part of the first invariant of the velocity gradient tensor ( $\bar{g}$ ) defined as:

$$\varsigma_{ij}^d = \frac{1}{2}(\bar{g}_{ij}^2 + \bar{g}_{ji}^2) - \frac{1}{3}\delta_{ij}\bar{g}_{kk}^2 \quad (2.2.30)$$

$$\bar{\Omega}_{ij} = \frac{1}{2}\left(\frac{\partial \bar{u}_i}{\partial x_j} - \frac{\partial \bar{u}_j}{\partial x_i}\right) \quad (2.2.31)$$

where  $\bar{g}_{ij}^2 = \bar{g}_{ik}\bar{g}_{kj}$  and  $\delta_{ij}$  stands for the Kronecker symbol. Considering the definition of the local strain rate ( $\bar{S}_{ij}$ ),  $\varsigma_{ij}^d$  can be written as a function of  $\bar{S}_{ij}$  and  $\bar{\Omega}_{ij}$  with zero trace.

$$\varsigma_{ij}^d = \bar{S}_{ik}\bar{S}_{kj} + \bar{\Omega}_{ik}\bar{\Omega}_{kj} - \frac{1}{3}\delta_{ij}[\bar{S}_{mn}\bar{S}_{mn} - \bar{\Omega}_{mn}\bar{\Omega}_{mn}] \quad (2.2.32)$$

The second invariant of this operator is finite and is proportional to  $\varsigma_{ij}^d\varsigma_{ij}^d$  which can be expressed as (in the same fashion as Nicoud and Ducros (1999)):

$$\varsigma_{ij}^d\varsigma_{ij}^d = \frac{1}{6}(S^2S^2 + \Omega^2\Omega^2) + \frac{2}{3}S^2\Omega^2 + 2IV_{S\Omega} \quad (2.2.33)$$

where  $S^2 = \bar{S}_{ij}\bar{S}_{ij}$ ,  $\Omega^2 = \bar{\Omega}_{ij}\bar{\Omega}_{ij}$ ,  $IV_{S\Omega} = \bar{S}_{ik}\bar{S}_{kj}\bar{\Omega}_{jl}\bar{\Omega}_{li}$ .

This operator is expected to capture turbulence dependent on large strain rate, rotation rate or both. Nicoud and Ducros further performed wall scaling and dimensional analysis to provide an expression for eddy viscosity in terms of the above defined operator.

$$\nu_t = (C_w \Delta)^2 \frac{(\varsigma_{ij}^d \bar{\varsigma}_{ij}^d)^{3/2}}{(\bar{S}_{ij} \bar{S}_{ij})^{5/2} + (\varsigma_{ij}^d \bar{\varsigma}_{ij}^d)^{5/4}} \quad (2.2.34)$$

where  $C_w$  is a constant coefficient which can be determined by comparing the ensemble sub-grid kinetic energy dissipation of WALE with the classic Smagorinsky model. Nicoud proposed a value for this constant to be 0.5 for isotropic homogenous turbulence and this value was successfully applied to turbulent channel flow by Weickert et al. (2010).

The WALE model is expected to alleviate most of the issues related to classic Smagorinsky model while it is easier to computationally evaluate due to its non-dynamical nature as compared to the Dynamic Smagorinsky model. However, for flows with high vorticity the model was seen to be unstable (e.g. cylinder wake flow).

#### 2.2.4 Models under location uncertainty

General classical models such as Smagorinsky or WALE viscosity model proceed through a deterministic approach towards modelling the SGS dissipation tensor. However, Mémin (2014) suggests a stochastic approach towards modelling the SGS contributions in the NS equation. Building a stochastic NS formulation can be achieved via various methods. The simplest way consists in considering an additional additive random forcing (Bensoussan and Temam, 1973). Other modelling approaches considered the introduction of fluctuations in the sub-grid models (Mason and Thomson, 1992; Schumann, 1995). Also, in the wake of Kraichnan's work (Kraichnan, 1959), another choice consisted in closing the large-scale flow in the Fourier space from a Langevin equation (Dubrulle et al., 2005; Laval et al., 2003). One such model is the LES-Langevin model developed by Laval and Dubrulle (2006) which was applied to channel flow by Dolganov et al. (2011). Lagrangian models based on Langevin equation in the physical space have been also successfully proposed for turbulent dispersion (Sawford, 1986) or for probability density function (PDF) modelling of turbulent flows (Haworth and Pope, 1986; Pope, 1994, 2000). These attractive models for particle based representation of turbulent flows are nevertheless not well suited to a large-scale Eulerian modelling.

The models under location uncertainty (MULU) relies on a different stochastic framework of the NS equation recently derived from the splitting of the Lagrangian velocity into a smooth component ( $\mathbf{u}$ ) and a highly oscillating random velocity component ( $\sigma(\mathbf{X}_t, t)\dot{\mathbf{B}}$ ) (i.e. the uncertainty in the parcel location expressed as velocity) (Mémin, 2014):

$$\frac{d\mathbf{X}_t}{dt} = \mathbf{u}(\mathbf{X}_t, t) + \sigma(\mathbf{X}_t, t)\dot{\mathbf{B}}. \quad (2.2.35)$$

$\mathbf{X}_t$  refers to the position of a fluid parcel at a given time  $t$ . It is important to note that while these models are classified here under LES, they do not involve filtering techniques and thus  $\bar{\mathbf{u}}$  for the classical LES and  $\mathbf{u}$  of the MULU models are different quantities in principle. They can be equated in specific cases, for example when the mesh is used as the intrinsic filter for the classical LES models and the same mesh is used to define the large-scale smooth component for the MULU models.

The first term on the right-hand side of eq. (2.2.35) represents the large-scale smooth velocity component, while the second term is the small-scale component. This latter term is a random field defined (formally) as the derivative of a Brownian function  $\dot{\mathbf{B}} = d\mathbf{B}/dt$  and

a diffusion tensor  $\boldsymbol{\sigma}$ . The small-scale component is rapidly de-correlating at the resolved time scale with spatial correlations (which might be inhomogeneous and non stationary) fixed through the diffusion tensor. It is associated with a covariance tensor ( $\mathcal{Q}$ ):

$$\mathcal{Q}_{ij}(\mathbf{x}, \mathbf{y}, t, t') = \mathbb{E}((\boldsymbol{\sigma}(\mathbf{x}, t)d\mathbf{B}_t)(\boldsymbol{\sigma}(\mathbf{x}, t)d\mathbf{B}_t)^T) = c_{ij}(\mathbf{x}, \mathbf{y}, t)\delta(t - t')dt, \quad (2.2.36)$$

where  $\mathbb{E}$  is the expectation, and  $c_{ij}$  is the spatial covariance. In the following the diagonal of the covariance tensor, termed here as the variance tensor, plays a central role; it is denoted as  $\mathbf{a}(\mathbf{x}) \triangleq c(\mathbf{x}, \mathbf{x}, t)$ . This tensor is a  $(3 \times 3)$  symmetric positive definite matrix with dimensions of viscosity [ $\text{m}^2 \text{s}^{-1}$ ].

With such a decomposition, the rate of change of a scalar ( $q$ ) within a material volume ( $\mathcal{V}$ ), is given through a stochastic representation of the Reynolds Transport Theorem (RTT) (M emin, 2014; Resseguier et al., 2017a). For an incompressible small-scale random component ( $\nabla \cdot \boldsymbol{\sigma} = 0$ ) the RTT has the following expression:

$$d \int_{\mathcal{V}(t)} q = \int_{\mathcal{V}(t)} \left( d_t q + (\nabla \cdot (q\mathbf{u}^*) - \frac{1}{2} \sum_{i,j=1}^d \partial_{x_i} (a_{ij} \partial_{x_j} q)) dt + \nabla q \cdot \boldsymbol{\sigma} d\mathbf{B}_t \right) d\mathbf{x}. \quad (2.2.37)$$

where the effective advection  $\mathbf{u}^*$  is defined as:

$$\mathbf{u}^* = \mathbf{u} - \frac{1}{2} \nabla \cdot \mathbf{a}. \quad (2.2.38)$$

The first term on the right-hand side represents the variation of quantity  $q$  with respect to time:  $d_t q = q(\mathbf{x}, t + dt) - q(\mathbf{x}, t)$ . It is similar to the temporal derivative. It is important here to quote that  $q$  is a non differentiable random function that depends among other things on the particles driven by the Brownian component and flowing through a given location. The second term on the right-hand side stands for the scalar transport by the large-scale velocity. However, it can be noticed that this scalar advection is not purely a function of the large-scale velocity. Indeed, the large-scale velocity is here affected by the inhomogeneity of the small-scale component through a modified large-scale advection (henceforth termed as velocity bias  $\mathbf{u}^*$ ), where the effect of the fluctuating component is taken into account via the small-scale velocity auto-correlations  $\mathbf{a} = (\boldsymbol{\sigma}\boldsymbol{\sigma}^T)$ . A similar modification of the large-scale velocity was also suggested in random walks Langevin models by MacInnes and Bracco (1992) who studied various stochastic models for particle dispersion - they concluded that an artificially introduced bias velocity to counter particle drift was necessary to optimise the models for a given flow. In the framework of modelling under location uncertainty, this term appears automatically. The third term in the stochastic RTT corresponds to a diffusion contribution due to the small-scale components. This can be compared with the SGS dissipation term in LES Modelling. This dissipation term corresponds to a generalisation of the classical SGS dissipation term, which ensues in the usual context from the Reynolds decomposition and the Boussinesq's eddy viscosity assumption. Here it figures the mixing effect exerted by the small-scale component on the large-scale component. Despite originating from a very different construction, in the following, for ease of reference, we keep designating this term as the SGS contribution. The final term in the equation is the direct scalar advection by the small-scale noise.

It should be noted that the RTT corresponds to the differential of the composition of two stochastic processes. The Ito formulae, which is restricted to deterministic functions of a stochastic process, does not apply here. An extended formulae know as Ito-Wentzell (or generalised Ito) formulae must be used instead (Kunita, 1997).

Using the Stochastic RTT, the large-scale flow conservation equations can be derived (for the full derivation please refer to Mémin (2014); Resseguier et al. (2017a)). The final conservation equations are presented below:

**Mass conservation:**

$$d_t \rho + \nabla \cdot (\rho \mathbf{u}^*) dt + \nabla \rho \cdot \boldsymbol{\sigma} d\mathbf{B}_t = \frac{1}{2} \nabla \cdot (\mathbf{a} \nabla \rho) dt, \quad (2.2.39)$$

which simplifies to the following constraints for an incompressible fluid:

$$\nabla \cdot \boldsymbol{\sigma} = 0, \quad \nabla \cdot \mathbf{u}^* = 0, \quad (2.2.40)$$

The first constraint maintains a divergence free small-scale velocity field, while the second imposes the same for the large smooth effective component. We observe that the large-scale component,  $\mathbf{u}$ , is allowed to be diverging, with a divergence given by  $\nabla \cdot \nabla \cdot \mathbf{a}$ . As we shall see, this value is in practice quite low. This weak incompressibility constraint results in a modified pressure computation, which is numerically not difficult to handle. Imposing instead a stronger incompressibility constraint on  $\mathbf{u}$  introduces an additional cumbersome constraint on the variance tensor ( $\nabla \cdot \nabla \cdot \mathbf{a} = 0$ ). In this work we will rely on the weak form of the incompressibility constraint. The large-scale momentum equation boils down to a large-scale deterministic equation after separation between the bounded variation terms (i.e. terms in "dt") and the Brownian terms, which is rigorously authorised - due to uniqueness of this decomposition.

**Momentum conservation:**

$$\left( \partial_t \mathbf{u} + \mathbf{u} \nabla^T (\mathbf{u} - \frac{1}{2} \nabla \cdot \mathbf{a}) - \frac{1}{2} \sum_{ij} \partial_{x_i} (a_{ij} \partial_{x_j} \mathbf{u}) \right) \rho = \rho \mathbf{g} - \nabla p + \mu \Delta \mathbf{u}. \quad (2.2.41)$$

Similar to the deterministic version of the NS equation, we have the flow material derivative, the forces, and viscous dissipation. The difference lies in the modification of the advection which includes the velocity bias and the presence of the dissipation term which can be compared with the SGS term present in the filtered NS equation. Both of the additional terms present in the stochastic version depend on the auto-correlation tensor  $\mathbf{a}$ . Thus to perform a LES, one needs to model, either directly or through the small-scale noise, the auto-correlation tensor. Two methodologies can be envisaged: the first would be to model the stochastic small-scale noise ( $\sigma(\mathbf{X}_t, t) \mathbf{B}$ ) and thus evaluate the auto-correlation tensor. We term such an approach as purely 'stochastic LES'. The second method would be to model the auto-correlation tensor directly as it encompasses the total contribution of the small scales. This method can be viewed as a form of 'deterministic LES' using stochastically derived conservation equations and this is the approach followed in this chapter. The crux of the 'deterministic LES' approach thus revolves around the characterisation of the auto-correlation tensor. The small-scale noise is considered subsumed within the mesh and is not defined explicitly.

This opens up various possibilities for turbulence modelling. The specification of the variance tensor  $\mathbf{a}$  can be performed through an empirical local velocity fluctuation variance multiplied by a decorrelation time, or by physical models/approximations or using experimental measurements. The options explored in this thesis include physical approximation based models and empirical local variance based models as described below. Note that this derivation can be applied to any flow model. For instance, such a modelling has been successfully applied to derive stochastic large-scale representation of geophysical flows by Resseguier et al. (2017a,b,c) and to Taylor Green vortex flow by Harouna and

Mémin (2017). For the Taylor Green vortex flow, the MULU were shown to out-perform the dynamic Smagorinsky model especially at higher Re.

A similar stochastic framework arising from a decomposition of the Lagrangian velocity has been proposed in Holm (2015) and analysed in Crisan et al. (2017) and Cotter et al. (2017). This framework leads to enstrophy conservation whereas the formulation under location uncertainty conserves the kinetic energy of a transported scalar (Resseguier et al., 2017a). A detailed comparison of the two frameworks can be found in Resseguier (2017). The next section analyses different modelling approaches towards characterising the variance tensor in a deterministic sense.

### Physical approximation based models:

Similar to Smagorinsky's eddy viscosity model, the variance tensor for the formulation under location uncertainty can also be specified using the strain rate tensor. Termed henceforth as the Stochastic Smagorinsky model (StSm), it specifies the variance tensor similar to the eddy viscosity in the Classical Smagorinsky model:

$$\mathbf{a}(\mathbf{x}, t) = (C_s * \Delta)^2 \|\mathbf{S}\| \mathbb{I}_3, \quad (2.2.42)$$

where  $\mathbb{I}_3$  stands for  $3 \times 3$  identity matrix,  $\|\mathbf{S}\|$  is the norm of the strain-rate-tensor and  $C_s$  is the Smagorinsky coefficient. The equivalency between the two models can be obtained in the following case (as shown by Mémin (2014)):

The SGS contribution (effective advection and SGS dissipation) for the StSm model is:

$$\begin{aligned} \mathbf{u}_j \partial_{x_j} (\partial_{x_j} a_{kj}) + \sum_{ij} \partial_{x_i} (a_{ij} \partial_{x_j} \mathbf{u}_k) &= \mathbf{u}_j \partial_{x_j} (\partial_{x_j} \|\mathbf{S}\| \delta_{kj}) + \sum_{ij} \partial_{x_i} (\|\mathbf{S}\| \delta_{ij} \partial_{x_j} \mathbf{u}_k), \\ &= \mathbf{u}_k \Delta \|\mathbf{S}\| + \|\mathbf{S}\| \Delta \mathbf{u}_k + \sum_j \partial_{x_j} \|\mathbf{S}\| \partial_{x_j} \mathbf{u}_k, \end{aligned} \quad (2.2.43)$$

and the SGS contribution for Smagorinsky model ( $\nabla \cdot \boldsymbol{\tau}$ ) is:

$$\begin{aligned} \nabla \cdot \boldsymbol{\tau} &= \sum_j \partial_{x_j} (\|\mathbf{S}\| \mathbf{S}), \\ &= \sum_j \partial_{x_j} (\|\mathbf{S}\| (\partial_{x_j} \bar{u}_k + \partial_{x_k} \bar{u}_j)), \\ &= \sum_j \partial_{x_j} \|\mathbf{S}\| \partial_{x_j} \bar{u}_k + \partial_{x_j} \|\mathbf{S}\| \partial_{x_k} \bar{u}_j + \|\mathbf{S}\| \Delta \bar{u}_k. \end{aligned} \quad (2.2.44)$$

An equivalency can be drawn between the two models, for the case when the filtered velocity  $\bar{\mathbf{u}}$  is equivalent to the large-scale smooth component  $\mathbf{u}$ , by adding  $\sum_j \partial_{x_j} \|\mathbf{S}\| \partial_{x_k} \mathbf{u}_j - \mathbf{u}_k \Delta \|\mathbf{S}\|$  to the StSm model. The additional term may also be written as:

$$\partial_{x_k} \sum_j \partial_{x_j} (\|\mathbf{S}\|) \mathbf{u}_j - \sum_j \partial_{x_j} \partial_{x_k} (\|\mathbf{S}\|) \mathbf{u}_j - \mathbf{u}_k \Delta \|\mathbf{S}\|, \quad (2.2.45)$$

where the first term represents a velocity gradient which can be included within a modified pressure term as is employed for Smagorinsky model. The other two terms can be neglected for smooth enough strain rate tensor function. For smooth deformations both models are equivalent in terms of dissipation. It is important to note here that even if the effective advection is ignored in the StSm model, the two models still differ in a general case due to the first two terms in (2.2.45).

### Local variance based models:

As the name states, the variance tensor can be calculated by an empirical covariance of the resolved velocity within a specified local neighbourhood. The neighbourhood can be spatially or temporally located giving rise to two formulations. A spatial neighbourhood based calculation (referred to as Stochastic Spatial Variance model (StSp)) is given as:

$$\mathbf{a}(\mathbf{x}, t) = \frac{1}{|\Gamma| - 1} \sum_{x_i \in \Gamma(x)} (\mathbf{u}(x_i, t) - \mathbf{u}_m(x, t))(\mathbf{u}(x_i, t) - \mathbf{u}_m(x, t))^T C_{sp}, \quad (2.2.46)$$

where  $\mathbf{u}_m(x, t)$  stands for the empirical mean around the arbitrarily selected local spatial neighbourhood defined by  $\Gamma$ . The constant  $C_{sp}$  is defined as (Harouna and Mémin, 2017):

$$C_{sp} = \left( \frac{\ell_{res}}{\eta} \right)^{\frac{5}{3}} \Delta t, \quad (2.2.47)$$

where  $\ell_{res}$  is the resolved length scale,  $\eta$  is the Kolmogorov length scale and  $\Delta t$  is the simulation time step.

A similar local variance based model can be envisaged in the temporal framework as,

$$\mathbf{a}(\mathbf{x}, t) = \frac{1}{|\beta| - 1} \sum_{t_i \in \beta(t)} (\mathbf{u}(x, t_i) - \mathbf{u}_m(x, t))(\mathbf{u}(x, t_i) - \mathbf{u}_m(x, t))^T C_{te}, \quad (2.2.48)$$

where  $\mathbf{u}_m(x, t)$  now stands for a temporally averaged mean field and the coefficient  $C_{te}$  is defined identical to the spatial variance model coefficient.

It is important to note that the prefix stochastic has been added to the Mulu (i.e. Stochastic Smagorinsky or Stochastic Spatial or Stochastic Temporal) to differentiate the Mulu version of the Smagorinsky model from its classical purely deterministic version. The model equations while derived using stochastic principles are applied here in a purely deterministic sense. The full stochastic formulation of Mulu has been studied by Resseguier et al. (2017a,b,c) and is the focus of study only in chapter 4.

### 2.2.5 Other Approaches

A complementary approach to LES modelling is the so called implicit-LES modelling. All of the above described models use explicit equations to characterise the sub-grid scale dissipation. However, Boris (1990) introduced a novel approach where the numerical scheme was used to account for the sub-grid scale dissipation effect, called the monotone integrated large eddy simulation (MILES) - the first such use of artificial viscosity, however, can be credited to Richtmyer (1948). This work was further enhanced by numerous researchers and now is renamed to include all such models under the name of implicit large eddy simulation (ILES).

ILES models are based on the basic principle of energy cascade - the energy in a turbulent flow flows from the larger scales to smaller scales and is dissipated at the smaller scales. LES which captures the large scales and models the sub-grid scales using models, is altered by the addition of artificial viscosity instead of an SGS model in ILES. This concept identified by Boris et al. (1992) was termed as the "*convenient conspiracy*" by Oran and Boris (1993). ILES schemes have the option of running without a viscous term in the NS equation but many schemes exist that apply explicitly viscosity through the viscous term to capture any dissipation in the larger scales (Fureby and Grinstein, 1999).

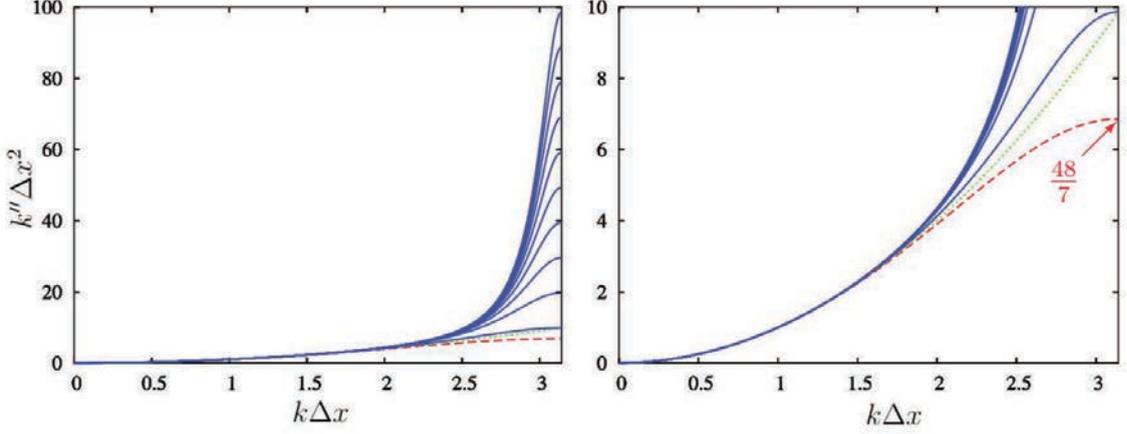


Figure 2.2: Comparison of the modified square wave-numbers  $k'' \Delta x^2$  (blue lines) for the sixth order compact scheme with the exact square wavenumber  $k^2 \Delta x^2$  (green line) and the conventional sixth order scheme (red line) (Lamballais et al., 2011)

The model proposed by Lamballais et al. (2011) uses a numerical substitute for the SGS models by adding a specified amount of artificial viscosity to the flow solver to mimic SGS dissipation. A novelty of this model, in context of Incompact3d (the solver in which it was applied), is the use of high order compact finite difference schemes for computing the second derivatives - the scheme being implicit in space provides wider control over the dissipation introduced. The second derivative sixth order scheme can be expressed in Fourier space as a function of the modified square wavenumber ( $k''$ ) expressed as:

$$k'' \Delta x^2 = \frac{2a[1 - \cos(k\Delta x)] + \frac{b}{2}[1 - \cos(2k\Delta x)] + \frac{2c}{9}[1 - \cos(3k\Delta x)]}{1 + 2\alpha\cos(k\Delta x)} \quad (2.2.49)$$

where the constants  $a$ ,  $b$ ,  $c$ , and  $\alpha$  define the order of the finite difference scheme for the second derivative.

Lamballais et al. (2011) showed that the use of this modified square wavenumber ( $k''$ ) with a finite difference scheme with added over dissipation compared to exact value ( $k^2$ ) is similar to the introduction of spectral viscosity  $\nu_s(k)$  given as:

$$\nu_s(k) = \nu \frac{k''(k) - k^2}{k^2} \quad (2.2.50)$$

where  $\nu$  stands for kinematic molecular viscosity. Figure 2.2 shows the curves for a set of  $k'' \Delta x^2$  for  $k_c'' \Delta x^2 = n\pi^2$  with  $n$  being an integer from 1 to 10 (where  $k_c''$  is the modified square wavenumber at cutoff) compared with the exact square wavenumber as well as the sixth order finite difference scheme with no modification (i.e.  $k_c'' \Delta x^2 = 48/7$ ). The curves match with the exact wavenumber for most of the regime except for a band of high wavenumber even for high over estimation by using  $n = 10$ . For more details about the scheme please refer to Lamballais et al. (2011). More details about the modified wave number concept can be found in Lele (1992).

The ILES scheme is implicit in space and this allows flexibility for  $k''(k)$  leading to any possible value for the spectral viscosity at the cut-off wavenumber ( $K_c = \pi/\Delta x$ ) while maintaining control on the kernel ( $\nu_s(k)/\nu_0$ ). This flexibility however has an associated issue of arbitrariness wherein the user has to choose  $\nu_0$  and the kernel for each specific case. Other interesting works in the field of ILES modelling can be found in Hickel et al. (2006); Margolin and Rider (2005); Margolin et al. (2006); Visbal et al. (2003).

All further models defined and described below are purely for sake of completeness and these models are not used within this thesis work. Vreman's model (Vreman, 2004), derived by investigating the algebraic properties of theoretical sub-grid dissipation, provides an alternative to the Smagorinsky set of models for which the eddy viscosity is defined as,

$$\nu_t = C_\nu \sqrt{\frac{Q_G}{\bar{\mathcal{A}}_{ij} \bar{\mathcal{A}}_{ij}}}, \quad (2.2.51)$$

where

$$\bar{\mathcal{A}}_{ij} = \frac{\partial \bar{u}_j}{\partial x_i}; \quad (2.2.52)$$

$$Q_G = \bar{\beta}_{11} \bar{\beta}_{22} - \bar{\beta}_{12}^2 + \bar{\beta}_{11} \bar{\beta}_{33} - \bar{\beta}_{13}^2 + \bar{\beta}_{22} \bar{\beta}_{33} - \bar{\beta}_{23}^2 \quad (2.2.53)$$

$$\bar{\beta}_{ij} = \sum_{m=1}^3 \Delta_m^2 \bar{\mathcal{A}}_{mi} \bar{\mathcal{A}}_{mj}, \quad (2.2.54)$$

where  $\Delta$  is the filter width. The model was defined with the intention of overcoming the drawbacks of the Smagorinsky model while maintaining a simple formulation. A constant value for the coefficient ( $C_\nu$ ) was proposed by Vreman as,

$$C_\nu = 2.5 C_s^2 \quad (2.2.55)$$

In a similar trend to the Smagorinsky model, this coefficient was found to be not universal and hence two dynamical update procedures were then suggested by Jiang et al. (2016) using first the Germano identity and the second based on the global equilibrium between the SGS dissipation and the viscous dissipation using two test-filters (Park et al., 2006). The second method was later modified and reduced to one test-filter by You and Moin (2006). The performance of various sub-sets of the Vreman model set can be found in Jiang et al. (2016); Vreman (2004); Lee et al. (2010).

A mathematical approach to large scale turbulence modelling was undertaken by Layton and Lewandowski (2003), where they developed a model using principles of scale similarity for periodic flows. For a constant flow  $\mathbf{u} = \bar{\mathbf{u}}$ , we get the following closure model,

$$\overline{\mathbf{u}\mathbf{u}} \cong \overline{\bar{\mathbf{u}}\bar{\mathbf{u}}}. \quad (2.2.56)$$

Historically, models built from scale-similarity principles tend to have stability issues when applied numerically. However, Layton and Lewandowski (2003) show mathematically the stability of their zeroth order model derived from scale-similarity principles. This mathematical model is further analysed in Layton and Lewandowski (2006a,b); Lewandowski (2006).

## 2.3 Hybrid approaches

For sake of completeness, this section describes the concept of hybrid model despite their lack of use in this thesis. RANS approaches and LES models both have certain advantages and disadvantages vis-à-vis computational cost, accuracy, simplicity, etc. A class of models, which fall under the 'hybrid RANS-LES models' classification, combine an LES approach with a RANS model depending on the nature of the flow. First developed for the Spalart-Allmaras model by Spalart (1997), this model, referred to as detached-eddy simulation, treats the near-wall region using a RANS model while the rest of the flow is

simulated using an LES model. The model emulates the Spalart-Allmaras RANS model near the wall while farther away the model is akin to a Smagorinsky LES model.

For a model to work in the hybrid sense, an appropriately turbulence length scale needs to be necessarily defined based on which the separation of regions into RANS or LES is done. Thus, not all RANS and LES models are compatible for such hybrid approaches. Baldwin-Barth model (Baldwin and Barth, 1990) is incompatible due to the use of extensive damping functions while Menter's Shear Stress Transport (SST) model (Menter, 1994) is compatible and has been employed by Strelets (2001) by replacing the model of Spalart with the SST model.

For cases where a clear definition of the turbulent length scale is difficult to obtain, an ad-hoc specification of regions for RANS modelling and regions for LES modelling can be specified thus overruling the distance function requirement. However, the use of a hybrid methodology is not a simple matter of specifying regions for model application. Such methodologies need to effectively address the difference in concept of the RANS and LES models. A review of DES approaches to turbulence modelling can be found in Spalart (2009).

For the framework of this thesis, only the LES modelling approaches are employed and analysed with turbulent flows in chapter 3. A preliminary application to homogenous channel flow is performed in §3.2. This is followed by an application to the more complex wake flow around a circular cylinder in §3.3. Model performance and physical characteristics are analysed in depth with respect to the latter flow. An optimal LES model is identified for this flow and this is used as the focus model for the 4D-Var<sub>les</sub> technique which is developed and applied in chapter 7 with the corresponding DA principles highlighted in chapter 5. Chapter 6 provides an innovative approach towards generating data-sets for use in 4D-Var<sub>les</sub>.



## Chapter 3

# Large eddy simulation - applications

The field of turbulence modelling has attracted significant attention from researchers leading to the development of a plethora of models. However, a universal approach to turbulence modelling or a universal model is yet to be developed which is applicable to all flows. The computational cost, mesh type, discretisation method, flow configuration, and turbulence intensity are a few parameters among many that define the model/approach optimal for a given flow. For DA studies, the aim is to explore higher Re flows which would require a significant reduction in computational cost. This can be ideally achieved using RANS models, however, at a significant loss of accuracy. In addition, the RANS models are not capable of reproducing the turbulent structures in a flow given their simulation of the averaged NS equation. Thus, the focus of this thesis is limited to LES modelling approaches.

Having established the approach under consideration, the LES model still needs to be chosen. Ideally, a 4D-Var<sub>les</sub> code should encompass many established models, thus providing an option for choosing an optimal model for different flow scenarios. For example, the WALE model is ideal for flows with moderate rotation while the dynamic Smagorinsky works well for flows with a homogenous flow direction. However, this implies a significant amount of work as careful consideration is needed for the merging of an LES SGS model with the 4D-Var assimilation procedure. In addition, the creation and validation of the required tangent/adjoint procedures (explained in chapter 7) requires considerable effort even using numerical tools. This thesis aims at proving the feasibility of such a merger as well as portraying the capabilities of a 4D-Var<sub>les</sub> code vis-à-vis accuracy, feasibility, and computational cost for one LES model.

This chapter focuses on identifying the optimal model to include in the 4D-Var<sub>les</sub> code. The models described under LES in the previous chapter, namely classic Smagorinsky (Smag), dynamic Smagorinsky using Germano's Identity (DSmag), WALE, Stochastic Smagorinsky (StSm), Stochastic Spatial (StSp), and Stochastic Temporal (StTe) models, are considered. A preliminary application to homogenous channel flow is performed in §3.2 followed by a more comprehensive study of the models for wake flow around a circular cylinder in §3.3 - this is the flow under study for DA. Before delving in to model performance, a description of the flow solver, Incompact3d, used for all CFD simulations in this thesis is provided in the next section (§3.1).

### 3.1 Flow solver - Incompact3d

Incompact3d (available at <http://www.incompact3d.com>) is a powerful, parallel, high order, fortran based flow solver developed by Laizet and Lamballais (2009) for simulating

fluid flows. Developed as a serial solver, the capabilities have been progressively enhanced to perform simulation using up to one million computational cores (Laizet and Li, 2011). Such a parallel scaling is possible due to the highly efficient 2D decomposition library 2decomp (available at <http://www.2decomp.org>) (Li and Laizet, 2010) and a distributed Fast Fourier Transform (FFT) interface.

The solver is based on a Cartesian mesh, the simplicity of which offers the ability to implement higher order spectral schemes for spatial discretisation. It also enables the implementation of the Immersed Boundary Method (IBM) to account for solid objects in the flow domain. The solving of the Poisson equation to ensure incompressibility in the spectral space using a modified wave number formalism represents the spirit of this code - efficient, fast, and accurate simulation of fluid flows. A concise description of the numerical schemes used by Incompact3d now follows.

### Governing equation

Incompact3d aims at solving the incompressible NS equations, recollected here,

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{1}{2}[\nabla(\mathbf{u}^k \otimes \mathbf{u}^k) + (\mathbf{u}^k \cdot \nabla)\mathbf{u}^k] = -\nabla p + \nu \Delta \mathbf{u} + \mathbf{f}, \quad (3.1.1)$$

$$\nabla \cdot \mathbf{u} = 0. \quad (3.1.2)$$

The convective term of the NS equation has been rewritten in its skew-symmetric form to reduce aliasing errors and  $\mathbf{f}$  stands, in general, for an additional forcing term corresponding to the IBM. The equation and the corresponding quantities are all non-dimensionalised with a density  $\rho$  of one.

### Solver parameters

The governing equations are solved in a computational domain measuring  $L_x \times L_y \times L_z$  discretised using a Cartesian mesh in the three principle directions with  $n_x \times n_y \times n_z$  nodes. An option is provided for stretching the mesh in the  $y$  direction with the choice of clustering node points at the boundaries or in the centre of the domain. 2D decomposition into pencil structures (see figure 3.1) is performed using the open source 2decomp library. Four different boundary conditions can be implemented in each principle direction, namely periodic, free-slip, no-slip, and open boundary conditions. For the dirichlet conditions, special care needs to be taken in imposing the values at the boundary within the time-advancement schemes. Incompressibility of the fluid is maintained by solving the Poisson equation using a classic fractional step method. The flow-chart for main processes within the code with corresponding 2D pencil changes can be seen in figure 3.2. The code provides the opportunity to save the current state of the simulation and to restart the simulation from said current state.

### Temporal scheme

Time advancement in Incompact3d of velocity  $\mathbf{u}$  from time  $t_k$  to time  $t_{k+1}$  can be expressed as,

$$\frac{\mathbf{u}^* - \mathbf{u}^k}{\Delta t} = a_k \mathbf{F}^k + b_k \mathbf{F}^{k-1} + g_k \tilde{\mathbf{f}}^{k-2} - c_k \nabla \tilde{p}^k, \quad (3.1.3)$$

$$\frac{\mathbf{u}^{**} - \mathbf{u}^*}{\Delta t} = c_k \nabla \tilde{p}^k, \quad (3.1.4)$$

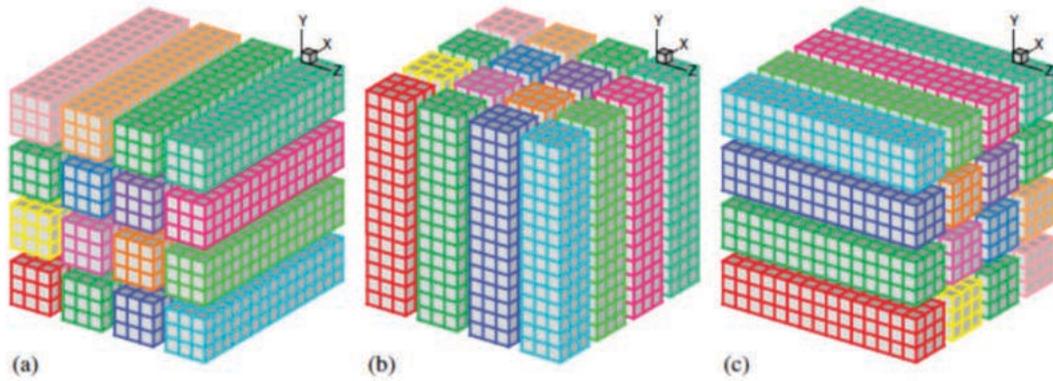


Figure 3.1: Pencil type decomposition of the computational domain in all three principle directions (x, y, z) respectively for a  $4 \times 4$  processor grid.

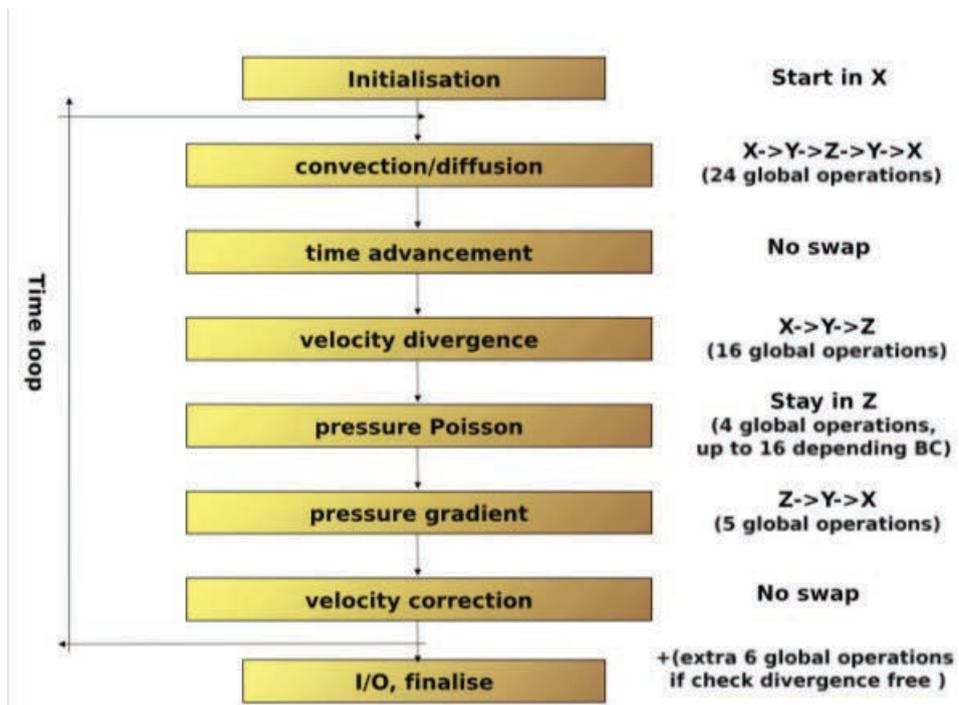


Figure 3.2: Incompact3d processes flow-chart with associated 2D pencil swaps (denoted using ->). Obtained from Incompact3d user-guide.

Table 3.1: Coefficients for temporal schemes for time advancement

| Scheme | $n_k$ | k | $a_k$           | $b_k$           | $c_k$ | $g_k$           |
|--------|-------|---|-----------------|-----------------|-------|-----------------|
| AB2    | 1     | 1 | 3/2             | -1/2            | 0     | 1               |
| AB3    | 1     | 1 | 23/12           | -16/12          | 5/12  | 1               |
| RK3    | 3     | 1 | 8/15            | 0               | 0     | 8/15            |
|        |       | 2 | 5/12            | -17/60          | 0     | 2/15            |
|        |       | 3 | 3/4             | -5/12           | 0     | 1/3             |
| RK4    | 5     | 1 | 0               | 0.1496590219993 | 0     | 0.1496590219993 |
|        |       | 2 | -0.4178904745   | 0.3792103129999 | 0     | 0.220741935365  |
|        |       | 3 | -1.192151694643 | 0.8229550293869 | 0     | 0.25185480577   |
|        |       | 4 | -1.697784692471 | 0.6994504559488 | 0     | 0.33602636754   |
|        |       | 5 | -1.514183444257 | 0.1530572479681 | 0     | 0.041717869325  |

$$\frac{\mathbf{u}^{k+1} - \mathbf{u}^{**}}{\Delta t} = -c_k \nabla \tilde{p}^{k+1}, \quad (3.1.5)$$

where,

$$\mathbf{F}^k = -\frac{1}{2}[\nabla(\mathbf{u}^k \otimes \mathbf{u}^k) + (\mathbf{u}^k \cdot \nabla)\mathbf{u}^k] + \nu \nabla^2 \mathbf{u}^k, \quad (3.1.6)$$

$$\tilde{p}^{k+1} = \frac{1}{c_k \Delta t} \int_{t_k}^{t_{k+1}} p dt, \quad \tilde{\mathbf{f}}^{k+1} = \frac{1}{g_k \Delta t} \int_{t_k}^{t_{k+1}} \mathbf{f} dt. \quad (3.1.7)$$

$\mathbf{F}^k$  represents the convection and diffusion contribution. In the classical LES framework,  $\mathbf{F}^k$  is given as,

$$\mathbf{F}^k = -\frac{1}{2}[\nabla(\mathbf{u}^k \otimes \mathbf{u}^k) + (\mathbf{u}^k \cdot \nabla)\mathbf{u}^k] + \nu \Delta \mathbf{u}^k + \nu_t \Delta \mathbf{u}^k, \quad (3.1.8)$$

where the final term stands for the SGS component and can take different forms depending on the model used. The formulation of equation (3.1.3) allows the use of four temporal schemes, namely Adams-Bashforth scheme in second (AB2) and third order (AB3), and the Runge-Kutta scheme in third (RK3, Williamson (1980)) and fourth order (RK4, Kennedy et al. (2000)), depending on the values of  $a_k$ ,  $b_k$ , and  $c_k$  and number of intermediary time-steps  $n_k$ . Table 3.1 tabulates the values of the variables for each scheme.

### Spatial scheme

The finite difference scheme for the first derivative is of the form,

$$\alpha f'_{i-1} + f'_i + \alpha f'_{i+1} = a \frac{f_{i+1} - f_{i-1}}{2\Delta x} + b \frac{f_{i+2} - f_{i-2}}{4\Delta x}. \quad (3.1.9)$$

A sixth-order accurate scheme can be obtained by setting the coefficient ( $\alpha$ ,  $a$ ,  $b$ ) as (1/3, 14/9, 1/9). Such a scheme has a quasi-spectral behaviour with a compact formulation. The second derivative can then be expressed as,

$$\alpha f''_{i-1} + f''_i + \alpha f''_{i+1} = a \frac{f_{i+1} - 2f_i + f_{i-1}}{\Delta x^2} + b \frac{f_{i+2} - 2f_i + f_{i-2}}{4\Delta x^2} + c \frac{f_{i+3} - 2f_i + f_{i-3}}{9\Delta x^2}. \quad (3.1.10)$$

A similar sixth-order accuracy with spectral-like resolution can be obtained with the values (2/11, 12/11, 3/11, 0) for the variables ( $\alpha$ ,  $a$ ,  $b$ ,  $c$ ). An alternate discretisation of the same order can be obtained with values (0.47959871686180711, 0.42090288706093404, 1.7020738409366740, -0.16377929427399390) which controls aliasing errors by being over-dissipative at the highest wave numbers at the cost of a less-compact formulation.

### Boundary conditions and treatment

Periodic, free-slip, or no slip/open boundary conditions or a combination of these can be imposed within Incompact3d. For the periodic condition and the free-slip condition, i.e. symmetric and anti-symmetric condition, the boundaries are discretised using the same schemes with the introduction of ghost nodes which are assigned values depending on the type of condition implemented. Thus, the function and their derivatives on the left boundary can be expressed (the right boundary conditions can be easily deduced in similarity).

Periodic condition:

$$f_o \rightarrow f_{n_x}; f_{-1} \rightarrow f_{n_x-1}; f'_0 \rightarrow f'_{n_x}; f''_0 \rightarrow f''_{n_x}. \quad (3.1.11)$$

Symmetric free-slip condition:

$$f_o \rightarrow f_2; f_{-1} \rightarrow f_3; f'_0 \rightarrow -f'_2; f''_0 \rightarrow f''_2. \quad (3.1.12)$$

Antisymmetric free-slip condition:

$$f_o \rightarrow -f_2; f_{-1} \rightarrow -f_3; f'_0 \rightarrow f'_2; f''_0 \rightarrow -f''_2. \quad (3.1.13)$$

For the dirichlet conditions of no slip and open boundary conditions, a one sided formulation is implemented within the flow domain. For the boundary node, the third-order discretisation scheme (Lele, 1992) is implemented as,

$$f'_1 + 2f'_2 = \frac{1}{2\Delta x}(-5f_1 + 4f_2 + f_3), \quad (3.1.14)$$

$$f''_1 + 11f''_2 = \frac{1}{\Delta x^2}(13f_1 - 27f_2 + 15f_3 - f_4), \quad (3.1.15)$$

and for the boundary adjacent node, fourth order accurate Padé schemes are employed as,

$$\frac{1}{4}f'_1 + f'_2 + \frac{1}{4}f'_3 = \frac{3}{2} \frac{f_3 - f_1}{2\Delta x}, \quad (3.1.16)$$

$$\frac{1}{10}f''_1 + f''_2 + \frac{1}{10}f''_3 = \frac{6}{5} \frac{f_3 - 2f_2 + f_1}{\Delta x^2}. \quad (3.1.17)$$

The use of such compact schemes provides computational advantages over purely spectral schemes due to the presence of the non-linear advection term. Especially with the introduction of LES models which could include further non-linear terms that require treatment, the compact schemes present a clear advantage over spectral scheme which are more preferable for linear equations (Laizet and Lamballais, 2009).

### Immersed boundary method

The IBM is used to account for solid bodies within the flow domain through an additional forcing term in the time advancement equation defined as,

$$g_k \tilde{f}^{k+1} = \epsilon \left( -a_k \mathbf{F}^k - b_k \mathbf{F}^{k-1} + c_k \nabla \tilde{p}^k + \frac{\mathbf{u}_0^{k+1} - \mathbf{u}^k}{\Delta t} \right), \quad (3.1.18)$$

where  $\epsilon$  is set as 1 inside and on the solid body and 0 elsewhere.

IBM method based on feedback forcing by Goldstein et al. (1993), as well as direct forcing method of Fadlun et al. (2000) have been implemented within Incompact3d to varying degrees of success - for example see the works of Silvestrini and Lamballais (2002, 2004); Pinto et al. (2011); Lamballais et al. (2008, 2010). However, for complex geometries both these IBM methods are not suitable to impose the no-slip boundary condition while maintaining compatibility with the higher-order schemes. A new IBM referred to as ‘alternating direction forcing (ADF) strategy’ was proposed by Gautier et al. (2014) which is used for the LES simulations in this thesis.

The main drawback of the direct forcing method is the lack of a smooth solution for the velocity derivatives at the interface between the solid and fluid boundaries. This arises as the velocity is enforced to be 0 at all points within the solid body thus affecting the continuity of spatial derivatives. This error, while minimal for low-order schemes, can result in spurious oscillations for high-order schemes used in Incompact3d. One method to negate this is to impose a non-zero velocity field within the solid body while maintaining a no-slip condition at the boundary - a feat possible only for the most simplistic solid body geometries. The ADF strategy is based on 1D expansion of the solution in the derivative direction within the solid body using a Lagrangian polynomial. The expansion is set such that the no-slip condition is maintained at the boundary while maintaining a smooth spatial derivative at the boundary and adjacent nodes. A sample Lagrangian polynomial expansion in 1D is shown in figure 3.3. For more details on the IBM method, please refer to Gautier et al. (2014).

### Pressure treatment

The incompressibility constraint ( $\Delta \cdot \mathbf{u}^{k+1} = 0$ ) is verified at each sub-time step by solving the Poisson equation,

$$\nabla \cdot \nabla \tilde{p}^{k+1} = \frac{\nabla \cdot \mathbf{u}^{**}}{c_k \Delta t}, \quad (3.1.19)$$

which gives the value of  $\tilde{p}^{k+1}$  required for time-advancement. For flows involving the IBM method to account for solid boundaries/ bluff bodies, the Poisson equation is expressed as,

$$\nabla \cdot \nabla \tilde{p}^{k+1} = \frac{\nabla \cdot [(1 - \epsilon) \mathbf{u}^{**}]}{c_k \Delta t}, \quad (3.1.20)$$

The computational effort for iteratively solving the Poisson equation especially with high-order schemes is generally prohibitive and requires the use of sophisticated methods when solved in the physical space (Mercier and Deville, 1981). Instead, if this inversion is performed in Fourier space the corresponding operation is easy to implement and computationally inexpensive with the help of FFT routines. The amalgamation of the Incompact3d with existing FFT libraries proved to be a tedious task resulting in the development from scratch of the generic FFT routines by Li and Laizet (2010) based on Glassmans’ general N FFT. The implementation of the Fourier approach for the periodic and free-slip conditions is straight-forward. However, the advantage of the method lies in its ability to

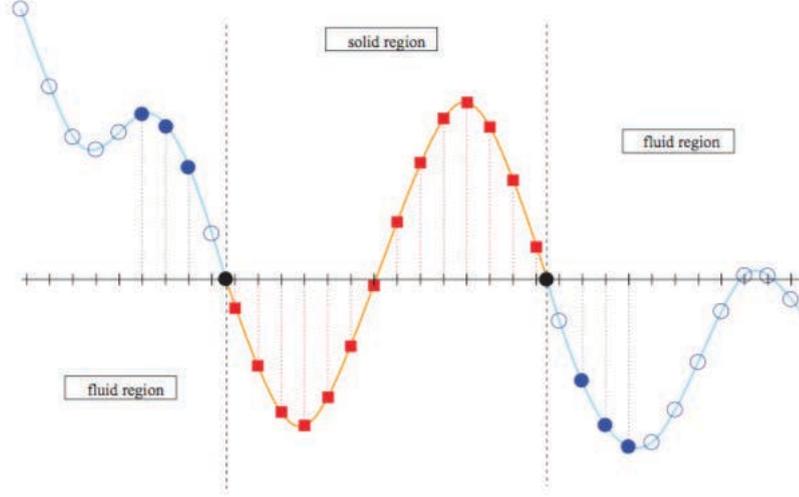


Figure 3.3: Lagrangian polynomial based reconstruction of flow within a solid body - a schematic. Blue filled points - mesh points considered for the reconstruction. Red filled points - points reconstructed within the solid boundary. Blue empty points - mesh points not considered for reconstruction. Black points - Boundary points between solid and fluid regions. Borrowed from Gautier et al. (2014).

obtain second-order accurate solutions from a Poisson equation for homogenous Neumann boundary conditions.

The pressure field is evaluated in the same number of mesh nodes as the velocity field but with a staggered placement of the mesh points as shown in figure 3.4. Such a staggered mesh helps reduce spurious oscillation in the flow and is easy to implement without any need for extensive interpolation procedures for the mid-points (Wilhelmson and Ericksen, 1977; Canuto and Funaro, 1988; Chen et al., 2010).

The corresponding mid-point first-order derivative is expressed as,

$$\alpha f'_{i-1/2} + f'_{i+1/2} + \alpha f'_{i+3/2} = a \frac{f_{i+1} - f_i}{\Delta x} + b \frac{f_{i+2} - f_{i-2}}{3\Delta x} \quad (3.1.21)$$

with values of  $(9/62, 63/62, 17/62)$  for  $(\alpha, a, b)$  providing a better spectral behaviour than its counter-part for non-staggered mesh points. A sixth order accurate scheme is obtained in combination with the interpolation procedure given as,

$$\alpha f'_{i-1/2} + f'_{i+1/2} + \alpha f'_{i+3/2} = a \frac{f_{i+1} + f_i}{2} + b \frac{f_{i+2} + f_{i-2}}{2} \quad (3.1.22)$$

when  $(\alpha, a, b)$  is  $(3/10, 3/4, 1/20)$  respectively. More details on the pressure treatment within Incompact3d can be found in Laizet and Lamballais (2009).

When the MULU models are used within incompact3d, a modification of the Poisson equation needs to be done as,

$$\nabla \cdot \nabla \tilde{p}^{k+1} = \frac{\nabla \cdot (\mathbf{u}^{**} - \frac{1}{2} \nabla \cdot \mathbf{a})}{c_k \Delta t}, \quad (3.1.23)$$

where  $a$  is the variance tensor of the small scale noise. This is required in order to ensure the weak incompressibility constraint is maintained,

$$\nabla \cdot \mathbf{u}^* = \nabla \cdot (\mathbf{u} - \frac{1}{2} \nabla \cdot \mathbf{a}) = 0. \quad (3.1.24)$$

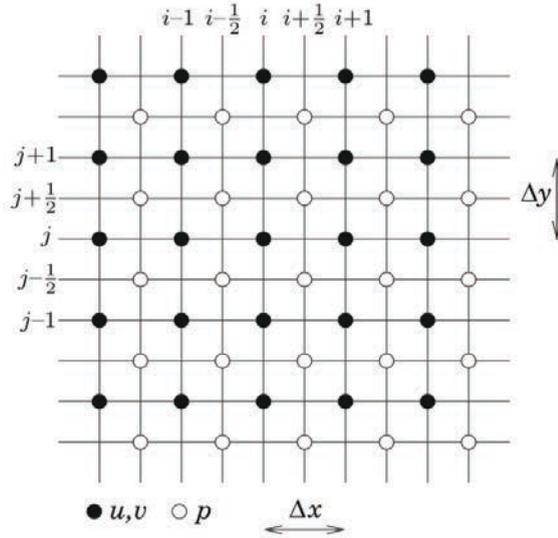


Figure 3.4: Staggered mesh schematic for velocity field (filled) and pressure field (empty) in a cartesian mesh.

Table 3.2: IGRIDA hardware architecture of infiniband clusters

| Cluster | Nodes | CPU                         | Reference                                   | RAM  | Disk       | Network               |
|---------|-------|-----------------------------|---|------|------------|-----------------------|
| Calda   | 5     | 2 × 8 cores<br>Sandy Bridge | Intel(R) Xeon(R) CPU<br>E5-2450 0 @ 2.10GHz | 48GB | 2 × 600 GB | Infiniband<br>+ 1GB/s |
| Lambda  | 11    | 2 × 6 cores<br>Westmere-EP  | Intel(R) Xeon(R) CPU<br>E5645 @ 2.40GHz     | 48GB | 2 × 600 GB | Infiniband<br>+ 1GB/s |

The enforcement of the strong incompressibility ( $\nabla \cdot \mathbf{u} = \nabla \cdot \nabla \cdot \mathbf{a} = 0$ ) is difficult as it is challenging to define a variance tensor  $\mathbf{a}$  that satisfies this condition.

### Computation resources

The computational grid, IGRIDA, available at INRIA - Rennes was used for the parallel computations with Incompact3d. IGRIDA was designed to provide a platform for code development and production runs. The infrastructure is made up of 138 computing nodes containing 1700 cores with a shared storage space of 3.4 TB for temporary storage. While 138 computing nodes exist for computational purposes, only 5 nodes ( $5 \times 16 = 80$  cores) of the Calda cluster and 11 nodes ( $11 \times 12 = 132$  cores) of the Lambda cluster are equipped with infiniband capabilities. Infiniband facilitates faster file transfer between processors. Considering the amount of data transfer between processors performed with each iteration in Incompact3d (denoted by 2D pencil swap in figure 3.2), simulation of flows with non-infiniband nodes is strictly prohibitive. It was observed that an increase in number of nodes was offset by the cost of data-transfer in non-infiniband nodes. The hardware architecture of the clusters used in this thesis are provided in table 3.2.

## 3.2 Channel flow at $Re_\tau = 395$

A channel flow (see figure 3.5) is a wall bounded flow with the dimensions in the wall normal plane being larger compared to the height (wall normal direction) of the channel ensuring statistical homogeneity along the streamwise and spanwise direction. This is generally achieved in a simulation by implementing periodicity boundary condition in the flow in transverse directions.

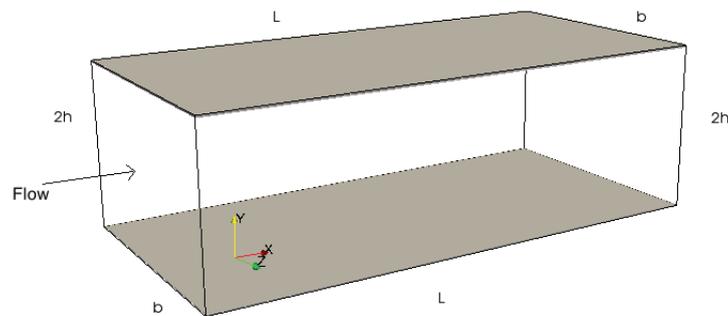


Figure 3.5: Channel Flow Set-up.

The flow, as depicted, is along the streamwise ( $x$ ) direction while the statistics of interest is generally the variation of any quantity along the height of the channel. The boundary conditions as applied in a simulation are generally periodic along streamwise and spanwise ( $z$ ) directions with no slip on the top and bottom walls. Channel flow, being one of the simplest wall bounded flow, has attracted a lot of research attention. Literature on channel flow is numerous - the works of Moin and Kim (1982); Hamba (2003); Völker et al. (2002); Kim et al. (1987) are a notable few.

Moser et al. (1999) performed a DNS of channel flow up to  $Re_\tau$  of 590, where  $Re_\tau = \frac{u_\tau H}{\nu}$  ( $u_\tau$  is the friction velocity). This work was further expanded by Lee and Moser (2015) up to a  $Re_\tau$  of 5200. These DNS statistics have been made available freely at (<http://turbulence.ices.utexas.edu/>) and have become a reference corner-stone for channel flow studies. The availability of validated reference statistics over a large range of Reynolds number makes channel flow a useful research tool for turbulence studies (Zhao et al., 2014; Nouri et al., 2012; Cziesla et al., 1999; Hamba, 2003). The simplicity of the flow and ease of simulation combined with the turbulent characteristics of the flow close to the wall of the channel makes it an ideal flow for performing model studies.

The flow, solver, and model parameters are described in §3.2.1 while the model performance is analysed in §3.2.2 through a statistical comparison.

### 3.2.1 Channel flow parameters

Channel flow at a  $Re_\tau$  of 395 is simulated using Incompact3d. The flow is constrained to a computational domain measuring  $2\pi \times 2 \times 3.14$  discretised into  $48 \times 81 \times 48$  cartesian mesh points. This discretisation is of particular interest as the cost of performing at this mesh level is only 1% the cost of the reference DNS of Moser et al. (1999). The coarseness of the mesh can be observed in the instantaneous streamwise velocity snapshot seen in figure 3.6. As such this is a much coarser mesh than generally used for LES which ranges

from 3% - 5% of the DNS mesh. The mesh is stretched in the wall normal ( $y$ ) direction with a larger clustering of points near the walls to capture well the turbulent boundary layer on the wall. The flow parameters and model parameters are tabulated in table 3.3. Time advancement is done using the RK3 scheme with a time-step of 0.002.

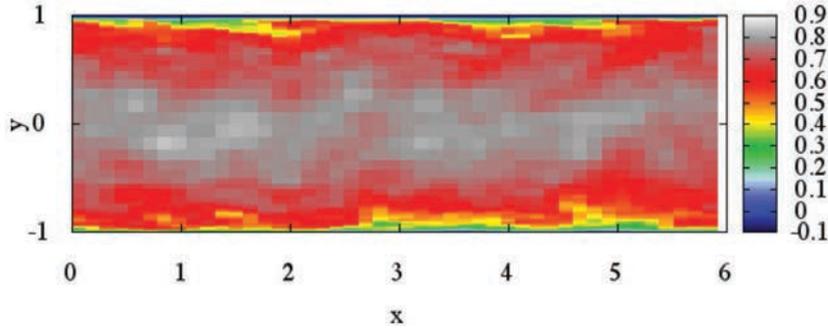


Figure 3.6: Instantaneous streamwise velocity snapshot for channel flow at  $Re_\tau$  395.

Incompact3d flow solver is inherently a DNS solver with the option of activating the ILES model of Lamballais et al. (2011) for LES studies. Thus, for the purpose of this thesis, the following LES models: Smag, DSmag, WALE, StSm, StSp, and StTe, in the form enumerated in chapter §2.2, were incorporated in incompact3d alongside the ILES model. The model coefficient for Smag and StSm for this flow configuration was set at 0.065 (Moin and Kim, 1982) while the constant for WALE was set as 0.5 (Nicoud and Ducros, 1999). A  $7 \times 7 \times 7$  local spatial neighbourhood was considered for StSp while a temporal neighbourhood of 7 was considered for StTe. The constant for StSp and StTe are defined based on the DNS mesh size, the LES resolution and the LES time step. A more detailed description of model inclusion in incompact3d is provided in §3.3.

The characterisation of channel flow is done in terms of either the friction velocity based Reynolds number (as done by Moser et al. (1999)) or the centre-line velocity based Reynolds number ( $Re_c$ ) (as required by Incompact3d). This poses a numerical problem as the friction velocity for a given  $Re_c$  cannot be accurately predicted before the simulation and hence an iterative methodology has to be employed. A initial guess for  $Re_c$ , based on theoretical calculations, is used to simulate the flow following which the friction velocity and  $Re_\tau$  are calculated.  $Re_c$  is then adjusted based on the value of  $Re_\tau$  obtained and the simulation is iteratively re-run until a  $Re_\tau \sim 395$  is obtained. This situation is further complicated for an LES study as each model performs differently near the wall and hence results in a different  $Re_\tau$  necessitating individual optimisation for each model. The final  $Re_c$  values for each model is tabulated in table 3.4.

The flow is initialised with a parabolic velocity profile in the  $y$  direction with a white noise incorporated initial condition - the noise is set to 10% of the centerline velocity. Periodic boundary conditions are enforced in the streamwise and spanwise direction with the no-slip condition enforced on the walls. In order to induce turbulence in the channel, the flow is rotated, after initialisation, for 5 non-dimensional time. All simulations are run until convergence and symmetry is established along the wall normal direction.

### 3.2.2 Channel flow results

The performance of the classical models namely Smagorinsky, its dynamic version and WALE for various flows has been established in literature. However, the ILES model of

|           | $n_x \times n_y \times n_z$ | $l_x \times l_y \times l_z$ | $\Delta x$ | $\Delta y$ | $\Delta z$ | $\Delta t$ | Model Coeff. |
|-----------|-----------------------------|-----------------------------|------------|------------|------------|------------|--------------|
| Smag      | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | 0.065        |
| DSmag     | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | eq. (2.2.28) |
| WALE      | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | 0.5          |
| StSm      | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | 0.065        |
| StSp      | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | eq. (2.2.47) |
| StTe      | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | eq. (2.2.47) |
| ILES      | 48×81×48                    | 6.28×2×3.14                 | 0.13       | 0.005-0.12 | 0.065      | 0.002      | -            |
| Reference | 256×257×256                 | 6.28×2×3.14                 | 0.024      | 0.0077     | 0.012      | -          | -            |

Table 3.3: Channel flow configuration and model parameters.

| Model  | Reference | Smag  | DSmag | Wale  | StSm  | StSp  | StTe  | ILES  |
|--------|-----------|-------|-------|-------|-------|-------|-------|-------|
| $Re_c$ | 10323     | 12000 | 10800 | 10900 | 11000 | 10200 | 10200 | 10500 |

Table 3.4: Centreline Reynolds number for different SGS models.

Lamballais et al. (2011) as well as the MULU of Mémin (2014) are relatively new models with limited literature studies on their performance. For the MULU, in the reformulation of the Navier-Stokes equations, the contributions of the sub-grid scale random component is split into an inhomogeneous turbulent diffusion and a velocity bias which corrects the advective transport due to resolved velocity field. Such a scheme has been shown to perform well on the Taylor Green vortex flow (Harouna and Mémin, 2017) at Reynolds number of 1600, 3000, and 5000. The new scheme was shown to outperform the established dynamic Smagorinsky model especially at higher  $Re$ . However, this flow is associated to an almost isotropic turbulence and no comparison with data is possible (as it is a pure numerical flow).

In this chapter, we wish to analyse the performance of these models for turbulent flows and classify the model performance in terms of the established classical SGS models. As a preliminary step towards more complex turbulent flow simulations, the model performance is analysed here for simple, homogenous turbulence in a plane channel. A good model performance can then be a precursor for simulating the more complex turbulent wake flow in §3.3.

Statistical comparison presented in figures 3.7-3.9 are time-averaged over 50000 time-steps and spatially-averaged along the homogenous streamwise ( $x$ ) and spanwise ( $z$ ) directions. The statistics presented have been normalised with respect to friction velocity and  $Re_\tau$ . All statistics are shown for half channel height to avoid redundancy due to symmetry. Considering the coarseness of the mesh, the first wall normal mesh-point is set at  $y^+ = 2$  and hence all profiles below this  $y^+$  are not captured in the LES statistics.

Figure 3.7 shows the mean streamwise velocity ( $u$ ) profile (the log law of the wall) for the channel normalised with respect to  $u_\tau$ . The iterative methodology for determining  $Re_c$  has biased the statistics for Smag due to its inherent inability to perform well close to the wall - the resultant  $Re_c$  is much higher than the theoretical value leading to a more energetic flow as seen in the higher velocity profile. This bias is moderately corrected by StSm, and DSMag, however, the resulting profile is significantly more energetic than the reference. The best statistical match with the reference is observed for the StSp and StTe models while the WALE model results in a slightly more energetic profile. The condition enforced

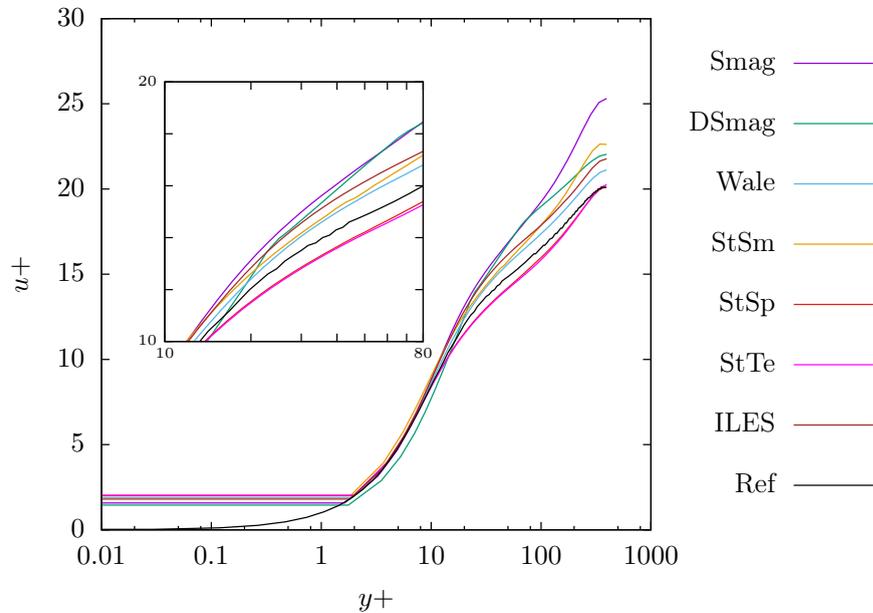


Figure 3.7: Mean velocity profile for turbulent channel flow at  $Re_\tau = 395$ .

on  $Re_\tau$  for all models can be seen in the near wall profile where all models estimate a similar magnitude thus resulting in a similar velocity gradient which defines the friction velocity and hence  $Re_\tau$ . A slight relaxation on the value of  $Re_\tau$  is the cause for the variations in profile close to the wall.

Figure 3.8 and figure 3.9 show the velocity and vorticity fluctuation profiles respectively. The solid lines indicate the streamwise fluctuation ( $u'or\omega'_x$ ) while the dash-dotted line is for the wall-normal fluctuation ( $v'or\omega'_y$ ) profiles and the dash-double dotted line is for the spanwise fluctuation ( $w'or\omega'_z$ ) profiles. A similar trend as seen in the log-law profile is observed in the velocity fluctuation profiles. Smag and StSm are seen to over-predict all profiles. ILES, WALE, StSp and StTe models predict a good match with the reference of the velocity fluctuation profiles. An opposite behaviour is observed for the WALE model and the other local-variance based models (StSp and StTe) - an over-prediction by WALE corresponds to an under-prediction by StSp and StTe, and vice versa. However, the mismatch at velocity maximum indicates that StSp and StTe better predict the velocity fluctuation statistics. The similar performance of the StSp and StTe models can be attributed to the flow itself which displays spatial and temporal homogeneity. Thus the variance calculated through a spatial averaging or a temporal averaging should result in a similar tensor. The slight difference between the two could be due to the difference in averaging periods for both models. The local-variance based models clearly outperform their Smagorinsky counter-part (StSm) for all profiles. A more physical analysis of the MULU and specifically the variance based models is provided in §3.3. It is interesting to note that ILES which has been fine-tuned for this Re also performs quite well. However, the amount of numerical dissipation to add requires ad-hoc specifications and this can be quite tedious requiring multiple iterations especially for complex flows Lamballais et al. (2011). There, currently, exists no methodology to determine the amount of numerical dissipation that needs to be added for a given flow at a given Reynolds number. Thus, an user defined estimate is used to simulate and depending on the performance, the amount of dissipation is tuned to obtain results. This is not feasible due to computational limitations.

Thus ILES is not considered further in this thesis for more complex flows.

The vorticity fluctuation profiles show more deviation between models and reference. A clear peak under-prediction for  $\omega_y$  is seen for all models while for  $\omega_x$  the maxima is well captured by StSp and StTe while the minima is better captured by the WALE model. Smag model's tendency to over-predict continues in these profiles with significant mismatch even in the centre of the channel. A better profile is obtained for StSp and StTe as we approach the centre of the channel. Clearly none of the models are able to capture accurately the vorticity profiles in all three principle directions. Individual matches suggests that ILES and StSp/StTe are best suited for estimating vorticity albeit with significant inaccuracies. The coarseness of the mesh could be a contributing factor towards the mismatch in these profiles. However, considering that this was a preliminary study, the inaccuracy in vorticity predictions are not analysed further in detail. It is interesting to note that in all statistics, the variance based StSp and StTe models both predict very similar profiles with minimal deviate from each other. This suggests that the local variance can be calculated either spatially or temporally indicating a statistical invariance between space and time.

From the mean and fluctuating velocity profiles, the ability of the MULES, especially the local variance based StSp and StTe models, to perform well as an SGS model in homogenous turbulence can be established. In the next section, model performance is analysed with respect to a more complex turbulent wake flow around a circular cylinder at a transitional Reynolds number of 3900 - this is the flow on which the 4D-Var<sub>ILES</sub> will be tested and validated.

### 3.3 Wake flow around a circular cylinder at $Re = 3900$

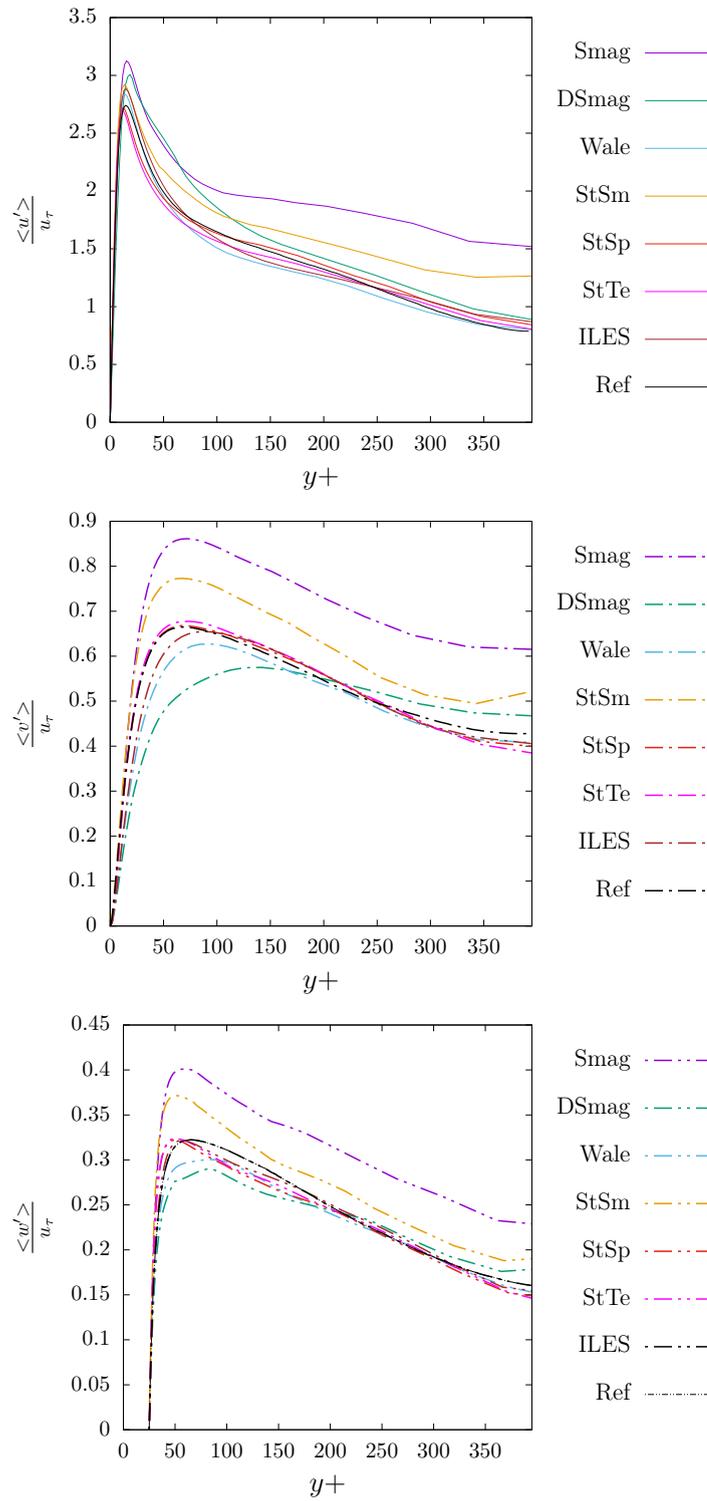
This section is organised as follows: an introduction to the flow and associated literature is provided in §3.3.1. A section on flow configuration and numerical methods (§3.3.2) follows and this section and chapter is concluded with results and discussions in §3.3.3.

#### 3.3.1 Introduction to wake flow

Cylinder wake flow has been studied extensively starting with the experimental works of Townsend (Townsend, 1949a,b) to the numerical works of Kravchenko and others (Singh and Mittal, 2004; Kravchenko and Moin, 2000; Dong and Karniadakis, 2005). The flow exhibits a strong dependance on Reynolds number  $Re = UD/\nu$ , where  $U$  is the inflow velocity,  $D$  is the cylinder diameter and  $\nu$  is the kinematic viscosity of the fluid. Beyond a critical Reynolds number  $Re \sim 40$  the wake becomes unstable leading to the well known von Kármán vortex street. The eddy formation remains laminar until gradually being replaced by turbulent vortex shedding at higher  $Re$ . The shear layers remain laminar until  $Re \sim 400$  beyond which the transition to turbulence takes place up to  $Re = 10^5$  - this regime, referred to as the sub-critical regime, is the focus of this section.

The transitional nature of the wake flow in the sub-critical regime, especially in the shear layers is a challenging problem for turbulence modelling and hence has attracted a lot of attention. The fragile stability of the shear layers leads to more or less delayed roll-up into von Kármán vortices and shorter or longer vortex formation regions. As a consequence significant discrepancies have been observed in near wake quantities, illustrated by U-shaped or V-shaped streamwise velocity profiles, both for numerical simulations (Ma et al., 2000) and experiments (Parnaudeau et al., 2008).

Within the sub-critical regime, 3900 has been established as a benchmark  $Re$  with numerous studies at this value. Rajagopalan and Antonia (2005) suggest that the choice

Figure 3.8: Velocity fluctuation profiles for turbulent channel flow at  $Re_\tau = 395$ .

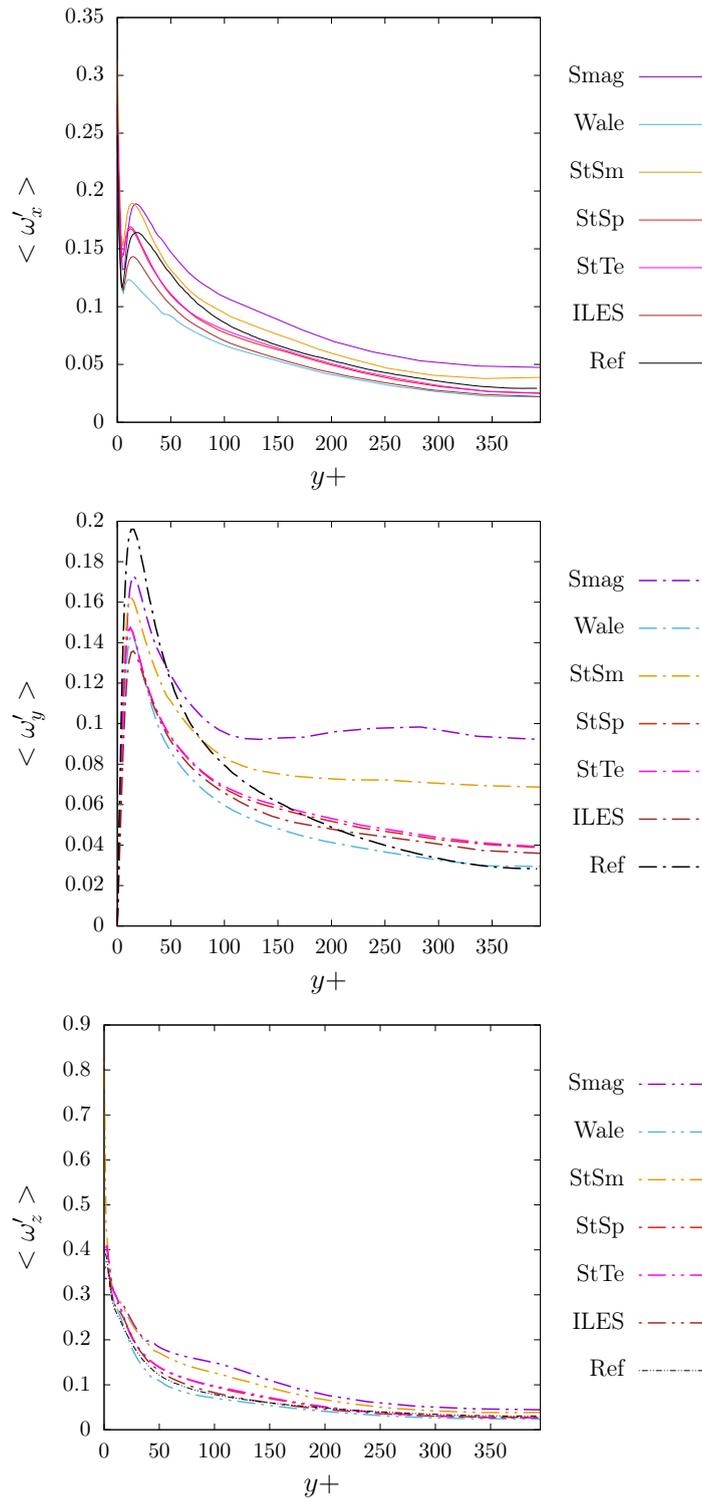


Figure 3.9: Vorticity ( $\Omega$ ) fluctuation profiles for turbulent channel flow at  $Re_\tau = 395$ .

of 3900 is due to a lack of organised structure in the von Kármán vortex street at higher Re. With increasing computational power, LES studies were performed by Beaudan and Moin (1994); Mittal and Moin (1997); Franke and Frank (2002) to name a few, at this Re. The statistical agreement of LES with experimental data was good far downstream of the cylinder while in the near wake, especially within the recirculation region, many anomalies were noted in early studies. In the recirculation region, the experimental statistics of Lourenco and Shih (1993) suggest a V-shaped mean streamwise velocity profile while the works of Beaudan and Moin (1994) showed a U-shaped mean streamwise velocity profile. The discrepancy in the profiles was attributed to inaccuracies in the experiment which was further confirmed by the studies of Mittal and Moin (1997) and Kravchenko and Moin (2000) who also obtained a U-shaped profile. They also showed that the numerical scheme plays an important role in the power spectra with overly dissipative schemes being unsuitable as was the case with Beaudan and Moin (1994) - a fact supported by Breuer et al. (2003). Further information regarding this can be found in Dong and Karniadakis (2005). Franke and Frank (2002) also suggest a dependency of the mean statistics on the averaging period - they identified erroneous results for short averaging periods. The work by Parnaudeau et al. (2008) provided accurate experimental data and LES results showing good agreement with previous numerical studies contrary to early experimental datasets. Their database is thus becoming useful for numerical validation studies. With increasing computation power, the LES data sets at  $Re = 3900$  have been further augmented with DNS studies performed by Ma et al. (2000) which shows good agreement with first and second order experimental statistics both in the near wake and far downstream of the cylinder.

The transition to turbulence nature of the flow combined with the availability of validated experimental and numerical (both LES and DNS) data-sets at  $Re = 3900$  makes this the ideal flow for model development and comparison. A RANS and LES comparison study was performed by Lübecke et al. (2001) with specific focus on explicit algebraic stress models. They found that RANS approaches (steady and unsteady) performed poorly for bluff body flows while LES were statistically accurate. Ouvrard et al. (2010) performed a comparative study on the effects of SGS parameters and grid resolution for LES and variational multi-scale (VMS) methods based LES - they concluded that the statistics had a strong dependence on the SGS model and the use of VMS did not necessarily improve statistics for all SGS models. A similar study for Spectral Vanishing Viscosity (SVV) methods was performed by Pasquetti (2005) who concluded that the arbitrariness in SVV parameters is not trivial and adversely affects the simulation results. The model parameterisation control the turbulent dissipation. A good SGS model should ensure suitable dissipation mechanism. Standard Smagorinsky model (Smagorinsky, 1963) based on an equilibrium between turbulence production and dissipation, has a tendency to overestimate dissipation in general (Meyers and Sagaut, 2006). In transitional flows, where the dissipation is weak, such a SGS model leads to laminar regimes. Different modifications of the model have been proposed to correct this wrong behaviour. As addressed by Meyers and Sagaut (2006), who introduced relevant improvements, the model coefficients exhibit a strong dependency both on the ratio between the integral length scale and the LES filter width, and on the ratio between the LES filter width and the Kolmogorov scale. In this context of SGS models, coarse LES remains a challenging issue.

The motivation for coarse LES is dominated by the general interest towards reduced computational cost which could pave the way for performing higher Re simulations and Data Assimilation (DA) studies - this remains the main focus of this thesis work. DA computational requirements restrict studies to low Reynolds number even with LES. There

has been a lot of focus on the concept of DA recently with the works of Combés et al. (2015); Kato et al. (2015); Yang et al. (2015). However, computational power requirements have limited DA studies to mainly 2D (for example, see the works of Gronsksis et al., 2013; Mons et al., 2016; Foures et al., 2014). Recent incursion into 3D DNS DA by Robinson (2015) has also been restricted to low Reynolds number flows (cylinder wake at  $Re = 300$ ) due to two orders of increase in computational cost. Both the commonly known ensemble DA methods and variational DA methods significantly increase the computational cost of simulations. In order to facilitate DA studies of realistic Reynolds number based flows, the shift to coarser resolution is necessary. An alternative to coarse LES is Detached Eddy Simulation (DES) where regions near solid boundary employ a RANS model while employing a LES construct otherwise. Breuer et al. (2003) performed a comparison study of LES, RANS, and DES and concluded that RANS failed to predict the unsteady characteristics of the flow while DES, which provided an improvement to RANS, showed remarkable deviations from LES results which were accurate. It is worthy of note that while studies exist for Detached Eddy Simulations (DES) of cylinder wake flow (Travin et al., 2000; Breuer et al., 2003), studies of coarse LES are very limited.

With the focus on coarse resolution, this section analyses the performance of LES models for transitional wake flow at  $Re = 3900$ . The models under location uncertainty which were seen to perform well for homogenous turbulence (§3.2) are analysed in depth for their performance at a coarse resolution and compared with classical models. The models are so called as the equations are derived assuming that the location of a fluid parcel is known only up to a random noise i.e. location uncertainty. Here we wish to assess the model skills with respect to a more complex situations (with laminar, transient and turbulent areas) and coarse resolution grids. We provide also a physical analysis of the solutions computed and compare them with classical LES schemes and experimental data. Although the models are applied to a specific Reynolds number, the nature of the flow generalises the applicability of the results to a wide range of Reynolds number from  $10^3 - 10^5$ , i.e. up to the pivotal point where the transition into turbulence of the boundary layer starts at the wall of the cylinder. The goal is to show the ability of such new LES approaches for simulation at coarse resolution of a wake flow in the subcritical regime. Note that recently for the same flow configuration, Resseguier et al. (2017d) have derived the Mulu in a reduced-order form using Proper Orthogonal Decomposition (POD), successfully providing physical interpretations of the local corrective advection and diffusion terms. The authors showed that the near wake regions like the pivotal zone of the shear layers rolling into vortices are key players into the modelling of the action of small-scale unresolved flow on the resolved flow.

In the results and discussion section (§3.3.3), it will be shown that the Mulu are able to capture, in the context of coarse simulation, the essential physical mechanisms of the transitional very near wake flow. This is due to the split of the SGS contribution into directional dissipation and velocity bias. The next section elaborates on the flow configuration and numerical methods used followed by a comparison of the elaborated models and the associated physics in the results section.

### 3.3.2 Flow configuration and numerical methods

The flow was simulated using the parallelised flow solver, Incompact3d. For simulating the circular cylinder in the computational domain, the Immersed Boundary Method (IBM) developed by Gautier et al. (2014) is used to emulate a body forcing (for more details on IBM refer to §3.1). The main advantage of using IBM is the ability to represent the

mesh in cartesian coordinates and the straightforward implementation of high-order finite difference schemes in this coordinate system. The IBM in Incompact3d has been applied effectively to cylinder wake flow by Parnaudeau et al. (2008) and to other flows by Gautier et al. (2014), and Pinto et al. (2011) among others.

The incompressibility condition is treated with a fractional step method based on the resolution of a Poisson equation in spectral space on a staggered pressure grid combined with IBM. While solving the Poisson equation for the stochastic formulation, the velocity bias was taken into account in order to satisfy the stochastic mass conservation constraints of eq. (2.2.40). It can be noted that although the solution of the Poisson equation in physical space is computationally heavy, the same when performed in Fourier space is cheap and easily implemented with Fast Fourier transforms.

The flow over the cylinder is simulated for a Re of 3900 on a domain measuring  $20D \times 20D \times \pi D$ . The cylinder is placed in the centre of the lateral domain at  $10D$  and at  $5D$  from the streamwise domain inlet. For statistical purposes, the centre of the cylinder is assumed to be  $(0, 0)$ . Two mesh resolutions were analysed: fine resolution LES (fLES) of  $481 \times 481 \times 48$  cartesian mesh points and a coarse mesh resolution of  $241 \times 241 \times 48$  is used for the coarse LES (cLES). cLES discretisation has been termed as coarse as this resolution is  $\sim 6.2\%$  the resolution of the reference LES of Parnaudeau et al. (2008) (henceforth referred to as LES - Parn). In terms of Kolmogorov units ( $\eta$ ), the mesh size for the cLES is  $41\eta \times 7\eta - 60\eta \times 32\eta$ . The Kolmogorov length scale has been calculated based on the dissipation rate and viscosity, where the dissipation rate can be estimated as  $\epsilon \sim U^3/L$  where  $U$  and  $L$  are the characteristic velocity scale and the integral length scale. A size range for  $y$  is used due to mesh stretching along the lateral ( $y$ ) direction which provides a finer mesh in the middle (see figure 3.10 for a view of the cartesian stretched mesh). Despite the stretching, the minimum mesh size for the cLES is still larger than the mesh size of particle image velocimetry (PIV) reference measurements of Parnaudeau et al. (2008) (henceforth referred to as PIV - Parn).

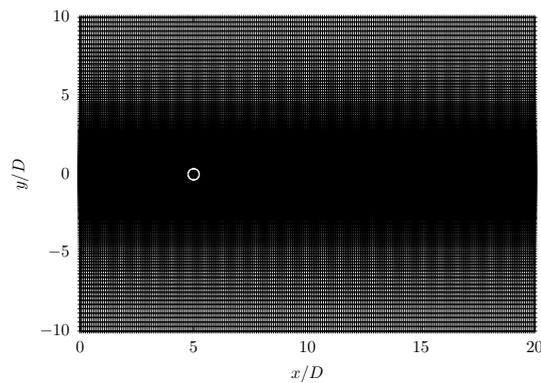


Figure 3.10: Cartesian stretched mesh for cylinder wake flow

For all simulations, inflow/outflow boundary condition is implemented along the streamwise ( $x$ ) direction with free-slip and periodic boundary conditions along the lateral ( $y$ ) and spanwise ( $z$ ) directions respectively - the size of the spanwise domain has been fixed to  $\pi D$  as set by Beaudan and Moin (1994), which was also validated by Parnaudeau et al. (2008) to be sufficient with periodic boundary conditions. The turbulence is initiated in the flow by introducing a white noise in the initial condition. Time advancement is performed using the third order Adams-Bashforth scheme.

A fixed coefficient of 0.1 is used for the classic Smagorinsky and the stochastic Smagorin-

Table 3.5: Model coefficients for wake flow around a circular cylinder at  $Re = 3900$ .

| DNS | Smag | DSmag        | StSm | StSp         |
|-----|------|--------------|------|--------------|
| -   | 0.1  | eq. (2.2.28) | 0.1  | eq. (2.2.47) |

sky model as suggested in literature (Ouvrard et al., 2010) while a spatial neighbourhood of  $7 \times 7 \times 7$  is used for the Stochastic Spatial model. For the dynamic Smagorinsky model, despite the lack of clear homogenous direction, a spanwise averaging is employed. In addition, the constant is filtered and a threshold on negative and large positive coefficients is also applied to stabilise the model. Note that the positive threshold is mesh dependent and needs user-intervention to specify the limits. An attempt at using the StTe model was initially performed, however, due to temporal inhomogeneity of the flow, a very large averaging domain was required to obtain relevant values for the variance tensor. This led to extremely large memory requirements as velocity at each point over a large number of time-steps needed to be stored which was not possible with the computational resources available. Thus, the StTe model is not analysed further in this section. The WALE model was also seen to be unstable for this flow at this coarse resolution due to large rotational strain rate in preliminary studies and shall not be discussed further. The ILES model is also difficult to implement for this flow as the amount of numerical dissipation to be added is user-fixed and requires time-consuming iterative work to fix. Thus, the MULES variants (StSm and StSp) along with the classical Smag and DSmag models are compared. The model coefficients are tabulated in table 3.5

The reference PIV was performed with a cylinder of diameter 12 mm and 280 mm in length placed  $3.5D$  from the entrance of the testing zone in a wind tunnel of length 100 cm and height 28 cm. Thin rectangular end plates placed 240 mm apart were used with a clearance of 20 mm between the plates and the wall. 2D2C measurements were carried out at a free stream velocity of  $4.6 \text{ m s}^{-1}$  ( $Re \sim 3900$ ) recording 5000 image pairs separated by  $25 \mu\text{s}$  with a final interrogation window measuring  $16 \times 16$  pixels with relatively weak noise. For more details about the experiment refer to Parnaudeau et al. (2008).

The high resolution LES of Parnaudeau et al. (2008) was performed on Incompact3d on the same domain measuring  $20D \times 20D \times \pi D$  with  $961 \times 961 \times 48$  cartesian mesh points. The simulation was performed with the structure function model of Métais and Lesieur (1992) with a constant mesh size. LES - Parn is well resolved, however, there is a distinct statistical mismatch between LES - Parn and PIV- Parn especially along the centre-line (see figure 3.11a and figure 3.11b). Literature suggests that the wake behind the cylinder at a  $Re \sim 3900$  is highly volatile and different studies predict slightly varied profiles for the streamwise velocity along the centre-line. The averaging time period, the type of model, and the mesh type all affect the centre-line velocity profile. As can be seen in figure 3.11a and figure 3.11b, each reference data set predicts a different profile/magnitude for the streamwise velocity profiles. The DNS study of Ma et al. (2000) does not present the centreline velocity profiles. This provided the motivation for performing a DNS study at  $Re \sim 3900$  to accurately quantify the velocity profiles and to reduce the mismatch between the existing experimental and simulation datasets. The DNS was performed on the same domain with  $1537 \times 1025 \times 96$  cartesian mesh points using Incompact3d with stretching implemented in the lateral ( $y$ ) direction.

From figure 3.11a we can see that the DNS and the PIV of Parnaudeau are the closest match among the data sets while significant deviation is seen in the other statistics. In

Table 3.6: Flow parameters.

|            | Re   | $n_x \times n_y \times n_z$ | $l_x/D \times l_y/D \times l_z/D$ | $\Delta x/D$ | $\Delta y/D$ | $\Delta z/D$ | $U\Delta t/D$ |
|------------|------|-----------------------------|-----------------------------------|--------------|--------------|--------------|---------------|
| cLES       | 3900 | 241×241×48                  | 20×20× $\pi$                      | 0.083        | 0.024-0.289  | 0.065        | 0.003         |
| fLES       | 3900 | 481×481×48                  | 20×20× $\pi$                      | 0.042        | 0.014-0.12   | 0.065        | 0.002         |
| DNS        | 3900 | 1537×1025×96                | 20×20× $\pi$                      | 0.013        | 0.0056-0.068 | 0.033        | 0.00075       |
| PIV - Parn | 3900 | 160×128×1                   | 3.6×2.9×0.083                     | 0.023        | 0.023        | 0.083        | 0.01          |
| LES - Parn | 3900 | 961×961×48                  | 20×20× $\pi$                      | 0.021        | 0.021        | 0.065        | 0.003         |

the rms of the fluctuating streamwise velocity profile, the only other data sets that exist are of Norberg (1998) who performed Laser Doppler Velocimetry (LDV) experiments at  $Re = 3000$  and  $Re = 5000$ . Among the remaining data-sets (LES of Parnaudeau, PIV of Parnaudeau, and current DNS) matching profiles are observed for the DNS and PIV despite a magnitude difference. These curves also match the profiles obtained by the experiments of Norberg in shape, i.e. the similar magnitude, dual peak nature. The LES of Parnaudeau is the only data-set to estimate an inflection point and hence is not considered further as a reference. The lower energy profile of the PIV may be attributed to the methods used for calculating the vector fields which employ a large-scale representation of the flow via interrogation windows similar to a LES resolution (Corpetti et al., 2006). The DNS, however, exhibits a profile similar to other references and a magnitude in between the two LDV experiments of Norberg. Considering the intermediate Reynolds number of the DNS compared to the Norberg experiments, this suggests good convergence and accuracy of the DNS statistics. Note that the cLES mesh is  $\sim 0.46\%$  the cost of the DNS. Table 3.6 concisely depicts all the important parameters for the flow configuration as well as the reference datasets.

Wake flow around a cylinder was simulated in the above enumerated configuration with the following SGS models: Classic Smagorinsky (Smag), Dynamic Smagorinsky (DSmag), Stochastic Smagorinsky (StSm), and Stochastic Spatial (StSp) variance. In accordance with the statistical comparison performed by Beaudan and Moin (1994), first and second order temporal statistics have been compared at 3 locations ( $x = 1.06D$ ,  $x = 1.54D$ , and  $x = 2.02D$ ) in the wake of the cylinder. All cLES statistics are computed (after an initial convergence period) over 90,000 time steps corresponding to 270 non-dimensional time or  $\sim 54$  vortex shedding cycles. The fLES statistics are computed (after an initial convergence period) over 130,000 time steps or 260 non-dimensional time or 52 vortex sheddings. All statistics are also averaged along the spanwise ( $z$ ) direction. The model statistics are evaluated against the PIV experimental data and the DNS for which the data has been averaged over 400,000 time steps corresponding to 60 vortex sheddings. The work of Parnaudeau et al. (2008) suggests that at least 52 vortex sheddings are needed for convergence which is satisfied for all the simulations. In addition, spanwise averaging of the statistics results in converged DNS statistics comparable with the PIV. Both DNS and PIV statistics are provided for all statistical comparison, however, the DNS is used as the principal reference when an ambiguity exists between the two references.

### 3.3.3 Wake flow results

In this section, the model results, performance analysis and physical interpretations are presented in that order. The fLES results along the centreline are presented first, followed by a complete analysis of the cLES results compared with the reference PIV and the DNS.

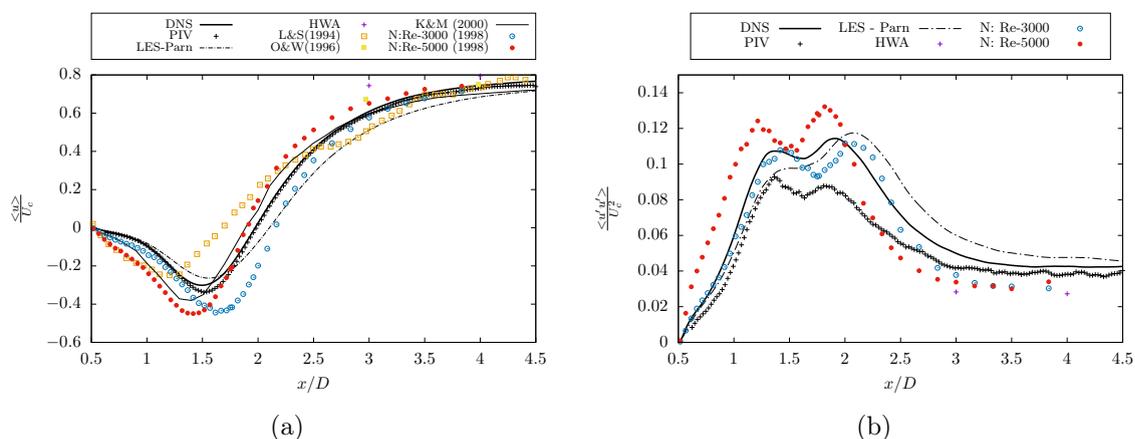


Figure 3.11: Mean streamwise velocity (a) and streamwise rms of the velocity ( $u'u'$ ) fluctuations (b) in the streamwise direction along the centreline ( $y = 0$ ) behind the cylinder for the reference data-sets. Legend: HWA - Hot wire anemometry of Parnaudeau et al. (2008), K&M - B-spline simulations [case II] of Kravchenko and Moin (2000), L&S - Experiment of Lourenco and Shih (1993), N - Experiment of Norberg (1998) at  $Re = 3000$  and  $5000$ , O&W - Experiment of Ong and Wallace (1996)

The focus on centreline results for certain comparisons is to avoid redundancy and focus on the curves with maximum statistical deviation. In addition, for fLES, all results converge to the reference statistics with minimal statistical mismatch for the lateral velocity profiles. Hence, the lateral profiles are only shown for cLES. This is followed by a characterisation and physical analysis of the velocity bias and SGS contributions for the models under location uncertainty. The section is concluded with the computation costs of the different models.

### Fine LES

First, the LES models are simulated with the finer resolution - the results have been compared with PIV - Parn and the DNS. Dynamic Smagorinsky has not been simulated due to restrictive computational cost for fLES. Figure 3.12a compares the streamwise mean velocity profiles along the centreline while figure 3.12b depicts the same for the streamwise root mean square (rms) component. For the mean profile, all profiles match well with the reference with minimal deviation except for a right-side shift in statistics for Smag. In the streamwise rms profile, the two references themselves display significant statistical deviation w.r.t the magnitude of the profile. However, both references indicate a clear dual-peak nature of comparable magnitude. Such a clear dual-peak nature is absent in fLES results, however, StSp appears to match the magnitude of the DNS reference while StSm under-predicts. Smag with a similar magnitude profiles to the DNS, once again displays a right side shift in statistics. This statistical shift for Smag is due to the over-dissipative nature of the model and will be analysed further in the cLES statistics. Model performance at coarser resolution, where LES tend to be inaccurate, is presented below in detail.

### Coarse LES

For cLES, the MULU have been compared with classic and dynamic version of the Smagorinsky model, DNS, and PIV - Parn. Figure 3.13 and figure 3.14 depict the mean streamwise and lateral velocity respectively plotted along the lateral ( $y$ ) direction. In the

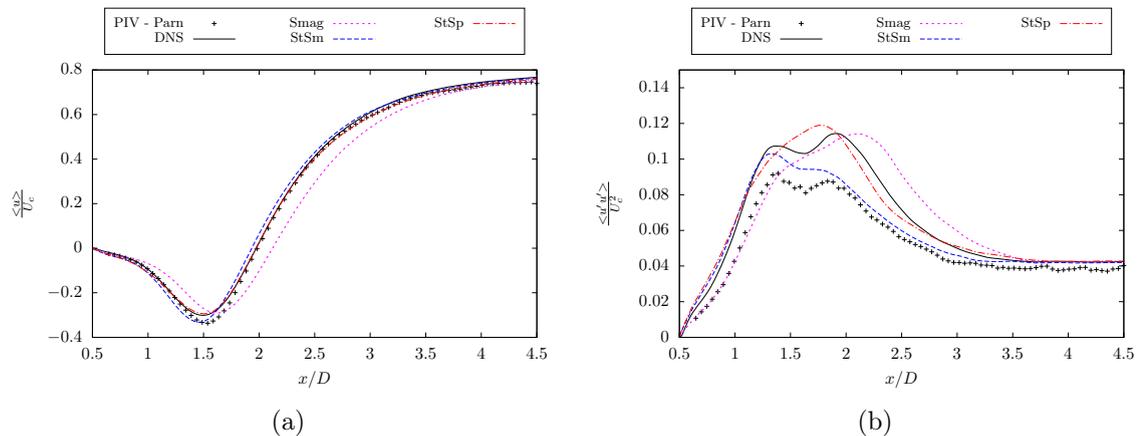


Figure 3.12: Mean streamwise velocity (a) and streamwise rms velocity ( $u'u'$ ) component (b) in the streamwise direction along the centreline ( $y = 0$ ) for fLES.

mean streamwise velocity profile (see figure 3.13), the velocity deficit behind the cylinder depicted via the U-shaped profile in the mean streamwise velocity is captured by all models. The expected downstream transition from U-shaped to a V-shaped profile is seen for all the models - a delay in transition is observed for Smag model which biases the statistics at  $x = 1.54D$  and  $2.02D$ . For the mean lateral component (see figure 3.14), all models display the anti-symmetric quality with respect to  $y = 0$ . Smag model shows maximum deviation from the reference DNS statistics in all observed profiles. All models but Smag capture the profile well while broadly StSp and DSmag models better capture the magnitude. As a general trend, Smag model can be seen to under-predict statistics while StSm model over-predicts. The relative error for StSp model with DNS reference at the first peak for  $x = 1.54D$  (the point of maximum statistical deviation) is 4.6% as compared to the 6% for StSm and DSmag, and 30% for Smag.

A better understanding of the model performance can be obtained through figures 3.15 - 3.17 which depict the second order statistics, i.e. the rms component of the streamwise ( $\langle u'u' \rangle$ ) and lateral ( $\langle v'v' \rangle$ ) velocity fluctuations and the cross-component ( $\langle u'v' \rangle$ ) fluctuations. The transitional state of the shear layer can be seen in the reference statistics by the two strong peaks at  $x = 1.06D$  in figure 3.15a. The magnitude of these peaks is in general under-predicted, however, a best estimate is given by MULU. DSmag and Smag models can be seen to under-predict these peaks at all  $x/D$ . This peak is eclipsed by a stronger peak further downstream due to the formation of primary vortices (see figure 3.15b) which is captured by all the models.

The maxima at the centreline for figure 3.16 and the anti-symmetric structure for figure 3.17 are seen for all models. Significant mismatch is observed between the reference and the Smag/StSm models especially in figure 3.16a and 3.17a. In all second-order statistics, StSm model improves in estimation as we move further downstream. No such trend is seen for StSp or DSmag models while a constant under-prediction is seen for all Smag model statistics. This under-prediction could be due to the inherent over-dissipativeness of the Smagorinsky model which smooths the velocity field. This is corrected by DSmag/StSm models and in some instances over-corrected by the StSm model. A more detailed analysis of the two formulations under location uncertainty (StSm and StSp) is presented in further sections.

The over-dissipativeness of Smag leads to smoothing of the vector fields seen in the instantaneous 2D velocity contours in figure 3.18. The snapshots for each model have been

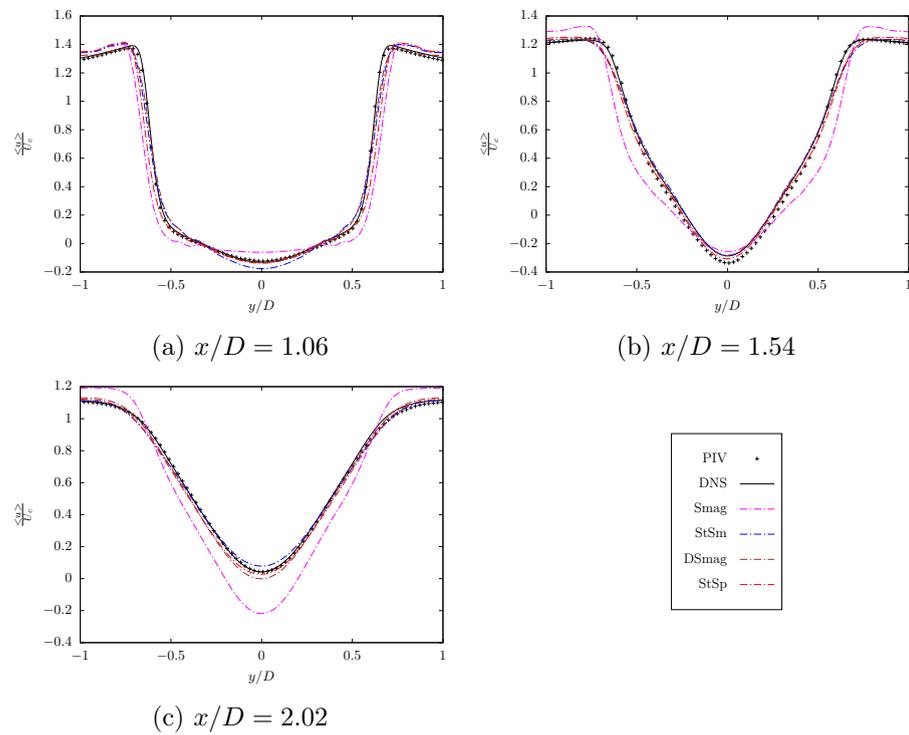


Figure 3.13: Mean streamwise velocity at 1.06D, 1.54D, and 2.02D in the wake of the circular cylinder.

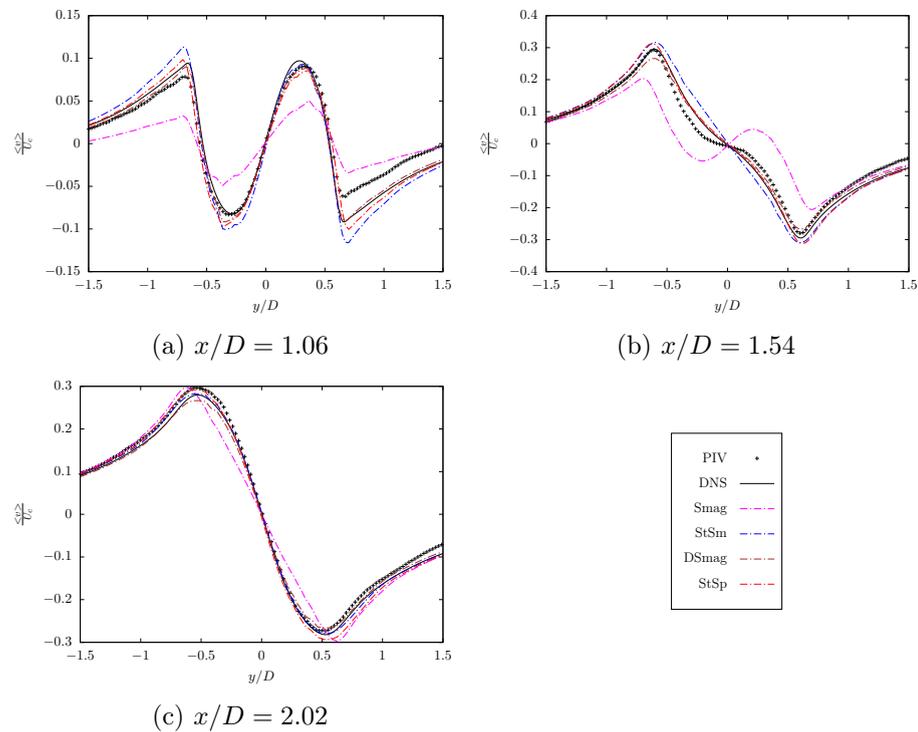


Figure 3.14: Mean lateral velocity at 1.06D, 1.54D, and 2.02D in the wake of the circular cylinder.

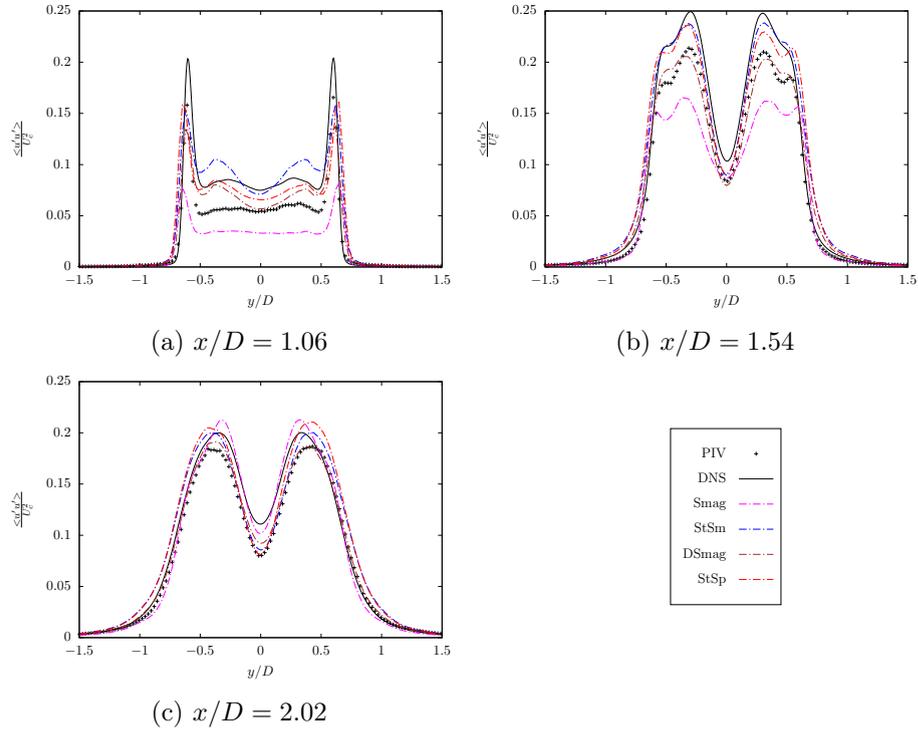


Figure 3.15: Streamwise rms velocity ( $u'u'$ ) fluctuations at 1.06D, 1.54D, and 2.02D in the wake of the circular cylinder.

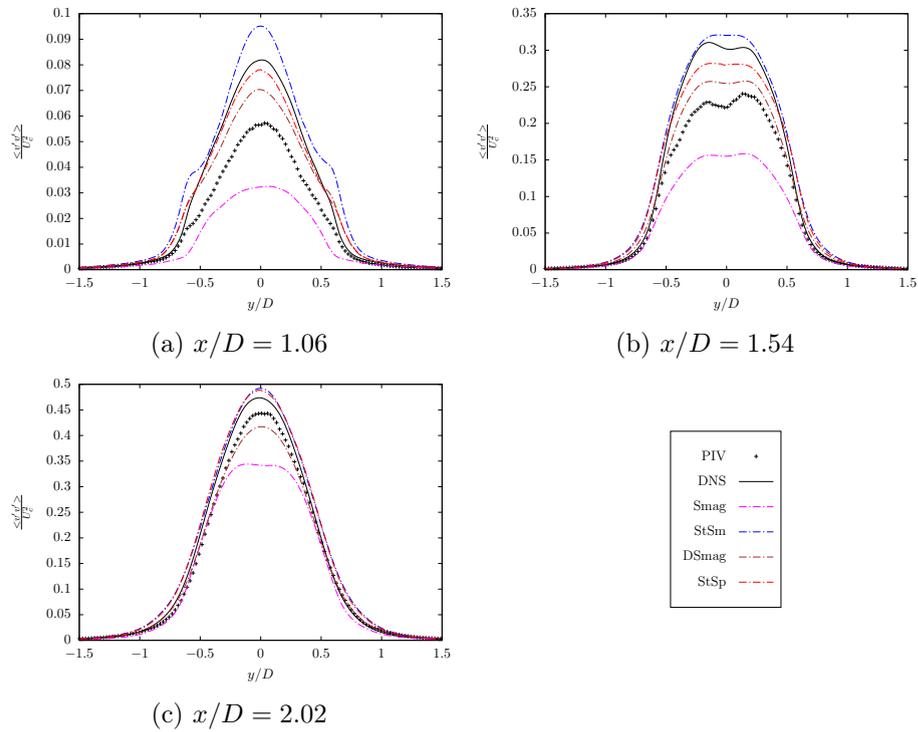


Figure 3.16: Lateral rms velocity ( $v'v'$ ) fluctuations at 1.06D, 1.54D, and 2.02D in the wake of the circular cylinder.

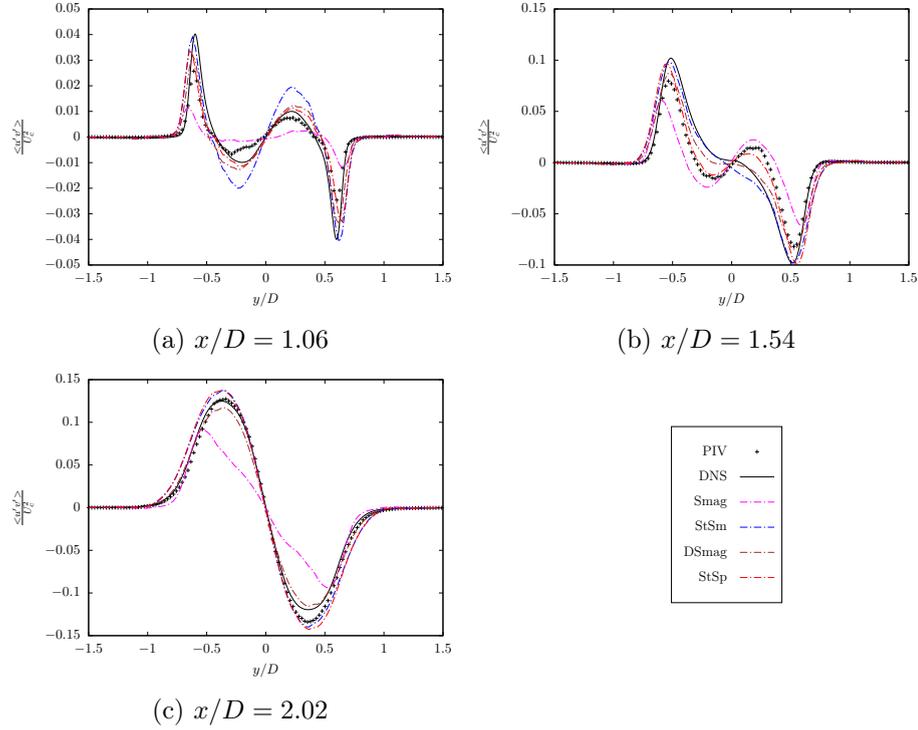


Figure 3.17: Rms velocity fluctuations cross-component ( $u'v'$ ) at 1.06D, 1.54D, and 2.02D in the wake of the circular cylinder.

selected such that the final recorded vortex is being shed from the top half of the cylinder to provide an accurate comparison. In Stochastic Smagorinsky model, more small scale fluctuations are observed as compared to the deterministic variations of the Smagorinsky model. This increase is enhanced in the Stochastic Spatial model where we see a further enhancement in small-scale structures within the flow. The spatial variance model effectively captures the vortex street behind the cylinder along with the small-scale fluctuations in the flow. The DNS contour has a well defined vortex street with maximum small-scale structures. The lengthy shear layer with an influx of back-velocity up to the cylinder wall observed in the DNS is seen in the StSp model with a marginal influx for the DSmag model while it is absent in the Smag and StSm models - this could be the reason for under-shot peaks in rms statistics for the Smag model.

The smoothing for each model is better observed in the 3D iso-contours of vorticity ( $\Omega$ ) plotted in figure 3.19. Plotted at non-dimensional  $\Omega = 7$ , the iso-contours provide an understanding of the dominant vortex structures within the flow. While large-scale vortex structures are observed in all flows, the small-scale structures and their spatial extent seen in the DNS are better represented by the MULU. The over-dissipativeness of the Smag model leads to smoothed iso-contours with reduced spatial extent. The large-scale vortex structures behind the cylinder exhibit the spanwise periodicity observed by Williamson (1996) for cylinder wake flow at low  $Re \sim 270$ . Inferred to be due to mode B instability by Williamson, this spanwise periodicity was associated with the formation of small-scale streamwise vortex pairs. It is interesting to observe here the presence of similar periodicity at higher  $Re$  - this periodicity will be further studied in §3.3.3.

A stable shear layer associated with higher dissipation is observed in Smag model with the shear layer instabilities beginning further downstream than the MULU. An accurate

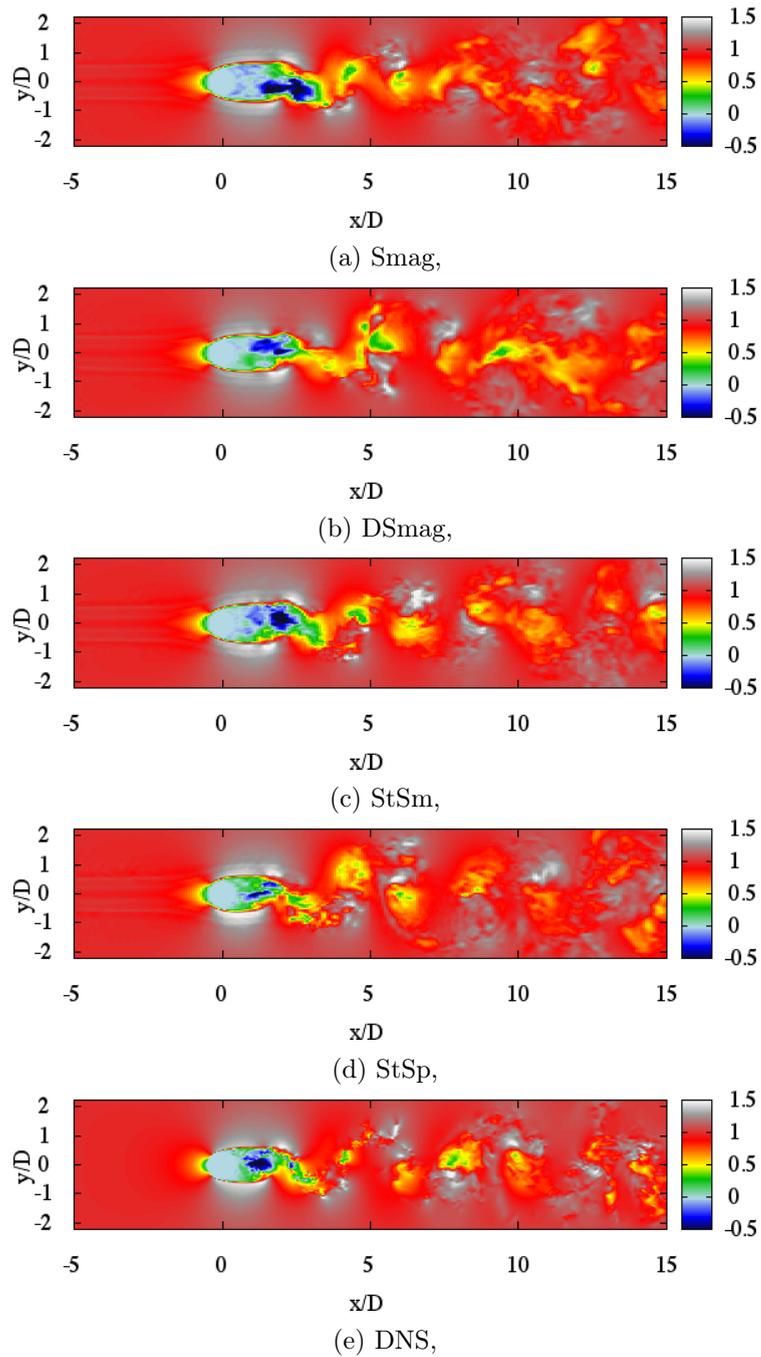
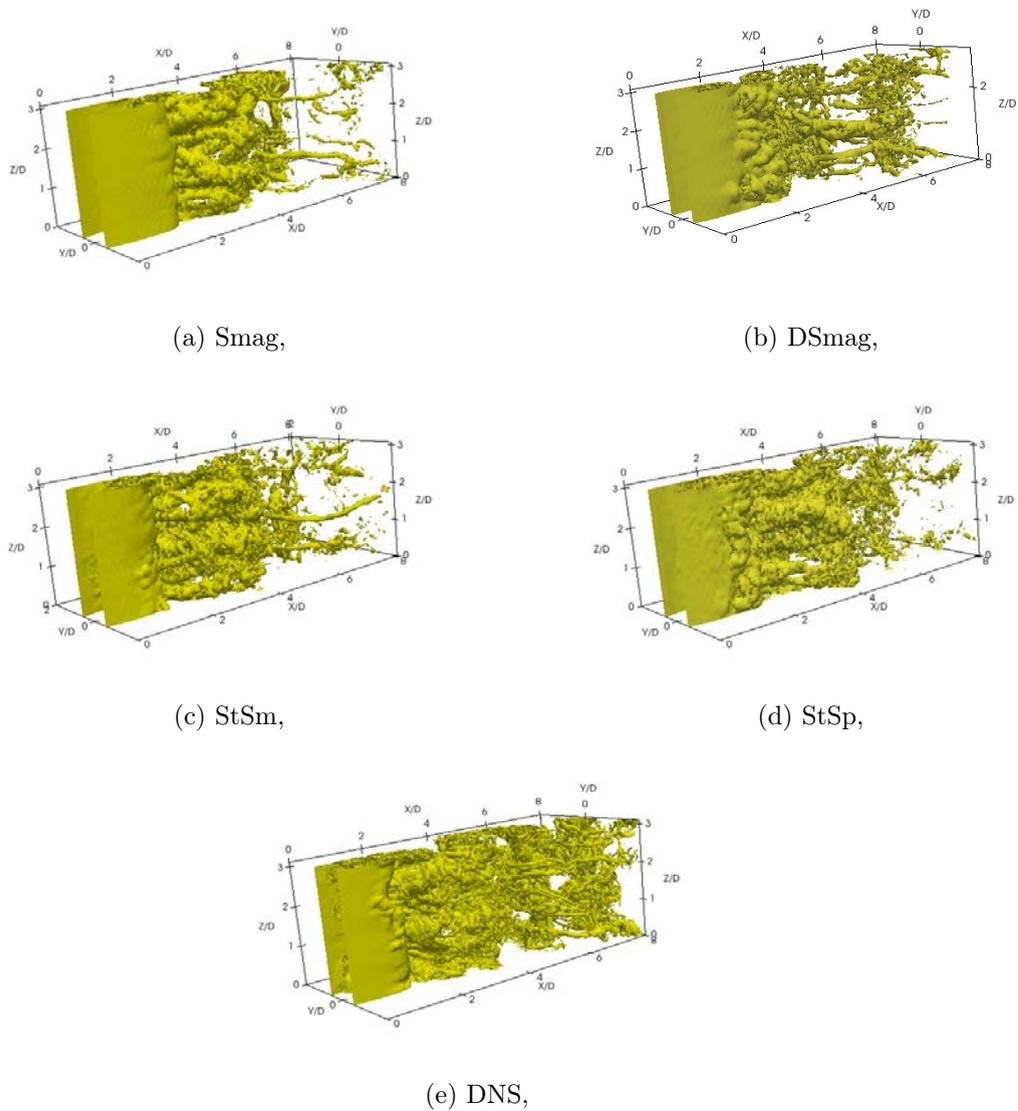


Figure 3.18: Instantaneous streamwise velocity contours behind the cylinder on a perpendicular plane to the cylinder axis.

Figure 3.19: 3D instantaneous vorticity iso-surface at  $\Omega = 7$ .

shear layer comparison can be done by calculating the recirculation length ( $L_r$ ) behind the cylinder. Also called the bubble length, it is the distance between the base of the cylinder and the point with null longitudinal mean velocity on the centreline of the wake flow. This parameter has been extensively studied due to its strong dependence on external disturbances in experiments and numerical methods in simulations (Chyu and Rockwell, 1996; Kravchenko and Moin, 2000). The effective capture of the recirculation length leads to the formation of U-shaped velocity profile in the near wake while the presence of external disturbances can lead to a V-shaped profile as obtained by the experiments of Lourenco and Shih (1993). Parnaudeau et al. (2008) used this characteristic to effectively parameterise their simulations.

The instantaneous contours can provide a qualitative outlook on the recirculation length based on shear layer breakdown and vortex formation. However, in order to quantify accurately the parameter, the mean and rms streamwise velocity fluctuation components

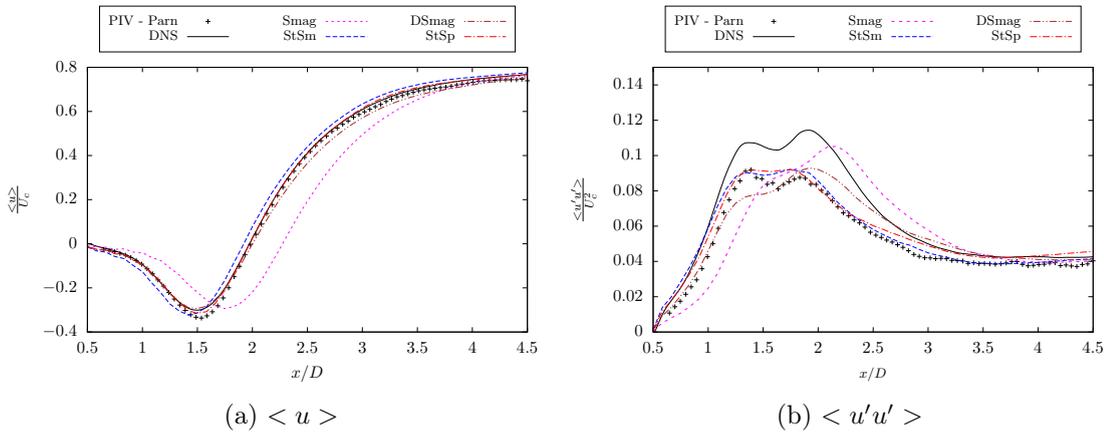


Figure 3.20: Mean streamwise velocity (a) and streamwise rms velocity ( $u'u'$ ) fluctuations (b) in the streamwise direction along the centreline behind the cylinder.

were plotted in the streamwise ( $x$ ) direction along the centreline (see figure 3.20a, and figure 3.20b). The recirculation length for each model is tabulated in table 3.7. StSp and DSmag models capture the size of the recirculation region with 0% error while the StSm model under estimates the length by 5.9% and the Smag model over estimates by 15.9%. The magnitude at the point of inflection is accurately captured by all the models (figure 3.20a).

For the rms centreline statistics of figure 3.20b, due to ambiguity between references, the DNS is chosen for comparison purposes. However, the similar magnitude, dual peak nature of the profile can be established, as enumerated before, through both the references. This dual peak nature of the model was also observed in the experiments of Norberg (1998) who concluded that within experimental accuracy, the secondary peak was the slightly larger RMS peak as seen for the DNS. The presence of the secondary peak is attributed to the cross over of mode B secondary structures within a Re regime (in the transitional regime) of regular shedding frequency with minimal undulations of the Kármán vortices along the spanwise direction. At higher Re above the transition range ( $Re > 5 \times 10^3$ ), where the shedding frequency is irregular accompanied by significant spanwise undulations of the streamwise vortices, the primary peak is smoothed into an inflection point. This is similar to the profile obtained for LES-Parn in figure 3.11b despite the simulation being within the transition regime.

The rms of the fluctuating streamwise velocity profile along the centerline for the deterministic Smagorinsky models display an inflection point unlike the references. The Mulu display a hint of the correct dual peak nature while under-predicting the magnitude matching with the PIV's large scale magnitude rather than the DNS. Although the Smag model has a second peak magnitude closer to the DNS, the position of this peak is shifted farther downstream. This combined with the inability of the model to capture the dual-peak nature speaks strongly against the validity of the Smag model statistics. The convergence of the models can be seen clearly by comparing figure 3.20a with figure 3.12a and figure 3.20b with figure 3.12b.

Further analysis can be done by plotting 2D iso-contours of the rms of the streamwise fluctuating velocity behind the cylinder, as shown in figure 3.21. The iso-contours are averaged in time and along the spanwise direction. The profiles show a clear distinction between the classical models and the Mulu in the vortex bubbles just behind the recirculation region. The vortex bubbles refer to the region in the wake where the initial fold-up

of the vortices start to occur from the shear layers. The MULU match better with the DNS iso-contours within this bubble as compared to the Smag or DSmag models. Along the centreline, MULU under-predict the magnitude, as depicted by the lower magnitude dual peaks in figure 3.20b. As we deviate from the centre-line, the match between the MULU and the DNS improves considerably. The mismatch of the iso-contours in the vortex bubbles for the Smag and DSmag models with the DNS suggests that a higher magnitude for the centreline profile is not indicative of an accurate model.

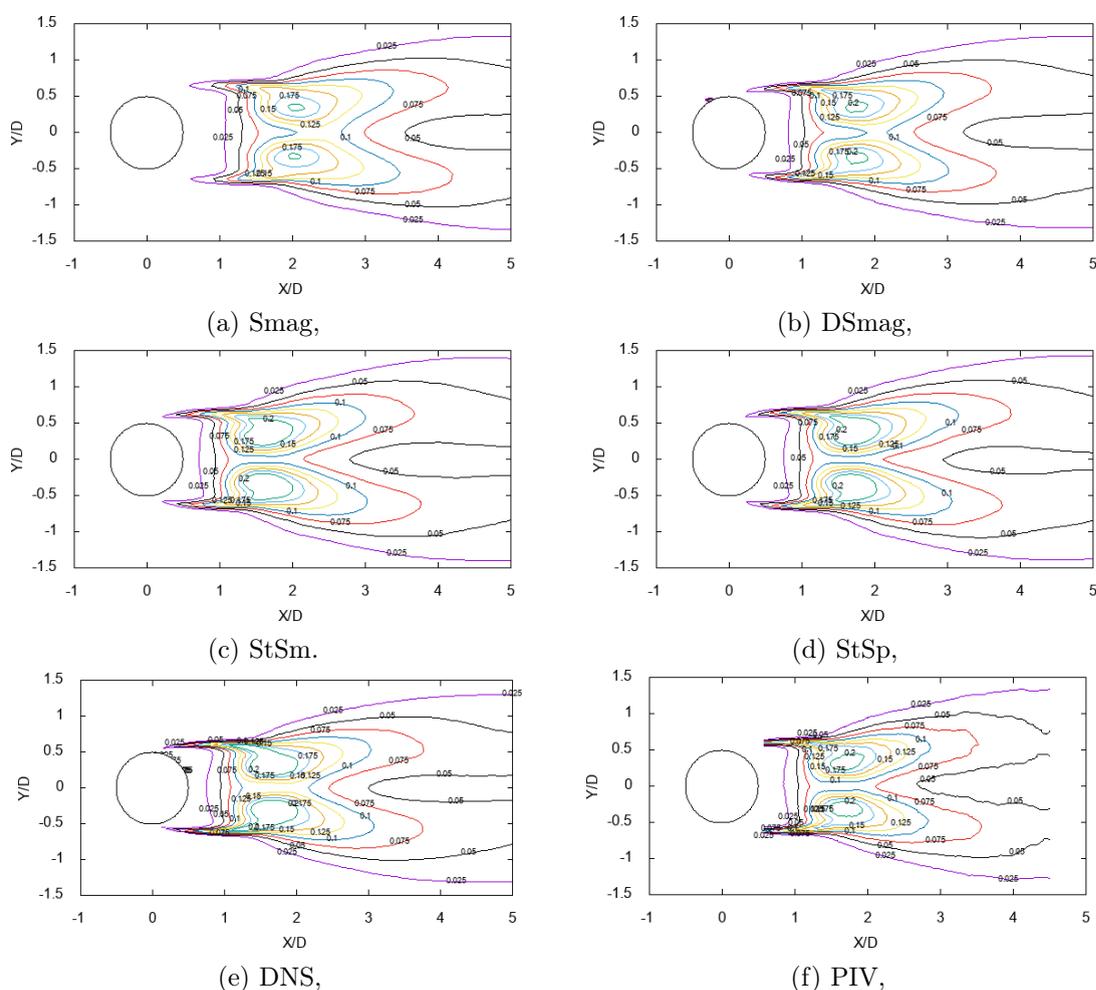


Figure 3.21: 2D iso-contours of the rms of the time averaged fluctuating streamwise velocity ( $u'u'$ ).

The dual peak nature of the streamwise velocity rms statistics show a strong dependence on the numerical model and parameters. This can be better understood via the constant definition within the StSp model formulation (refer to (2.2.47)). The constant requires definition of the scale ( $l_{res}$ ) of the simulation which is similar to  $\Delta$  used in classic Smagorinsky model, i.e. it defines the resolved length scale of the simulation. In the case of stretched mesh, the definition of this  $l_{res}$  can be tricky due to the lack of a fixed mesh size. It can be defined as a maximum ( $\max(dx, dy, dz)$ ) or a minimum ( $\min(dx, dy, dz)$ ) or an average  $(dx \times dy \times dz)^{(1/3)}$ . A larger value of this parameter would signify a coarser mesh (i.e. a rough resolution) while a small value indicates a finer cut off scale or a finer mesh resolution. When  $l_{res}$  is large, corresponding to a "PIV resolution", the centreline

Table 3.7: Recirculation lengths for cLES.

| Model   | PIV - Parn | DNS  | Smag | DSmag | StSm | StSp |
|---------|------------|------|------|-------|------|------|
| $L_r/D$ | 1.51       | 1.50 | 1.75 | 1.50  | 1.42 | 1.50 |

streamwise rms statistics display a dual peak nature with a larger initial peak similar to PIV reference. A smaller value for  $\ell_{res}$ , corresponding to a "higher resolution LES", shifts this dual peak into a small initial peak and a larger second peak similar to the DNS and of higher magnitude. The statistics shown above have been obtained with an  $\ell_{res}$  defined as  $(\max(dx, dy, dz))$  to emulate the coarseness within the model.

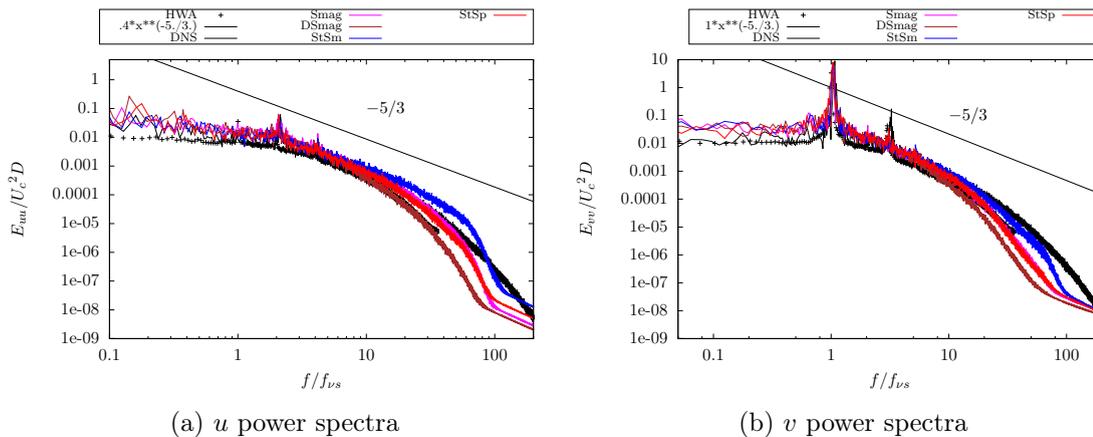


Figure 3.22: Power spectra of streamwise (a) and lateral (b) velocity component at  $x/D = 3$  behind the cylinder.

Figure 3.22a and figure 3.22b show the power spectra of the streamwise and lateral velocity fluctuation calculated over time using probes at  $x/D = 3$  behind the cylinder along the full spanwise domain. For the model power spectra, 135,000 time-steps were considered corresponding to a non-dimensional time of 405 which encapsulates  $\sim 81$  vortex shedding cycles. The Hanning methodology is used to calculate the power spectra with an overlap of 50% considering 30 vortex shedding cycles in each sequence. The reference energy spectra (namely HWA) have been obtained from Parnaudeau et al. (2008) while the DNS energy spectra has been calculated similar to the cLES. The process of spectra calculation for both reference and models are identical. All the values have been non-dimensionalised.

The fundamental harmonic frequency ( $f/f_s = 1$ ) and the second harmonic frequency are captured accurately by all models in the v-spectra. Let us recall that the cLES mesh is coarser than the PIV grid. Twice the vortex shedding frequency is captured by the peak in u-spectra at  $f/f_s \sim 2$  as expected - twice the Strouhal frequency is observed due to symmetry condition at the centreline (Ma et al., 2000). The HWA measurement has an erroneous peak at  $f/f_s \sim 1$  which was attributed to calibration issues and the cosine law by Parnaudeau et al. (2008). All models match with both the reference spectra. One can observe a clear inertial subrange for all models in line with the expected  $-5/3$  slope. The models in the order of increasing energy at the small scales is  $DSmag < Smag = StSp < StSm$ . For the StSm model, an accumulation of energy is observed at the smaller scales in the u-spectra, unlike the StSp model. This suggests that the small-scale fluctuations seen in vorticity or velocity contours for the StSp model (i.e. in figure 3.19) are physical

structures and not a numerical accumulation of energy at the smaller scales known to occur for LES.

The statistical comparisons show the accuracy and applicability of the MULU over classical LES models. The next section focuses on the physical characterisation of the MULU - SGS dissipation, velocity bias and their contributions are studied in detail.

### Velocity bias characterisation

The functioning of the MULU is through the small-scale velocity auto-correlation  $\mathbf{a}$ . The effect of this parameter on the simulation is threefold: firstly, it contributes to a velocity bias/correction which is a unique feature of the MULU. Secondly, this velocity correction plays a vital part in the pressure calculation to maintain incompressibility. Finally, it contributes to the SGS dissipation similar to classical LES models - this signifies the dissipation occurring at small scales. This threefold feature of the MULU is explored in this section.

The contribution of the velocity bias can be characterised by simulating the MULU (StSm and StSp) with and without the velocity bias (denoted by Nad for no advection bias) and comparing the statistics (see figure 3.23a - 3.23b). Only the centre-line statistics have been shown for this purpose as they display maximum statistical variation among the models and provide an appropriate medium for comparison. For the simulations without the velocity bias, the convective part in the NS equations remains purely a function of the large-scale velocity. In addition, the weak incompressibility constraint (2.2.40) is not enforced in the simulations with no velocity bias and the pressure is computed only on the basis of large-scale velocity. Similar to the Smagorinsky model, where the gradients of the trace of the stress tensor are considered subsumed within the pressure term, the divergence of the velocity bias is considered subsumed within the pressure term. The simulation parameters and flow configuration remain identical to cLES configuration.

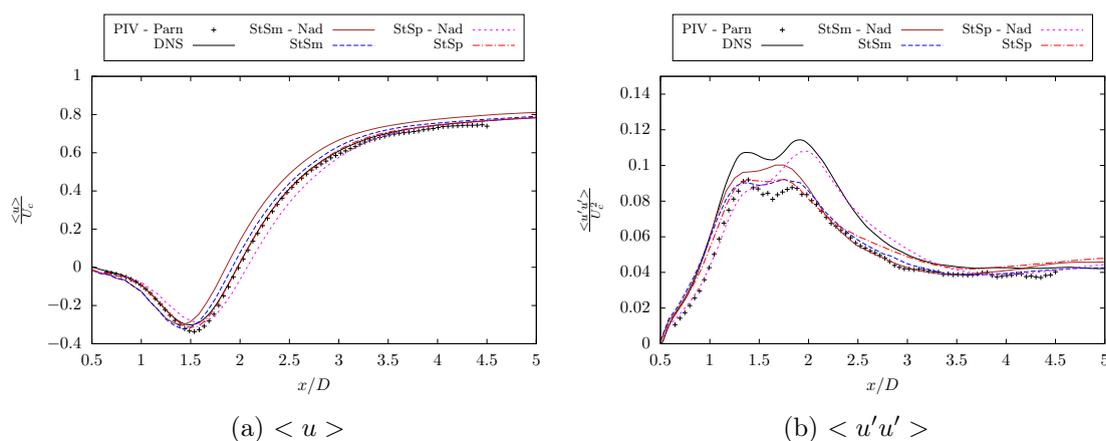


Figure 3.23: Effect of velocity bias on centre-line mean (a) and rms of the fluctuating (b) streamwise velocity behind the cylinder.

The statistics show improvement in statistical correlation when the velocity bias is included in the simulation - all statistical profiles show improvement with inclusion of velocity bias but only the centre-line statistics have been shown to avoid redundancy. In the mean profile, the inclusion of velocity bias appears to correct the statistics for both models to match better with the reference. For the StSm model, there is a right shift in the statistics while the opposite is seen for the StSp model. The correction for the StSp

model appears stronger than that for the StSm model. This is further supported by the rms fluctuation profile where without the velocity bias, the StSp model tends to the Smag model with an inflection point while the inclusion of a connection between the large-scale velocity advection and the small-scale variance results in the correct dual peak nature of the references. For the StSm model, figure 3.23b suggests a reduction in statistical correlation with the inclusion of the velocity bias - this is studied further through 2D iso-contours.

Figure 3.24 plots 2D iso-contours for the rms of the streamwise velocity fluctuations for the MULU. Once again an averaging is performed in time and along the spanwise direction. A clear distinction between the models with and without velocity bias is again difficult to observe. However, on closer inspection, within the vortex bubbles, we can see that including velocity bias improves the agreement with the DNS by reducing the bubble size for the StSm model and increasing it for the StSp model. The higher magnitude prediction along the centreline seen for the StSm - Nad model could be the result of an overall bias of the statistics and not due to an improvement in model performance - the presence of an inflection point in the profile further validates the model inaccuracy. This error is corrected in the model with velocity bias. This corrective nature of the bias is further analysed.

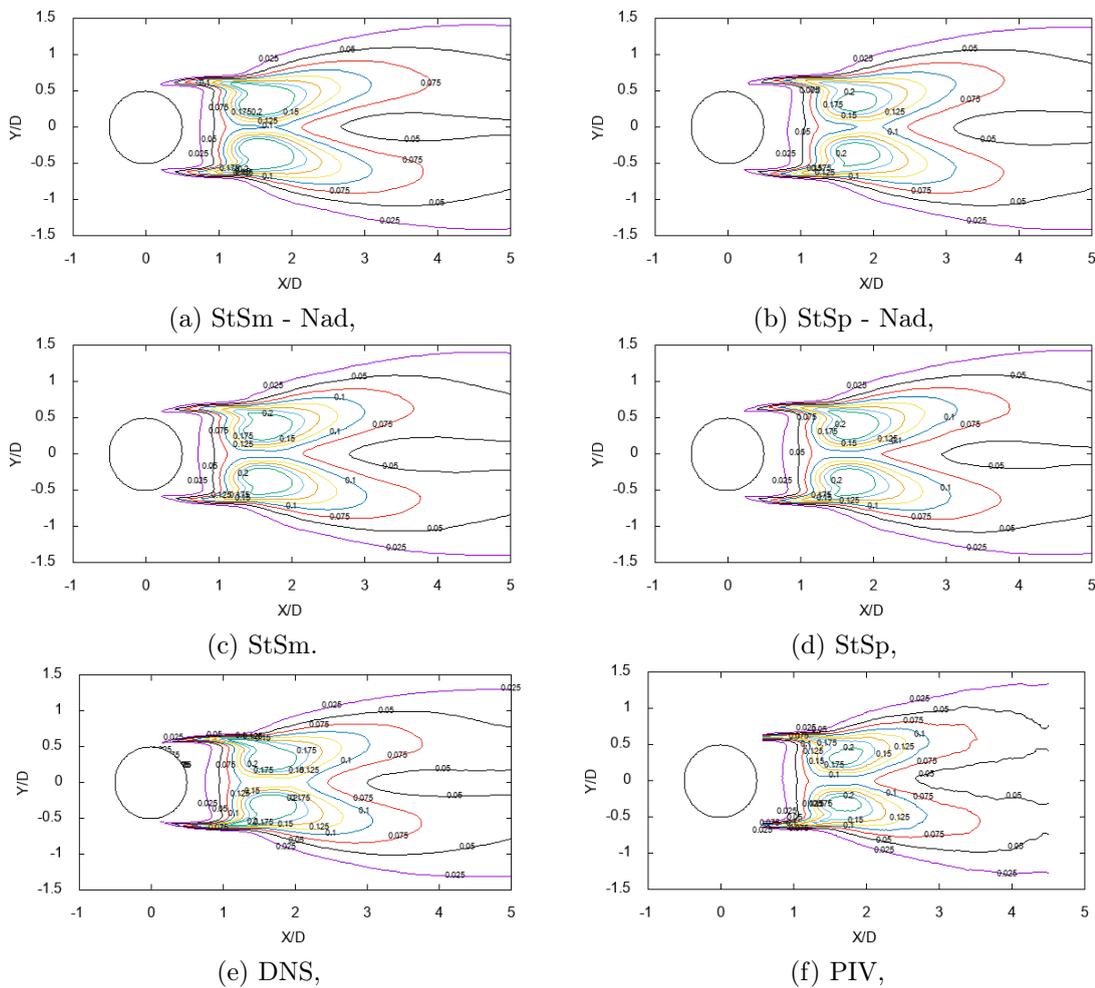


Figure 3.24: Effect of velocity bias on the 2D iso-contour of the rms of the time averaged fluctuating streamwise velocity ( $u'u'$ ).

Table 3.8: Recirculation lengths with and without velocity bias.

| Model   | PIV - Parn | DNS  | StSm - Nad | StSm | StSp - Nad | StSp |
|---------|------------|------|------------|------|------------|------|
| $L_r/D$ | 1.51       | 1.50 | 1.42       | 1.42 | 1.58       | 1.50 |

For the StSp model, the simulation without the bias has a larger recirculation zone or is "over dissipative" and this is corrected by the bias. This result supports the findings of Lesieur et al. (2000) whose structure model, when employed in physical space, applies a similar statistical averaging procedure of square-velocity differences in a local neighbourhood. They found their model to also be over-dissipative in free-shear flows and did not work for wall flows as too much dissipation suppressed development of turbulence and had to be turned off in regions of low three-dimensionality. To achieve that, Ducros et al. (1996) proposed the filtered-structure-function model, which removes the large-scale fluctuations before computing the statistical average. They applied this model with success to large-eddy simulation and analysis of transition to turbulence in a boundary layer. For the StSp model, which also displays this over dissipation quality (without velocity bias), the correction appears to be implicitly done by the velocity bias. Such a velocity correction is consistent with the recent findings of Resseguier et al. (2017d) who provided physical interpretations of the local corrective advection due to the turbulence inhomogeneity in the pivotal region of the near wake shear layers where transition to turbulence takes place. The recirculation length for all cases is tabulated in table 3.8. Data are obtained from the centre-line velocity statistics shown in figure 3.23a. The tabulated values further exemplify the corrective nature of the velocity bias where we see an improved estimation of the recirculation length with the inclusion of the velocity bias. Also, a marginal improvement in statistical match similar to figure 3.23a is observed with the inclusion of the velocity bias for all lateral profiles (not shown here). It can be concluded that the inclusion of velocity bias provides, in general, an improvement to the model.

The physical characteristics of the velocity bias (expressed henceforth as  $\mathbf{u}^* = \frac{1}{2}\nabla \cdot \mathbf{a}$ ) are explored further. The bias  $\mathbf{u}^*$ , having the same units as velocity, can be seen as an extension or a correction to velocity. Extending this analogy, the divergence of  $\mathbf{u}^*$  is similar to "the divergence of a velocity field". This is the case in the MULLU where to ensure incompressibility, the divergence free constraints (eqn. (2.2.40)) are necessary. The stability and statistical accuracy of simulations were improved with a pressure field calculated using the modified velocity  $\tilde{\mathbf{u}}$ , i.e. when the weak incompressibility constraint was enforced on the flow. This pressure field can be visualised as a true pressure field unlike the Smagorinsky model where the gradients of the trace of the stress tensor are absorbed in an effective pressure field.

This can be seen through the pressure contours, shown in figure 3.25, where an improved pressure field is observed for the stochastic models. The pressure distribution, range and offset are better estimated for the MULLU compared to classic or dynamic Smagorinsky where a strong damping of the pressure is observed. The presence of directional dissipation in the StSp model further improves the estimate. However, an offset of 1 non-dimensional pressure is still observed even for StSp compared with DNS.

Stretching the  $\mathbf{u}^*$  and velocity analogy, we can also interpret the curl of  $\mathbf{u}^*$  ( $\nabla \times \mathbf{u}^*$ ) as vorticity or more specifically as a vorticity bias (applying a similar terminology to velocity bias). The curl of  $\mathbf{u}^*$  plays a role in the wake of the flow where it can be seen as a correction to the vorticity field. The divergence and curl of  $\mathbf{u}^*$  are features solely of the MULLU and

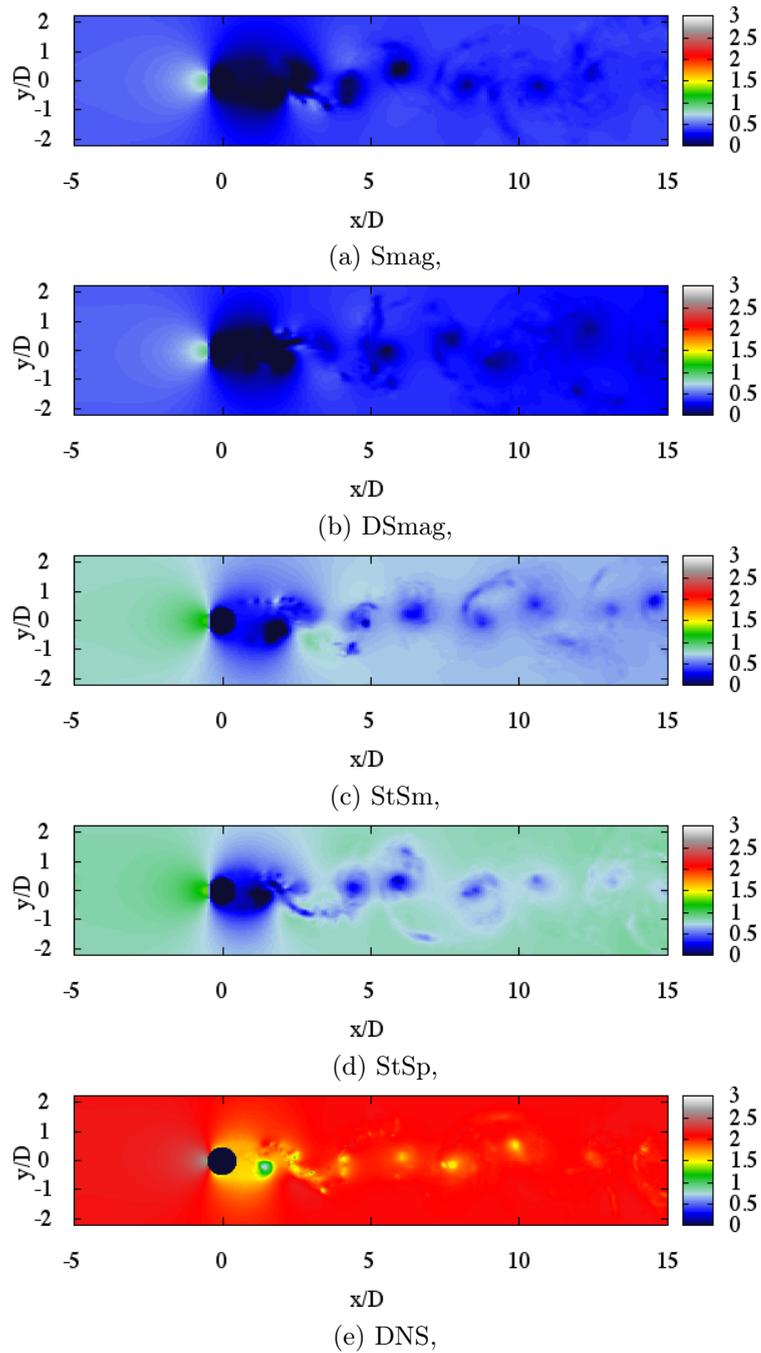


Figure 3.25: Instantaneous pressure contours on a perpendicular plane to the cylinder axis.

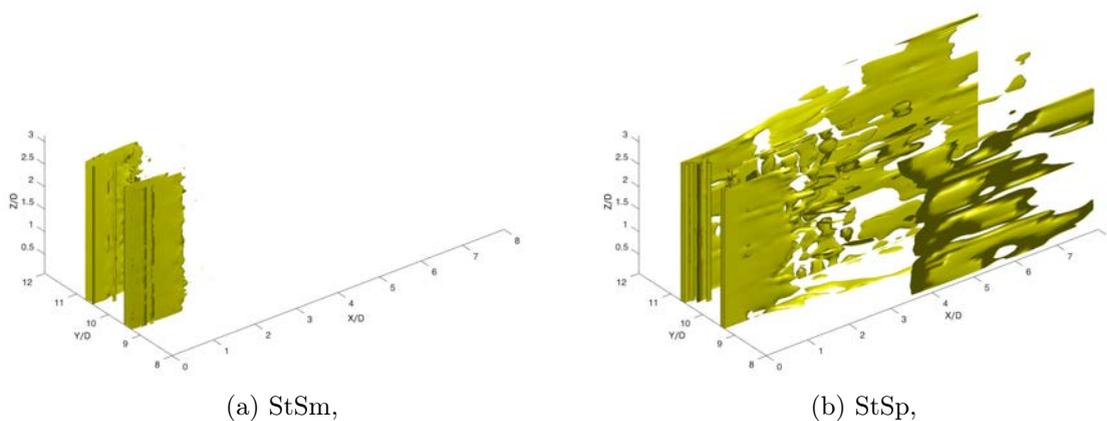


Figure 3.26: 3D iso-contour of the divergence of  $\mathbf{u}^*$  at  $\nabla \cdot (\mathbf{u}^*) = 0.02$ .

their characterisation defines the functioning of these models.

Figure 3.26 depicts the mean iso-contour of  $\nabla \cdot (\mathbf{u}^*) = 0.02$  for the two MULU. This divergence function is included in the Poisson equation for pressure calculation in order to enforce the weak incompressibility constraint. In the StSm model the contribution is strictly limited to within the shear layer while in the StSp model the spatial influence extends far into the downstream wake. The stark difference in the spatial range could be due to the lack of directional dissipation in the StSm model which is modelled on the classic Smagorinsky model. This modelling results in a constant diagonal auto correlation matrix, and the trace elements simplifying to a laplacian of  $\mathbf{a}$  ( $\Delta \mathbf{a}$ ) for  $\nabla \cdot (\mathbf{u}^*)$ . This formulation contains a no cross-correlation assumption (zero non-diagonal elements in the auto correlation matrix) as well as ignoring directional dissipation contribution (constant diagonal terms provides equal SGS dissipation in all three principle directions). These Smagorinsky like assumptions place a restriction on the form and magnitude of  $\mathbf{u}^*$  which are absent in the StSp model. The existence of cross-correlation terms in  $\mathbf{a}$  for the StSp model results in a better defined and spatially well-extended structure for the divergence.

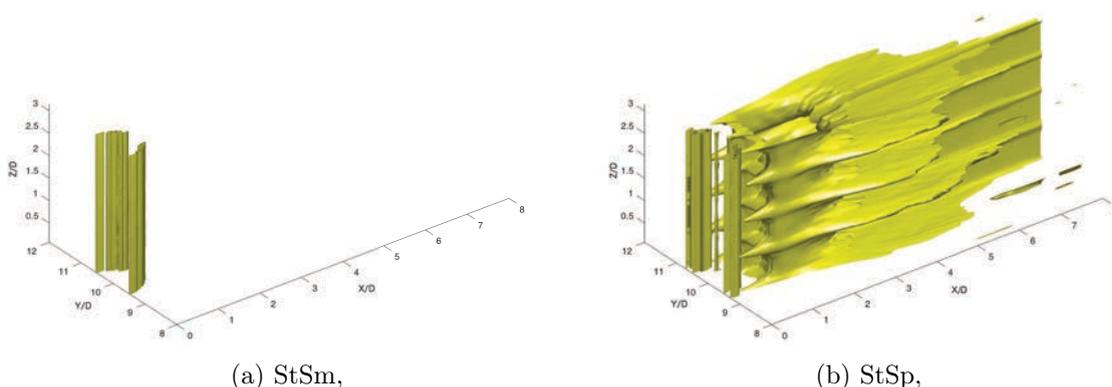


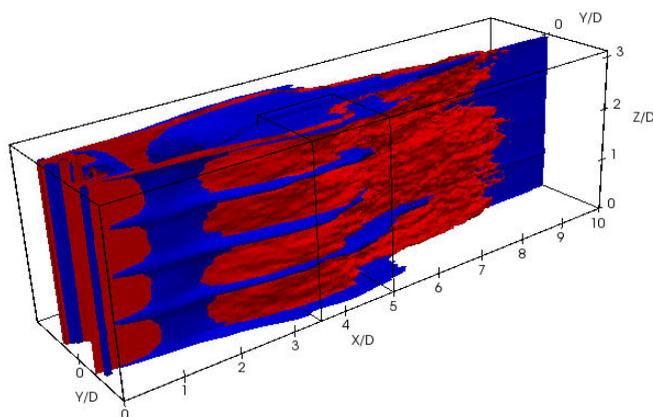
Figure 3.27: 3D iso-contour of the curl of  $\mathbf{u}^*$  at  $\nabla \times (\mathbf{u}^*) = 0.01$  (for StSm),  $0.05$  (for StSp).

The importance of the cross-correlation terms are further amplified in the mean curl iso-contour of  $\mathbf{u}^*$  (see figure 3.27) where once again a spatial limitation is observed for

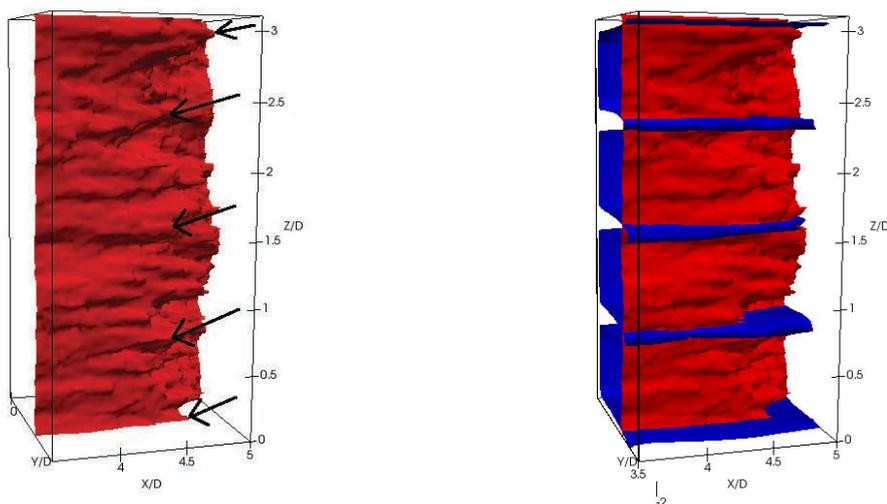
the StSm model. However, the more interesting observation is the presence of spanwise periodicity for the curl of  $\mathbf{u}^*$  observed in the StSp model. The curl parameter is analogous to vorticity and is in coherence with the birth of streamwise vortices seen in figure 3.19, a spanwise periodicity is observed with a wavelength  $\lambda \sim 0.8$ . Figure 3.28 superimposes the iso-contour of the curl of  $\mathbf{u}^*$  for the StSp model at  $\nabla \times (\mathbf{u}^*) = 0.05$  with the mean DNS flow vorticity iso-contour at  $\Omega = 3$ . Figure 3.28a shows the superposition over a large spatial domain behind the cylinder. The outlined box within figure 3.28a indicates the area of zoom for figure 3.28b and figure 3.28c. Figure 3.28b is the zoomed in view of the mean DNS vorticity iso-contour at  $\Omega = 3$ . The arrows in figure 3.28b indicate the points of match between the mean vorticity peaks and the peaks in the mean curl iso-contour of  $\mathbf{u}^*$  for the StSp model. The superposed iso-contours are shown in figure 3.28b. The structural match between the peaks of the two iso-contours are satisfactory with the periodic peaks in the curl of  $\mathbf{u}^*$  matching that of the noisier vorticity iso-contour. While clear periodicity is not observed for the mean vorticity, alternate peaks and troughs can be seen which match with the peaks in the mean curl iso-contour. The wavelength of this periodicity is comparable with the spanwise wavelength of approximately 1D of mode B instabilities observed by Williamson (1996) for  $Re \sim 270$ . The footprint of mode B instabilities is linked to secondary instabilities leading to streamwise vortices observed for  $Re$  ranging from 270 to  $\sim 21000$  (Bays-Muchmore and Ahmed, 1993). These results demonstrate the ability of the spatial variance model to capture the essence of the auto-correlation tensor.

The regions of the flow affected by the auto-correlation term can be characterised by plotting the contours of SGS dissipation density  $((\nabla \mathbf{u})\mathbf{a}(\nabla \mathbf{u})^T)$  of the MULU averaged in time and along the spanwise direction. These have been compared with the dissipation densities for the Smag and DSmag models  $((\nabla \mathbf{u})\nu_t(\nabla \mathbf{u})^T)$  (see figure 3.29). A ‘reference’ dissipation density has been obtained by filtering the DNS dissipation density at the cLES resolution (see figure 3.29e) and by plotting the difference. The StSp model density matches best with the DNS compared with all other models - a larger spatial extent and a better magnitude match for the dissipation density is observed. The high dissipation density observed just behind the recirculation zone is captured only by the StSp model while all Smag models under-predict the density in this region. The longer recirculation zone for Smag model can be observed in the density contours. A few important questions need to be addressed here: firstly, the Smag model is known to be over-dissipative, however, in the density contours, a lower magnitude is observed for this model. This is a case of cause and effect where the over-dissipative nature of the Smag model smooths the velocity field thus reducing the velocity gradients which inversely affects the value of the dissipation density. Secondly, in the statistical comparison only a marginal difference is observed especially between the DSmag and StSp models while in the dissipation density contours we observe considerable difference. This is because the statistical profiles are a result of contributions from the resolved scales and the sub-grid scales. The dissipation density contours of figure 3.29 represent only a contribution of the sub-grid scales, i.e. the scales of turbulence characterised by the model. Thus, larger differences are observed in this case due to focus on the scales of model activity.

Finally, we observe in figure 3.21 that within the vortex bubbles behind the cylinder the MULU perform better than the Smag or DSmag models. For the StSp model, this improvement is associated with the higher magnitude seen within this region in the SGS dissipation density. For the StSm model, no such direct relation can be made with the SGS dissipation density. However, when we look at the resolved scale dissipation  $((\nabla \mathbf{u})\nu(\nabla \mathbf{u})^T)$  for the models (see figure 3.30), a higher density is observed in the vortex bubbles for this model. For the classical models high dissipation is observed mainly in the shear layers. As



(a) Full scale view of the iso-contour superposition with the outlined zoom area,



(b) Zoomed in view of the DNS mean vorticity iso-contour,

(c) Zoomed in view of the iso-contour superposition,

Figure 3.28: 3D iso-contour superposition of the mean curl of  $\mathbf{u}^*$  (blue) for the StSp model at  $\nabla \times (\mathbf{u}^*) = 0.05$  with the mean vorticity (red) for DNS at  $\Omega = 3$ .

the kinematic viscosity ( $\nu$ ) for all models is the same, the density maps are indicative of the smoothness of the velocity gradient. For the classical models, we see a highly smoothed field while for the MULU, we see higher density in the wake. This difference could induce the iso-contour mismatch seen in figure 3.21. These results are consistent with the findings of Resseguier et al. (2017d), who applied the MULU in the context of reduced order model and observed that MULU plays a significant role in the very near wake where important physical mechanisms take place.

For the MULU, the SGS contributions can be split into velocity bias ( $\mathbf{u}\nabla^T(-\frac{1}{2}\nabla \cdot \mathbf{a})$ ) and dissipation ( $\frac{1}{2}\sum_{ij}\partial_{x_i}(a_{ij}\partial_{x_j}\mathbf{u})$ ). Figure 3.31 shows the contribution of the two via 3D iso-contours (dissipation in yellow and velocity bias in red). The contribution of velocity bias is limited for the StSm model as expected while in the StSp model it plays a larger role. Velocity bias in StSp model is dominant in the near wake of the flow especially in and around the recirculation zone. It is important to outline that this is the region where a statistical mismatch is observed for Smag statistics which are possibly corrected by the

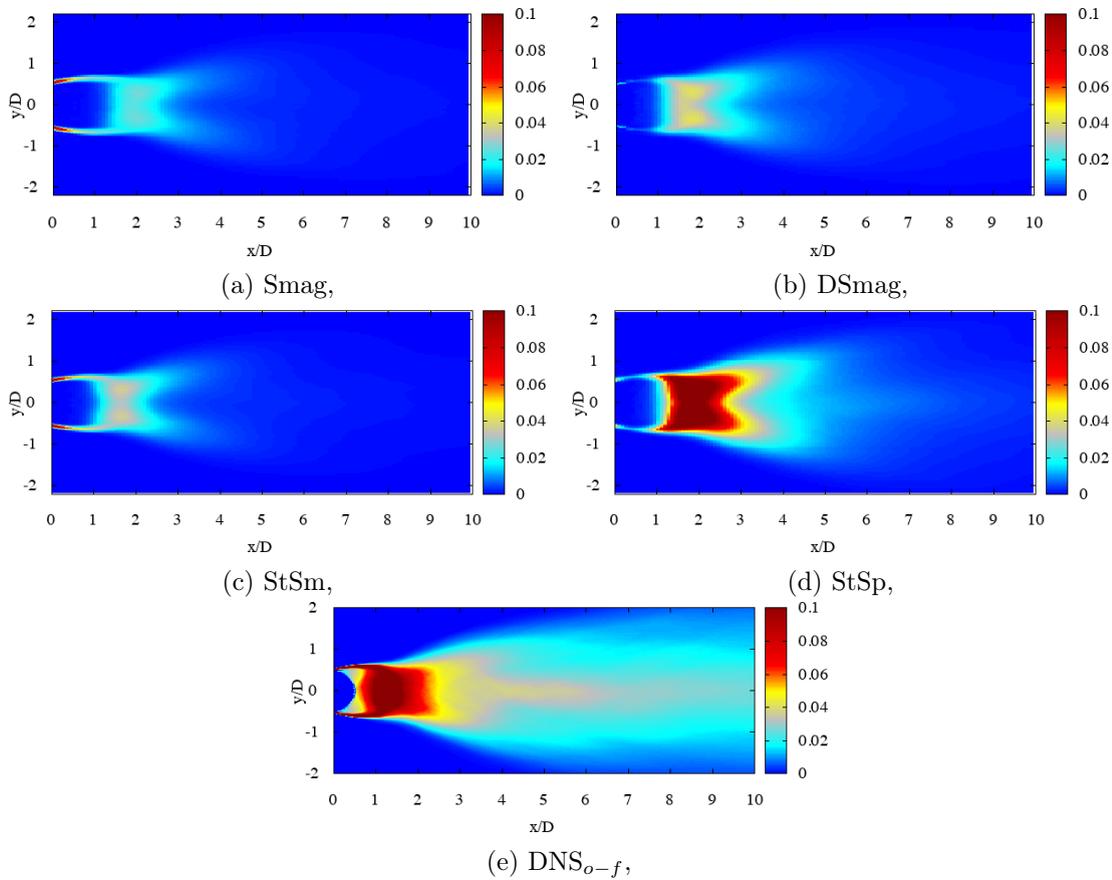


Figure 3.29: Sub-grid scale dissipation density in the wake of the cylinder.  $o$  stands for the original DNS dissipation and  $f$  stands for filtered (to cLES resolution) DNS dissipation.

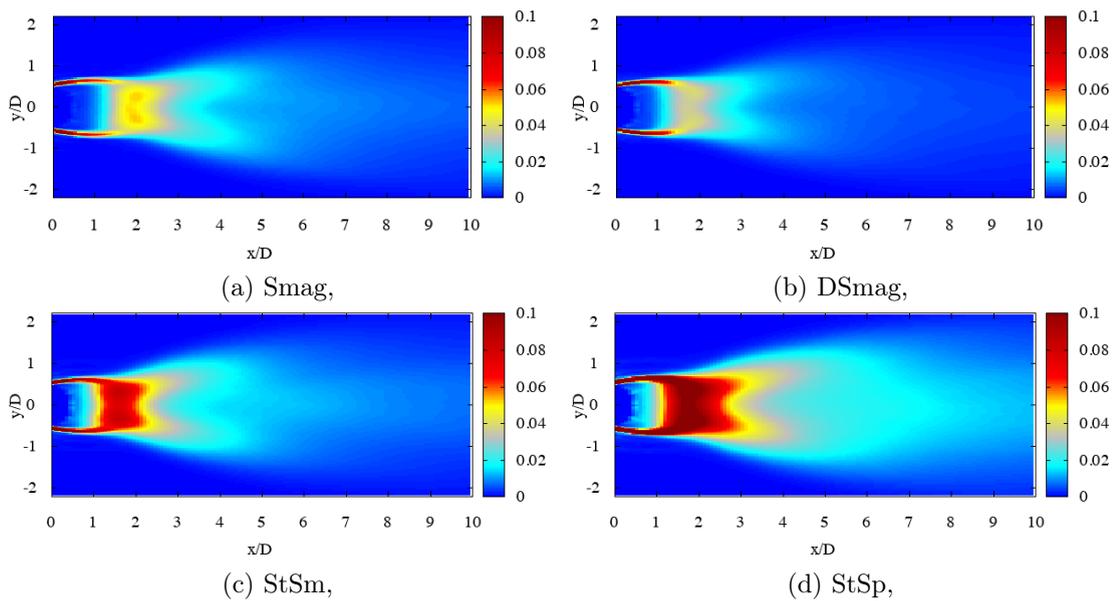


Figure 3.30: Resolved scale dissipation density in the wake of the cylinder for each model.

velocity bias in the StSp model.

Under the context of cLES, the statistical accuracy and stability of the MULU have been

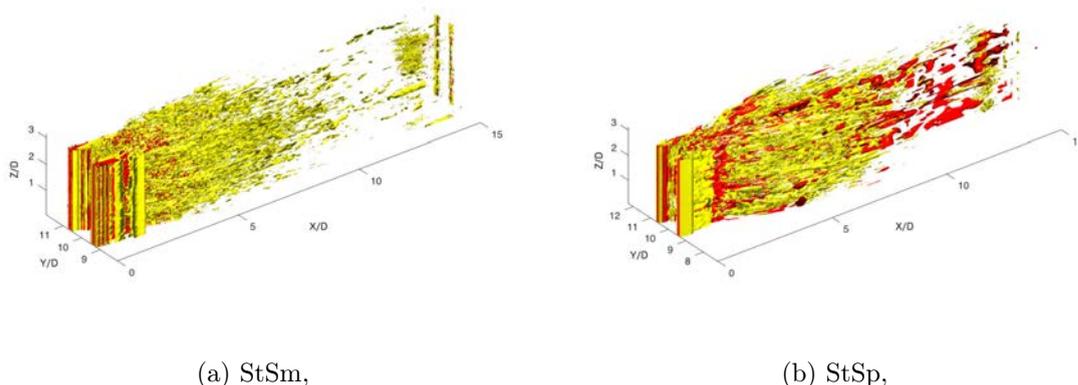


Figure 3.31: 3d SGS contribution iso-surface along the primary flow direction ( $x$ ) with dissipation iso-surface in yellow (at 0.002) and velocity bias in red (at 0.001)

Table 3.9: Computational cost [s] per iteration for each model.

| Model                            | w/o model | Smag  | DSmag | StSm  | StSp  |
|----------------------------------|-----------|-------|-------|-------|-------|
| Computational Cost [s/iteration] | 1.34      | 2.877 | 15.15 | 3.654 | 5.711 |

established with this study to be comparable, if not better, than the dynamic Smagorinsky model. The aim of performing cLES is to reduce the computational cost of simulations and hence opening newer avenues of research. This necessitates a study of the computational requirement for each model (see table 3.9). All simulations were performed with the same hardware and the simulation time, presented in table 3.9, is the time per iteration for each model averaged over two simulations of 50 time-steps each. Without a model LES or under resolved DNS is the fastest as there is no extra model calculations involved. For the cLES, MULU are computationally more expensive than classic Smagorinsky model as they involve calculation of both SGS dissipation and the velocity bias terms - this leads to a marginally increased cost for the StSm model compared to the Smag model. The StSp model is roughly twice ( $\sim 1.98$ ) as expensive as Smag models due to statistical variance calculations for  $\mathbf{a}$  performed over a local neighbourhood. However, the slight increase in computational cost is off-set by the statistical improvements of the StSp model over the Smag model. The cost of performing the StSp model is 37% the cost of performing the DSmag model. This highlights the most important feature of the StSp model - the model captures accurately the statistics at only 0.37 the cost of performing the DSmag model. However, it is important to note that while the computational cost for the Smagorinsky models stay fixed despite changes in model parameters, the cost for the StSp model strictly depends on the size of the local neighbourhood used. A smaller neighbourhood reduces the simulation cost but could lead to loss of accuracy and vice versa for a larger neighbourhood. The definition of an optimal local neighbourhood is one avenue of future research that could be promising. StSm model, which also provides comparable improvement on the classic Smag model, can be performed at 24% the cost of the DSmag model. Thus, the MULU provide a low cost (greater than 2/3rd reduction) alternative to the dynamic Smagorinsky model while improving the level of statistical accuracy.

### 3.4 Concluding remarks

In this chapter, two turbulent flows, namely homogenous turbulent channel flow and wake flow around a circular cylinder, was simulated in a coarse mesh construct using various LES models. The aim was to identify an optimal LES model which could be implemented in the 4D-Var algorithm for performing DA of high Reynolds number flows.

The preliminary study with channel flow indicated towards the capabilities of the MULU, whose formulation introduces a velocity bias term in addition to the SGS dissipation term, to perform well as an LES SGS model when applied in the deterministic sense. These models were shown to match, if not better, in statistical accuracy, the established classical LES models such as the classic and dynamic Smagorinsky, and the WALE model. The performance of the MULU was seen as a precursor for the model capabilities but an application to a more complex turbulent flow was warranted. This was achieved by simulating wake flow around a circular cylinder in the transitional regime

The wake flow simulations were performed on a mesh 54 times coarser than the DNS study. The simulation resolution is of comparable size with PIV resolution - this presents a useful tool for performing DA where in disparity between the two resolutions can lead to difficulties.

The MULU were compared with the classic and dynamic Smagorinsky model. The MULU were shown to perform well with a coarsened mesh - the statistical accuracy of the spatial variance based model was, in general, better than the other compared models. The spatial variance based model and DSmag model both captured accurately the volatile recirculation length. The 2D streamwise velocity iso-contours of the MULU matched better with the DNS reference than the Smagorinsky models. Additionally, the physical characterisation of the MULU showed that the velocity bias improved the statistics - considerably in the case of the StSp model. The analogy of the velocity bias with vorticity was explored further through divergence and curl functions. The spanwise periodicity observed at low Re in literature was observed at this higher Re with the StSp model through the curl of  $\mathbf{u}^*$  (analogous with vorticity) and through mean vorticity albeit noisily. The SGS contribution was compared with Smagorinsky models and the split of the velocity bias and dissipation was also delineated through iso-contours.

It was shown that the performance of the MULU under a coarse mesh construct could provide the necessary computational cost reduction needed for performing LES of higher Re flows. The higher cost of the StSp model compared with Smagorinsky, is compensated by the improvement in accuracy obtained at coarse resolution. In addition, the StSp model performs marginally better than the currently established DSmag model at just 37% the cost of the DSmag model. This cost reduction could pave the way for different avenue of research such as sensitivity analyses, high Reynolds number flows, etc.

Of particular interest, to this thesis work, is the possible expansion of Data Assimilation studies from the currently existing 2D assimilations (Gronskis et al., 2013) or low Re 3D DNS assimilations (Robinson, 2015) to a more informative 3D LES assimilation at realistic Re making use of advanced experimental techniques such as tomo-PIV and/or reconstruction techniques 6. Also and more importantly, the simplistic definition of the MULU facilitate an easy variational assimilation procedure unlike the dynamic Smagorinsky model whose filtering/thresholding procedures are difficult to linearise and to perform the adjoint which are necessary for variational data assimilation.

To summarise, the StSp model, within the MULU framework, provides an improvement in statistical results for a marginal cost with no ad-hoc requirements or procedures needed for stability. The model coefficient is set on the basis of scaling assumption using DNS

parameters. It is important to note that a tuning of this coefficient could be done to improve further the model capabilities. On a qualitative comparison of instantaneous results, the model is shown to reproduce physical small-scale structures for the same resolution at which classical LES tends to smooth the velocity field. This property of the model can be further enhanced by using the full-stochastic formulation of the Mulu which is analysed in next chapter - the fully stochastic version of the model in DA could pave the way for a truly stochastic, physically meaningful ensemble based DA algorithms. In such a case the dynamical model and the flow evolution would be truly stochastic rather than having an ensemble of realisations obtained by perturbing the initial condition or model parameters as is generally done for ensemble based data assimilation methods such as the ensemble Kalman filter.



## Chapter 4

# Stochastic Simulations - vortex forces and Langmuir circulation

**Focus:** The contents of this chapter are a deviation from the general trend of this thesis in the sense that they do not pertain directly to variational data assimilation. However, there exists a correlation between the analysis of stochastic models in fluid flows and the ability to perform ensemble data assimilation using techniques like ensemble Kalman filter. These DA methods require multiple realisations which are currently produced by perturbing, using random vectors, the initial condition or more recently by perturbing model parameters which seems to generate a more accurate spread of realisations Yang (2014). However, such an ensemble of realisations are not constrained by any physical laws. This would be redressed by using a stochastic model such as Mulu which could produce an ensemble of physically meaningful realisations obeying the laws of fluid dynamics. Towards this, this chapter assesses one of the capabilities of the Mulu model in its stochastic formulation towards structuring fluid flows in the hope that it paves the way for future expansion in to ensemble DA methods.

### 4.1 Introduction

The congregation of floating substance into streaks is a common sight on large bodies of water under favourable wind conditions. These windrows were described by Langmuir as due to a series of counter-rotating vortices aligned with the wind (Langmuir, 1938). Langmuir circulation (LC) occurs when wind induced shear interacts with surface gravity waves and is expressed mathematically by the Craik-Leibovich (CL) flow dynamics model (Craik and Leibovich, 1976). The CL equations parametrise the effect of surface gravity waves by introducing a vortex force ( $\mathbf{u}_s \times \omega$ ) into the (large-scale) evolution equation driving the mean current velocity ( $\bar{\mathbf{u}}$ ). The vortex force, responsible for streak formation, is the curl of the Stokes drift ( $\mathbf{u}_s$ ) with the mean current vorticity ( $\omega$ ). The Stokes drift is defined as the difference between the Lagrangian flow velocity of a fluid parcel and the averaged Eulerian flow velocity at a fixed point.

The CL equations explain the creation of streaks in oceanic flows through the action of vortex force. Such streaks are an important feature in numerous other flows of interest: turbulent boundary layer, channel flow, pipe flow, among many others. The origin and organisation of streaks in wall-bounded flows remains till date an open question of significant import (Jiménez, 2018). Literature provides two conceptual frameworks for the creation of streaks - the organisation of streaks remains less studied. The more established of the two

frameworks is the lift-up theory: one theory for lift-up is due to longitudinal vortices which induce alternative wall normal fluid motion with lift up from the wall and push down on to the wall (Stuart, 1965) while another approach associates hair-pin vortices at the wall to the lift-up motion of the fluid (Smith and Walker, 1995). The second and relatively new framework conceptualises the streak formation as a combination of lift-up, shear, and dissipation using linearised equations (Chernyshenko and Baig, 2005). A different outlook is provided by Jimènez (2018) suggesting a decorrelation between wall effects and streak formation in wall-bounded flows. The formation of streaks is instead credited to the role of shear which gets enhanced by the presence of a wall.

These contrasting theories suggest that the formation and spatial arrangement of streaks in wall-bounded fluid flows is still open for interpretation. This work attempts to explain the formation and structuring of streaks in channel flows (expandable to a turbulent boundary layer) by borrowing and generalising concepts developed to describe wave-current interactions in ocean modelling. This generalisation allows us to emphasise the key role played by the spatial inhomogeneity of the unresolved velocity scales to shape large scale physical structures in the flow.

## 4.2 Mathematical formulations

### 4.2.1 CL model

The Craik-Leibovich (CL) flow dynamics model parametrises the effect of waves by introducing a *vortex force* into a (large-scale) representation of NS equation driving the mean current velocity  $\bar{\mathbf{u}}(\mathbf{x}, t)$ :

$$\partial_t \bar{\mathbf{u}} + (\bar{\mathbf{u}} \cdot \nabla) \bar{\mathbf{u}} = -\nabla p' + \mathbf{u}_s \times \boldsymbol{\omega} + \nu \nabla^2 \bar{\mathbf{u}}, \quad \nabla \cdot \bar{\mathbf{u}} = 0, \quad p' = p + \frac{1}{2}(\mathbf{u}_s \cdot \bar{\mathbf{u}}) \quad (4.2.1)$$

where  $\boldsymbol{\omega} = \nabla \times \bar{\mathbf{u}}$  the curl of the mean current and the pressure,  $p'$ , is a modified pressure that includes the pressure  $p$  as well as a correction term due to Stokes drift  $\mathbf{u}_s$ . For a divergence-free large-scale velocity ( $\bar{\mathbf{u}}$ ) the pressure ( $p'$ ) is the solution of the following Poisson equation,

$$-\nabla^2 \left( p + \frac{1}{2}(\mathbf{u}_s \cdot \bar{\mathbf{u}}) \right) = \nabla \cdot \left( (\bar{\mathbf{u}} \cdot \nabla) \bar{\mathbf{u}} - \mathbf{u}_s \times \boldsymbol{\omega} \right). \quad (4.2.2)$$

The Stokes drift velocity is related to the velocity of the surface waves,  $\mathbf{u}'$ , as per CL theory, by:

$$\mathbf{u}_s = \overline{\left( \int_{t_0}^t \mathbf{u}' ds \right) \cdot \mathbf{u}'},$$

where the over-bar represents a time average over fast variation at fixed Eulerian position. Leibovich demonstrated that for nearly rapid irrotational oscillations and surface waves with small slope ( $\epsilon$ ) the Eulerian mean and velocity  $\bar{\mathbf{u}}$  is related to Lagrangian mean  $\mathbf{u}_L$  through the Stokes drift:

$$\bar{\mathbf{u}} = \mathbf{u}_L - \mathbf{u}_s + O(\epsilon^4). \quad (4.2.3)$$

### 4.2.2 Large-scale drift MULLU

The modelling under location uncertainty of large-scale velocity relies on the decomposition of the fluid velocity in terms of a large-scale drift,  $\mathbf{u}$ , and a small-scale random term

assumed uncorrelated in time at the characteristic time-scale of the large scale drift. The dynamics of the large-scale drift expression is given by,

$$\partial_t \mathbf{u} + \underbrace{\left( \left( \mathbf{u} - \frac{1}{2} \nabla \cdot \mathbf{a} \right) \cdot \nabla \right) \mathbf{u}}_{\text{Modified advection}} - \underbrace{\frac{1}{2} \nabla \cdot \left( \mathbf{a} \nabla \mathbf{u} \right)}_{\text{Mixing by unresolved scales}} + 2\Omega \times \mathbf{u} = -\nabla p + \nu \nabla^2 \mathbf{u}. \quad (4.2.4)$$

This model introduces a large-scale sub-grid term accounting for the mixing of the large-scale velocity by the small-scale (unresolved) velocity and a modified advection that describes the statistical influence of inhomogeneity at the unresolved scales encoded through the matrix valued diffusion tensor,  $\mathbf{a}$ . This tensor, which has the dimension of an eddy-viscosity ( $m^2 s^{-1}$ ), is related to the variance of the small-scale velocity on the Eulerian grid times a decorrelation time.

In the following, the relation of this large-scale model with the CL model is mathematically investigated. To that end, the MULU dynamics are expressed in terms of the modified advection  $\mathbf{u}^* = \mathbf{u} - \frac{1}{2} \nabla \cdot \mathbf{a}$  giving,

$$\begin{aligned} \partial_t \mathbf{u}^* + (\mathbf{u}^* \cdot \nabla) \mathbf{u}^* - \frac{1}{2} \nabla \cdot \left( \mathbf{a} \nabla \mathbf{u}^* \right) + 2\Omega \times \left( \mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a} \right) + \\ \partial_t (\nabla \cdot \mathbf{a}) + (\mathbf{u}^* \cdot \nabla) \frac{1}{2} \nabla \cdot \mathbf{a} - \frac{1}{2} \nabla \cdot \left( \mathbf{a} \nabla \frac{1}{2} \nabla \cdot \mathbf{a} \right) = -\nabla p + \nu \nabla^2 \left( \mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a} \right), \end{aligned} \quad (4.2.5)$$

with a modified pressure gradient,

$$\nabla p' = \frac{1}{\rho} \nabla p + \frac{1}{2} \nabla (\mathbf{u}^* \cdot \nabla \cdot \mathbf{a}) \quad (4.2.6)$$

$$\begin{aligned} = \frac{1}{\rho} \nabla p + (\mathbf{u}^* \cdot \nabla) \frac{1}{2} \nabla \cdot \mathbf{a} + \left( \frac{1}{2} (\nabla \cdot \mathbf{a}) \cdot \nabla \right) \mathbf{u}^* + \\ \mathbf{u}^* \times \nabla \times \left( \frac{1}{2} \nabla \cdot \mathbf{a} \right) + \frac{1}{2} \nabla \cdot \mathbf{a} \times \omega^*, \end{aligned} \quad (4.2.7)$$

and assuming the field  $\nabla \cdot \mathbf{a}$  is stationary (the unresolved scale inhomogeneity is stationary – or corresponds to a temporal mean) equation (4.2.5) reads:

$$\begin{aligned} \partial_t \mathbf{u}^* + (\mathbf{u}^* \cdot \nabla) \mathbf{u}^* - \frac{1}{2} \nabla \cdot \left( \mathbf{a} \nabla \mathbf{u}^* \right) - \frac{1}{2} \nabla \cdot \left( \mathbf{a} \nabla \frac{1}{2} \nabla \cdot \mathbf{a} \right) + 2\Omega \times \left( \mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a} \right) = \\ -\nabla p' + \left( \frac{1}{2} \nabla \cdot \mathbf{a} \cdot \nabla \right) \mathbf{u}^* + \mathbf{u}^* \times \nabla \times \left( \frac{1}{2} \nabla \cdot \mathbf{a} \right) + \frac{1}{2} \nabla \cdot \mathbf{a} \times \omega^* + \nu \nabla^2 \left( \mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a} \right). \end{aligned} \quad (4.2.8)$$

This can be rewritten as

$$\begin{aligned} \partial_t \mathbf{u}^* + \underbrace{\left( \mathbf{u}^* - \frac{1}{2} \nabla \cdot \mathbf{a} \right) \cdot \nabla}_{\text{turbophoresis}} \mathbf{u}^* - \underbrace{\frac{1}{2} \nabla \cdot \left( \mathbf{a} \nabla \left( \mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a} \right) \right)}_{\text{SGS}} = -\nabla p' \\ - \underbrace{\left( 2\Omega + \nabla \times \left( \frac{1}{2} \nabla \cdot \mathbf{a} \right) \right) \times \mathbf{u}^*}_{\text{modified Coriolis}} - \underbrace{2\Omega \times \frac{1}{2} \nabla \cdot \mathbf{a}}_{\text{Coriolis Stokes}} + \underbrace{\frac{1}{2} \nabla \cdot \mathbf{a} \times \omega^*}_{\text{vortex force}} + \nu \nabla^2 \left( \mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a} \right). \end{aligned} \quad (4.2.9)$$

Compared to the original equation, this equation still includes a turbophoresis term, a large-scale dissipation and modified Coriolis term. A vortex force appears together with a modified pressure term. The modification of the Coriolis term shows the velocity  $\mathbf{u}^*$  can be interpreted as a velocity in a rotating frame driven by Earth rotation and the

curl of turbophoresis term – related to the statistical effect of the inhomogeneity at the small-scales.

This equation can be simplified with additional assumption on this inhomogeneity. Due to incompressibility condition,  $\nabla \cdot \mathbf{u}^* = \nabla \cdot (\mathbf{u} + \frac{1}{2} \nabla \cdot \mathbf{a}) = 0$ . Therefore, if the large scale drift is assumed incompressible this implies  $\nabla \cdot \nabla \cdot \mathbf{a} = 0$ . In that case, the summation of the SGS contribution on  $\mathbf{u}^*$  and the turbophoresis term gives:

$$-\frac{1}{2} \nabla \cdot ((\mathbf{a} \nabla) \mathbf{u}^*) - (\frac{1}{2} \nabla \cdot \mathbf{a} \cdot \nabla) \mathbf{u}^* = -\frac{1}{2} \nabla \cdot \nabla \cdot (\mathbf{a} \mathbf{u}^{*T}). \quad (4.2.10)$$

and the turbophoresis SGS term reads:

$$-\frac{1}{2} \nabla \cdot ((\mathbf{a} \nabla) \frac{1}{2} \nabla \cdot \mathbf{a}) = -\frac{1}{2} \nabla \cdot \nabla \cdot (\mathbf{a} \frac{1}{2} (\nabla \cdot \mathbf{a})^T) + \frac{1}{2} ((\nabla \cdot \mathbf{a}) \cdot \nabla) (\frac{1}{2} \nabla \cdot \mathbf{a}). \quad (4.2.11)$$

The last term can be written:

$$\frac{1}{2} ((\nabla \cdot \mathbf{a}) \cdot \nabla) (\frac{1}{2} \nabla \cdot \mathbf{a}) = \frac{1}{2} \nabla \cdot (\|\frac{1}{2} \nabla \cdot \mathbf{a}\|^2) - \frac{1}{2} \nabla \cdot \mathbf{a} \times \nabla \times \frac{1}{2} \nabla \cdot \mathbf{a}. \quad (4.2.12)$$

In that case we get thus :

$$\begin{aligned} \partial_t \mathbf{u}^* + (\mathbf{u}^* \cdot \nabla) \mathbf{u}^* - \frac{1}{2} \nabla \cdot \nabla \cdot (\mathbf{a} \mathbf{u}^T) &= -\nabla(p' + \frac{1}{2} \|\frac{1}{2} \nabla \cdot \mathbf{a}\|^2) + \frac{1}{2} \nabla \cdot \mathbf{a} \times \nabla \times (\frac{1}{2} \nabla \cdot \mathbf{a}) - \\ &\underbrace{(2\Omega + \nabla \times (\frac{1}{2} \nabla \cdot \mathbf{a})) \times \mathbf{u}^*}_{\text{modified Coriolis}} - \underbrace{2\Omega \times \frac{1}{2} \nabla \cdot \mathbf{a}}_{\text{Coriolis Stokes}} + \frac{1}{2} \nabla \cdot \mathbf{a} \times \omega^* + \nu \nabla^2 (\mathbf{u}^* + \frac{1}{2} \nabla \cdot \mathbf{a}). \end{aligned} \quad (4.2.13)$$

If in addition the turbophoresis term is sufficiently smooth (with a negligible curl) – which together with the null divergence, implies that  $\nabla \cdot \mathbf{a}$  is quasi-harmonic, i.e.  $\nabla^2 (\nabla \cdot \mathbf{a}) = 0$ , we obtain:

$$\begin{aligned} \partial_t \mathbf{u}^* + (\mathbf{u}^* \cdot \nabla) \mathbf{u}^* - \frac{1}{2} \nabla \cdot \nabla \cdot (\mathbf{a} \mathbf{u}^T) &= -\nabla(p' + \frac{1}{2} \|\frac{1}{2} \nabla \cdot \mathbf{a}\|^2) - \\ &\underbrace{2\Omega \times (\mathbf{u}^* - \frac{1}{2} \nabla \cdot \mathbf{a})}_{\text{Coriolis + Stokes}} + \underbrace{\frac{1}{2} \nabla \cdot \mathbf{a} \times \omega^*}_{\text{vortex force}} + \nu \nabla^2 \mathbf{u}^*. \end{aligned} \quad (4.2.14)$$

This equation has exactly a CL form with a pressure term

$$\pi = p + \frac{1}{2} \|\mathbf{u}^*\|^2 + \frac{1}{2} \|\nabla \cdot \mathbf{a}\|^2 - \frac{1}{2} \|\mathbf{u}^*\|^2 = p + \frac{1}{2} \|\mathbf{u}\|^2 - \frac{1}{2} \|\mathbf{u}^*\|^2, \quad (4.2.15)$$

accounting for the kinetic energy of the Stokes drift. Compared to the original CL form (Leibovich, 1980), this modified pressure entails an additional Bernoulli term stemming from the incompressibility of the turbophoresis term. This pressure term corresponds exactly to the pressure correction as found in Harcourt (2015); McWilliams et al. (1997); Holm (1996).

This derivation stands under the assumption of a stationary  $\mathbf{a}$  and a quasi-harmonic assumption on the divergence of  $\mathbf{a}$  (i.e. the drift correction) similar to the statistically stationary waves, small wave slope and irrotational Stokes drift assumption in the CL equation. A non-constant  $\mathbf{a}$  implies an inhomogeneous random field in space with a non-zero divergence ( $\delta \cdot \mathbf{a} \neq 0$ ). The quasi-harmonic assumption assumes that the divergence and curl of this  $\delta \cdot \mathbf{a}$  is nearly zero. The derived equation is similar to the CL equation with added dissipation terms, Coriolis term and the modified pressure. The vortex force

of the CL equation is also present in this new formulation with the turbophoretic term  $\frac{1}{2}\delta \cdot \mathbf{a}$  playing the role of Stokes drift ( $\mathbf{u}_s$ ). In addition, this term satisfies the same type of decomposition as the Stokes drift, in terms of Lagrangian ( $\mathbf{u}^*$ ) and Eulerian ( $\mathbf{u}$ ) mean velocities:

$$\mathbf{u} = \mathbf{u}^* - \mathbf{u}_s = \mathbf{u}^* - \frac{1}{2}\nabla \cdot \mathbf{a}, \quad (4.2.16)$$

The inhomogeneous turbophoretic correction term in the LU can thus be mathematically identified as a generalised Stokes drift, henceforth referred to as Ito-Stokes drift. This equivalency is drawn here for geophysical flows - the type of flows for which the CL equation is abundantly used. However, it is applicable to all fluid flows with a stationary  $\mathbf{a}$  and a quasi-harmonic Ito-Stokes drift. Stokes drift in LC can thus be seen as a specific example of this inhomogeneity at the unresolved scales triggered by wind induced waves. By expressing this via the divergence of the variance tensor, we generalise the formulation associated to the specific case of wave induced Stokes drift. Let us note that the harmonic condition on the Ito-Stokes drift, intrinsic to the CL equation (through the small wave slope and irrotational Stokes drift assumption), can be relaxed at the price of a generalised model with a modified friction term, a Coriolis correction and a large-scale sub-grid tensor

### 4.3 Vortex force theory for streak formation

A wall-bounded channel flow with modelled small-scale velocity, in the simplest case, through a stationary variance tensor and an induced quasi-harmonic Ito-Stokes drift would imply the reproduction of the vortex forces mentioned in the CL equation. If these vortex forces are indeed present, then it should be a major contributing factor towards the structuration of the flow into streaks similar to what is observed with LC. In addition, if the Stokes drift can be represented by the variance tensor, the modified advection of the LU framework should drive implicitly the structuring effect of the small scales inhomogeneity in flow simulations (including oceanic ones). There would be then no need to incorporate explicitly the vortex force in the flow governing equations as suggested in numerous ocean models (Harcourt, 2015; McWilliams et al., 1997). As a consequence, the Ito-Stokes drift would be a leading actor, either to encode in numerical simulation or to unveil from the data, the driving role played by the small-scale velocity inhomogeneity in shaping coherent large-scale structures in turbulent flows as first wonderfully intuited by Phillips (1977).

Existing theories contrast on the role of different components of the vortex forces for streak formation in LC. The original work on LC by Craik and Leibovich (1976) equates the vortex force to a buoyancy like force acting normal to the fluid surface. Two different mechanisms are theorised for the creation of such a buoyancy like vortex force: spanwise variation or vertical shear of the Stokes drift. An alternative outlook is presented in Leibovich (1983) where the spanwise component of the vortex force is held responsible for the creation of streaks by forcing the fluid towards planes of maximum  $\mathbf{u}_s$ . Due to continuity, this causes fluid to sink at these focal planes and leads to counter rotating vortices which produce streaks.

Clearly, there is a lack of unified view on the component responsible for streak formation and its effect of the vortex forces. There is also an associated limitation due to the assumption of unidirectional Stokes drift (streamwise) for LC. Within the LU framework, such a limitation is not present and it is possible to obtain a three-component vortex force along all three spatial directions. In the following, we present and analyse analytically each component of the vortex force and their corresponding role in the formation of streaks for

wall-bounded channel flow. The analytical study is followed by a numerical application of the LU framework with the Ito-Stokes drift to large-scale flow simulation.

A high resolution (fully resolved direct numerical simulation) simulation is performed on channel flow in an attempt to identify the signature of vortex forces and their contribution to streak formation. Channel flow is simulated at the friction velocity-based Reynolds number of 590 using the flow solver Incompact3d (Laizet and Lamballais, 2009). The flow domain measures  $2\pi H$  by  $2H$  by  $\pi H$  and is discretised into  $384 \times 257 \times 384$  mesh points with stretching implemented in the wall normal direction (minimum mesh size  $\sim$  one wall unit at the wall). After convergence, 500 flow snapshots are recorded every  $50\delta t$  for analysis. The small-scale information can be extracted by filtering and decomposing the data using proper orthogonal decomposition (POD) method proposed by Sirovich (1987).

The instantaneous velocity fields of channel flow obtained from the high-resolution simulation are decomposed using the snapshot POD method. After removal of the mean (zeroth mode), the combined 3D three component flow field ( $\mathbf{u}'$ ) is decomposed as:

$$\mathbf{u}'(x, y, z, t) = \sum_{i=1}^N \phi_i(x, y, z) \psi_i(t), \quad (4.3.1)$$

where  $N$  is the total number of modes,  $\phi$  refers to the spatial modes in decreasing order of energy from 1 to  $N$ , and  $\psi$  stands for the time coefficients. The variance tensor  $\mathbf{a}$  can now be constructed from the spatial modes containing spatial information of the various scales of flow-motion. By excluding the high energy spatial modes (up to 90%), it is possible to extract the spatial information of only the small scales of motion, which need to be represented by the  $\mathbf{a}$  matrix, as:

$$a_{ij} = \sum_{m=n_s}^{n_f} (\lambda_m \Delta t)^{0.5} \phi_{i,m} \phi_{j,m} (\lambda_m \Delta t)^{0.5}, \quad (4.3.2)$$

where  $\lambda_m$  stands for the singular value of the corresponding  $m^{th}$  mode,  $n_s$  stands for the first small-scale mode and  $n_f$  for the final small-scale mode used to construct the variance tensor. The choice for the first and last small-scale mode is guided by the cumulative energy content of the modes - modes containing 90%-99% of the energy are used. The variance tensor is thus formulated from the small-scale information and its divergence is evaluated. This enables the *a posteriori* computation of the vortex forces.

Consider a fully developed flow in a channel at a friction velocity-based Reynolds number of 590. The action of wall friction combined with the mean flow shear induces spatial inhomogeneities, a causal effect of which is the Ito-Stokes drift. The Ito-Stokes drift combined with the vorticity gives rise to the vortex forces (see figure 4.1). Unlike LC, where only a unidirectional Stokes drift is considered, channel flow gives rise to all three components of the Ito-Stokes drift, each of comparable magnitude (first row, figure 4.1). It is interesting to note that the streamwise component of the Ito-Stokes drift is, comparably, the weakest component. The flow induced shear and the wall friction create inhomogeneities along all directions. In the spanwise and wall-normal direction, we observe such strong inhomogeneities especially near the wall. However, a strong mean flow along the streamwise direction dilutes the streamwise inhomogeneity. For LC, there is a strong source of streamwise inhomogeneity in the surface wind which is lacking for a channel where the source of inhomogeneity is wall friction.

Each resultant component of the vortex force (say along  $x$ ) are related to the corresponding perpendicular components of the Ito-Stokes drift (along  $y$  and along  $z$ ) via a cross

product with the mean vorticity, i.e. oriented along the third perpendicular direction (along  $z$  and along  $y$  respectively). Given the multitude of combinations possible for causing the vortex forces, multiple mechanisms can be theorised for their creation and corresponding effect on the flow and streak formation. Here, we focus on the CL1 and CL2 mechanism of Leibovich (15) for LC and their adaptability to channel flow while also introducing a new mechanism (CL3) on the role of streamwise vortex force.

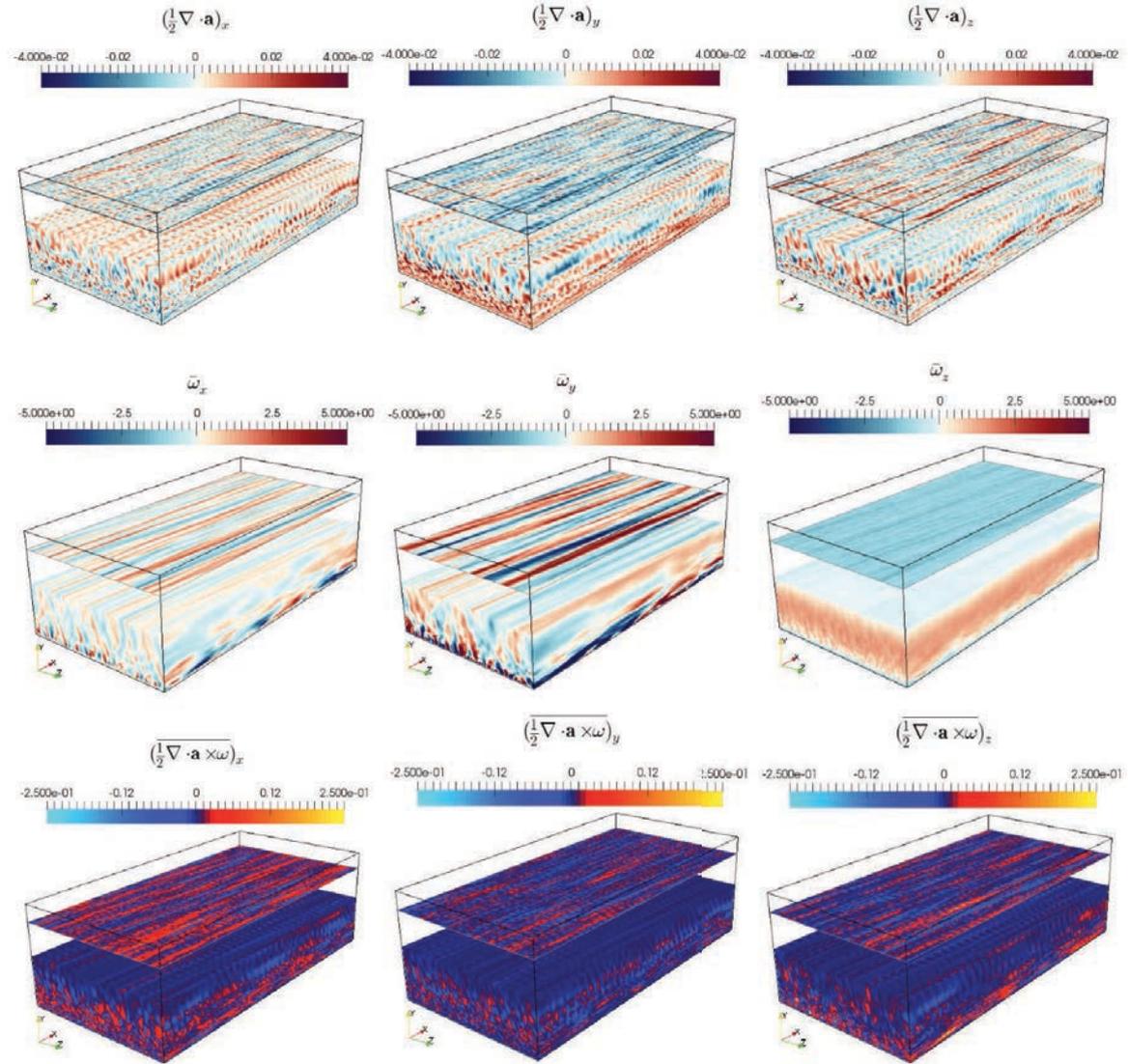


Figure 4.1: The three components of the Ito-Stokes drift (top row), the mean vorticity (middle row), and the resultant vortex force (bottom row) for a channel flow at  $Re_\tau = 590$ , extracted from DNS data.

Under a ‘heuristic discussion’ of vortex force mechanisms, Leibovich (1983) explains the CL1 mechanism: For a flow with unidirectional Stokes drift and a similarly aligned mean velocity (along streamwise direction), the variation along height of the mean velocity induces a spanwise vorticity leading to a vertically (along  $y$ ) oriented vortex force. For a horizontally periodic (along  $z$ ) Stokes drift, the variation of the vortex force produces a torque that drives directly the roll motions responsible for streaks. This theory is applicable to a channel flow whose mean flow variation along height also induces a span-

wise vorticity which combined with the Ito-Stokes drift gives rise to a wall-normal vortex force. On a given XZ plane along the spanwise (z) direction, we observe an alternating structure for the streamwise Ito-Stokes drift and a unidirectional spanwise vorticity. The alternating structure of the Ito-Stokes drift is observed in the wall-normal vortex force. This counter-balancing forces would produce a torque driving the vortex rolls that lead to streak formation, similar to the mechanism described for LC.

The spanwise vortex forces are triggered either by the streamwise component of the Ito-Stokes drift and the wall-normal vorticity as theorised in the CL2 mechanism for LC or ensues from the Ito-Stokes drift normal component and the streamwise vorticity. For LC, the force is believed to act such that the fluid is forced to converge on to spanwise planes of maximum Stokes drift. In these convergence zones on the surface, by continuity, the fluid dives creating regions of downwelling and alternatively regions of upwelling away from these. This up and downwelling leads to the parallel system of vortices in the spanwise direction which is a signature of LC. The CL2 mechanism can be extended to channel flow with respect to the waves on the XZ plane. However, an exact parallel cannot be drawn here as the streamwise Ito-Stokes drift is not unidirectional. Thus, if we consider an alternating form for the streamwise Ito-Stokes drift within the CL2 mechanism, the vortex force is oriented such that it forces the fluid to convergence on to planes of absolute maximum of the Ito-Stokes drift. The spanwise vortex force converges the fluid on to high speed (maximum Ito-Stokes drift) and low speed (minimum Ito-Stokes drift) zones. In these planes, the fluid is forced to converge (the zones separating red and blue streaks of the spanwise vortex force in figure 4.1) and is sheared downstream by the mean flow resulting in streaks. This is seen in the time-averaged spanwise vortex force on the plane at  $y^+ = 100$ , where, in a time-averaged sense, the fluid is acted on by opposing forces (alternative red and blue streaks) to converge on to a line. Such well-defined streaks of alternate forcing are not observed in an instantaneous snapshot of the vortex force (see figure 4.2). However, over-time, there is a strong indication that the spanwise vortex force plays a similar role as conceived in the CL2 mechanism for LC.

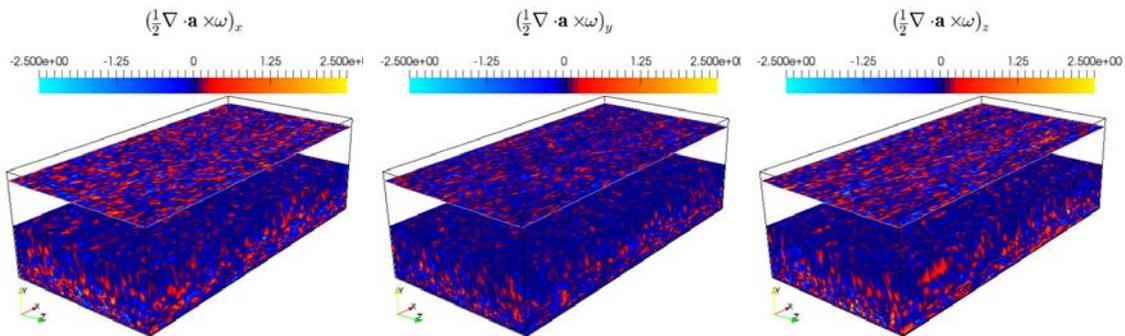


Figure 4.2: 3D instantaneous contours of the vortex force components along the three principle directions. Note the absence of streak like structures and wave pattern in all three components in contrast with the time-averaged contours.

In these convergence zones for LC, only downwelling is observed as the force is acting on the surface. However, for a channel, in these zones, the fluid upon converging could be lifted up towards the middle of the channel or pushed down on to the wall. However, in an averaged sense, the fluid should be lifted up due to the restriction imposed by the wall. This is observed in the wall-normal Ito-Stokes drift component which, near the wall, is predominantly directed away from the wall towards the middle of the channel. This uni-

directional Ito-Stokes drift can combine with the alternating streamwise vorticity resulting in spanwise convergent/divergent vortex forces. These forces can augment the convergence zones created by the CL2 mechanism. Thus, this can be treated as a complementary mechanism to the CL2 mechanism. Within these convergence zones, the fluid is lifted away from the wall and sheared downstream leading to streaks.

In LC with only streamwise Stokes drift, there could be no streamwise vortex forcing. However, with the introduction of the three component Ito-Stokes drift, we obtain a three-component vortex force. Consider the vertical Ito-Stokes drift which near the wall is oriented majorly away from the wall. This combined with the spanwise vorticity of uniform sign gives rise to a streamwise vortex force that is aligned with the flow direction. The variation in strength of the Ito-Stokes drift along the width of the channel leads to a varying strength of the vortex force. On regions with strong streamwise vortex force, the fluid speed is increased leading to high-speed streaks and on regions with weak vortex force, the fluid is acted on only by the mean flow leading to low-speed streaks. This new mechanism, introduced here, is termed as the CL3 mechanism.

We have defined three mechanisms for the vortex forces, each describing the action of one principle component of the vortex force. Within a channel, all three mechanisms are applicable, and could combine together giving rise to the high and low speed streaks observed within the channel: The CL1 mechanism gives rise to the spanwise vortex rolls which is the driving mechanism for streaks in LC. The CL2 mechanism produces convergence zones of fluid due to alternating spanwise forcing where the fluid converges and is forced to move vertically away from the wall due to continuity. This up-welling of fluid is carried downstream by the mean flow and forms into high-speed streaks when augmented strongly by the streamwise vortex forces as described in the CL3 mechanism.

Considering that inhomogeneity arises at the wall, the vortex forces should orient and organise the fluid and create larger and wider streaks as we move away from the wall ? such an organisation is observed in a channel. The vertical extent of organisation does not, however, extend up to the middle of the channel. Larger streaks are formed only up to a certain height and beyond this, they are merely convected downstream by the mean flow. This vertical region of streak formation is now analysed.

There is a limited vertical region of influence (up to  $y^+ \approx 200$  wall units) of the vortex forces extending into the log-law regime of the channel. It is interesting to note that above this spatial limit streaks no longer merge to form larger streaks - we see only an attenuation of the streaks as energy is distributed by the mean flow (such motion streaks are termed as very large-scale motions (VLSMs) by Sharma and McKeon (2013)). In this region, shear predominates over wall effects (Craik and Leibovich, 1976) and the vortex forces play no role. The sweeping up of small-scale streaks into larger and larger streaks is enumerated in Leibovich (1983) as one of the qualitative features of LC. While in LC this is observed on the surface, in a channel, we observe the merging of streaks across the channel half-width effected by the vortex forces.

The spatial influence of the vortex forces can be better analyzed through the spatial wave spectrum. The spectrum provides information on the regions of the flow with active vortex forces and it also determines the dominant component of the vortex force within these active regions. The wave spectrum of each component of vortex force at varying distances from the wall in terms of wall units are shown in figure ???. Each spectrum is calculated along the spanwise line at the given  $y^+$  and averaged along the streamwise direction and in time. In close proximity to the wall ( $y^+ = 5$ ), the vortex-forces play no role, as here the friction forces dominate to create lift-up motion and hairpin vortices. As we move away from the wall, the vortex forces gain potency. At  $y^+$  of 50, all three

vortex forces are strong and are active at all scales of motion. The spanwise component (in yellow) is strongest in the region associated with streak formation, i.e. from  $y^+$  of 15-100. At distances greater than  $y^+$  of 100, shearing occurs and the streamwise component (in blue) becomes stronger. As we approach the middle of the channel, vortex forces once again lose their potency and play no significant role as the pressure forces dominate.

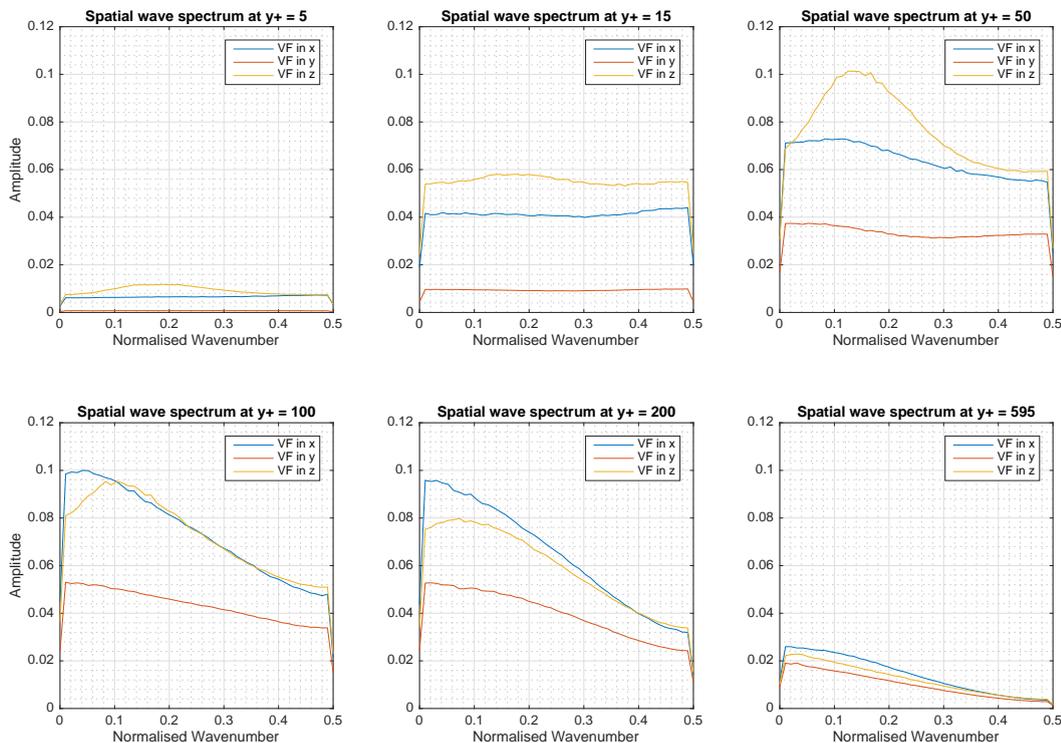


Figure 4.3: 1D FFT in space of the vortex forces along spanwise lines at different heights from the wall averaged along the streamwise direction and in time. Blue line corresponds to streamwise vortex force, red to wall-normal, and yellow to spanwise vortex force.

## 4.4 Large-scale application

To demonstrate the relevance of the previous conclusions, let us consider the large-scale, low-resolution simulations of the channel under three frameworks: first, we implement the MULU in its full stochastic representation with a stationary POD based noise extracted from the high-resolution simulation, second, the MULU are represented in a deterministic sense, with a scale separation principle between the smooth in time large scale component and the highly oscillating Brownian terms, as applied by Chandramouli et al. (2018) using the spatial variance model, and finally, we use the established classical large-scale model of Germano et al. (1991) termed the ‘dynamic Smagorinsky model (DSM)’ to simulate the channel flow. The flow domain is identical to the high-resolution simulation and is discretised into  $96 \times 97 \times 96$  mesh points with stretching implemented in the wall-normal direction.

The full stochastic version of the MULU, contrary to its deterministic version, does

not rely on a scale separation assumption between the small-scale unresolved random component and the large-scale drift. This large-scale velocity is here not necessarily smooth (i.e. of finite variation). The complete dynamics with no scale separation between the two components of velocity reads formally as:

$$\partial_t \mathbf{u} + \left( \mathbf{u} - \frac{1}{2} \nabla \cdot \mathbf{a} \right) \cdot \nabla (\mathbf{u} + \dot{\boldsymbol{\eta}}) - \frac{1}{2} \nabla \cdot ((\mathbf{a} \nabla) \mathbf{u}) + 2\Omega \times (\mathbf{u} + \dot{\boldsymbol{\eta}}) = -\nabla \tilde{p} + \nu \nabla^2 (\mathbf{u} + \dot{\boldsymbol{\eta}}), \quad (4.4.1)$$

where the small-scale random velocity ( $\dot{\boldsymbol{\eta}}$ ) is defined using POD as,

$$\dot{\boldsymbol{\eta}} = \sum_{m=n_s}^{n_f} (\lambda_m \Delta t)^{\frac{1}{2}} \gamma_m \boldsymbol{\phi}_m, \quad (4.4.2)$$

where  $\gamma_m$  is a standard zero mean scalar Gaussian random variable. Note that the pressure in the full stochastic formulation is related both to the large scale and small-scale incompressibility. This stochastic partial differential equation is implemented numerically using the simple Euler-Maruyama scheme with a small time-step.

The flow domain is identical to the high-resolution simulation and is discretised into 96 x 97 x 96 mesh points with stretching implemented in the wall-normal direction. The main differences between the three models need to be highlighted:

- DSM assumes isotropy of the unresolved scales while the two LU versions are capable of modelling isotropic and anisotropic unresolved scales.
- Mulu versions include a modification of the advection by the Ito-Stokes drift which is absent in the purely dissipative DSM.
- The full stochastic Mulu scheme enables small-scale energy injection (balanced by the sub-grid diffusion term) whereas the deterministic Mulu implementation is a purely diffusive (though with anisotropic directions) model that prevents any energy backscattering.

Figure 4.4 plots the streamwise velocity fluctuations at three different heights ( $y^+ = 50, 100, 200$ ) from the wall for all three cases and the reference high-resolution simulation. High speed streaks are represented in red while low-speed streaks are in blue. High intensity, elongated streaks are visible in the high-resolution snapshots with increasing wavelength of streaks as we move farther away from the wall. At  $y^+ = 50$ , the vortex forces are just starting to play a role in structuring the flow. Hence, at this height from the wall, we observe similar structures with all three models and they all compare well with the DNS - the use of a stretched mesh provides a fine resolution this close to the wall and hence most of the physics is inherently resolved. The structures in the DSM, as we move away from the wall, are seen to deteriorate rapidly to become too smooth with wavy structures that are inexistent in the DNS. The presence of small-scale structures in the two Mulu models is an added advantage of the model brought by the back-scattering and/or anisotropic form of the model. As we move farther away from the wall ( $y^+ = 100$ ) into the region where the vortex forces are predominantly active, we observe better structuration in the form of streaks with the inclusion of the Ito-Stokes drift. The DSM model produces wavy regions of low speed flow (blue blobs) within a global high-speed flow. With the Mulu models, well-defined alternating high and slow speed streaks, as observed in the DNS, are visible. This pattern is seen to propagate to  $y^+$  of 200.

The stochastic Mulu scheme differs from the dynamic Smagorinsky model in its ability to include an inhomogeneous diffusion and a modified advection driven by inhomogeneity

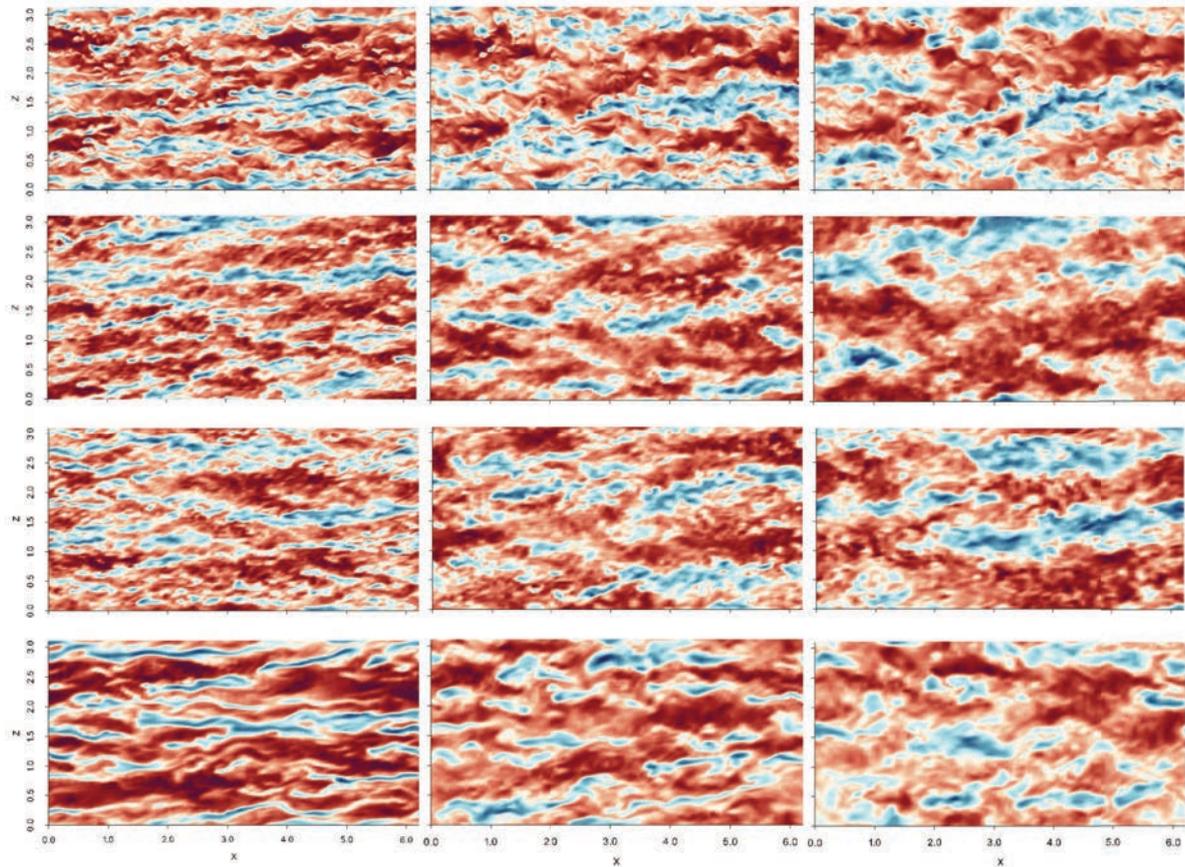


Figure 4.4: 2D instantaneous streamwise velocity contours at  $y^+ = 50, 100,$  and  $200$  as the first, second and third column respectively. From top to bottom, we have the high-resolution simulation, the MULU fully stochastic framework, the MULU deterministic framework, and the DSM framework. Blue indicates low-speed streaks and red indicates high-speed streaks.

of the unresolved small-scales. The deterministic MULU scheme is purely dissipative with a vanishing turbulence on the long run contrary to the stochastic scheme in which occurs a backscattering phenomenon through the energy intake brought by the small-scale random component. The Ito-Stokes drift, caused by inhomogeneity at the unresolved scales, is responsible for the streaking phenomenon and its effect is incorporated further by the backscattering introduced in the stochastic MULU. Backscattering, thus, plays only a secondary role in maintaining large scale-structures shaped by the small-scale inhomogeneity.

The comparison suggests that an effective modelling of the Ito-Stokes drift via the MULU framework is capable of explaining streak formation in wall-bounded flows through borrowed ideologies from oceanic modelling. The modelling framework at low-resolution captures this mechanism with a better physical representation of the flow than an established classical model. An application of this MULU framework to oceanic flows could, thus, redress the complexity of coupling the Craik-Leibovich equation with the oceanic flow evolution equation to simulate LC and similar flows.

This chapter analysed the capability of the fully stochastic version of the MULU as the driving mechanism for streak formation in a channel flow. Such a modelling has direct application to the field of oceanic modelling. In addition, the stochastic nature of the model also provides certain advantages for performing ensemble based DA. This chapter is to act

as a precursor for future studies on ensemble DA implementing the Mulu framework. However, such methods are not presented or analysed further in this thesis work. This thesis restricts itself to variational data assimilation methods and the implementation of a deterministic LES model within the VDA framework. With this in mind, the rest of this thesis manuscript focuses on the concepts, application and performance of VDA.



## Chapter 5

# Data assimilation techniques

The limitations of CFD and EFD techniques highlighted in §1.3 clearly endorse the need for a merger of the two techniques which can alleviate errors inherent to each methodology. Data Assimilation is one such compendium of techniques in which the observations or measurements of a system obtained from EFD are used to guide a numerical model of the system with CFD. First developed for atmospheric sciences and geosciences, mainly for numerical weather prediction (NWP), DA techniques are now becoming popular among fluid flow researchers as can be seen by the amount of publication in this field over the past decade (Gronskis et al., 2013; Artana et al., 2012; Misaka et al., 2008; Papadakis and Mémin, 2008).

The aim to DA is to achieve three main purposes (Robinson, 2015):

- Improve knowledge of the current state of the system based on past observations.
- Predict an accurate future state of the system from current and past observations.
- Provide an accurate estimate of the system state evolution i.e. the state trajectory.

A first DA attempt, aimed at NWP, was the optimal interpolation (OI) method based on the works of Bergthórsson and Döös (1955); Cressman (1959); Gandin (1965). OI is a fairly simple yet powerful methodology for data assimilation and remained the most used DA methodology for NWP throughout the 1970s and 80s. Since OI, numerous methodologies have been developed for DA categorised into two main, independent categories:

(i) The sequential approach aims at propagating forward the model state while correcting this state estimate when observation becomes available through the so called Kalman gain matrix which is expensive to estimate.

(ii) The variational approach arises from optimal control theory with the aim of minimising a cost function for discrepancies between observations and model trajectory. Such methods require the calculation of a gradient which is computationally expensive and requires special methods to obtain.

The mathematical formulation of a data-assimilation system is given in §5.1. §5.2 provides a review of existing data-assimilation methods and studies from a literature point of view with little to no mathematical derivations. §5.3 delves into 4D-Var DA approach for fluid flows formulated by Gronskis et al. (2013) in 2D and extended by Robinson (2015) to 3D. The incremental approach, conditional control, adjoint formulation, and validation is explored in this section.

## 5.1 Mathematical Formulation

Through DA techniques, the aim is to obtain the actual or true initial state space  $\mathcal{X}_0^{tr}(x)$  and control parameters  $\vartheta^{tr}(x)$  if any, given a dynamical evolution model  $\mathbb{M}$  of the system from an *a priori* known background state  $\mathcal{X}_0^b$  and, usually sparse, set of observations  $\mathcal{Y}$ . This can be expressed through the following set of equations:

$$\partial_t \mathcal{X}_t(x) + \mathbb{M}(\mathcal{X}_t(x)) = q_t, \quad (5.1.1)$$

$$\mathcal{X}_0(x) = \mathcal{X}_0^b(x) + \eta, \quad (5.1.2)$$

$$\mathcal{Y}_t = \mathbb{H}(\mathcal{X}_t(x)) + \epsilon_t, \quad (5.1.3)$$

where the sub-script denotes the temporal state of the system, and the super-script denotes the type of system state, i.e. background (*b*), analysis, or true (*tr*), to differentiate from time  $t$ ). All variables are functions in space  $x$  unless specified otherwise. The temporal evolution of the state in space ( $x$ ) and time ( $t$ ) through the dynamical model  $\mathbb{M}$  is denoted by eq. (5.1.1) up to a model error  $q_t$ . In a fluid dynamics context, the state of the system would denote the velocity and pressure fields while the dynamical model could be the non-linear Navier-Stokes equations or the filtered NS equation for an LES. The error  $q$  in the dynamical evolution equation can stand for discretisation errors, numerical errors, or model errors. However, within this thesis, as assumed by Robinson (2015), the model is assumed to be perfect with zero error.

The second equation equates the state of the system at time  $t = 0$ , i.e.  $\mathcal{X}_0(x)$  to the *a priori* known background state  $\mathcal{X}_0^b(x)$  up to an error  $\eta(x)$ . The background error is assumed to be of zero mean and associated to the background covariance matrix  $\mathbf{B}$  as,

$$\mathbf{B} = \mathbb{E}((\mathcal{X}_0^b - \mathcal{X})(\mathcal{X}_0^b - \mathcal{X})^T) = \mathbb{E}(\eta\eta^T). \quad (5.1.4)$$

The definition of this background covariance matrix can be of significance in a DA study, especially in the context of variational data assimilation. A method for defining this background covariance matrix through proper orthogonal decomposition (POD) techniques is explored in chapter 7.

The final equation relates the observations  $\mathcal{Y}$  with the state variable  $\mathcal{X}$  at each occurrence of the observations through the observation operator  $\mathbb{H}$  which can be non-linear. This model is assumed to be accurate up to an observation error  $\epsilon(t, x)$ . This error arises due to sparseness of the observations, or due to scale dissimilarity between the observations and the state space or due to a noisy set of observations. It covers equipment errors as well as errors arising from the observation operator. The error is assumed to be a zero mean Gaussian random field with the associated covariance tensor  $\mathbf{R}$ . For the fluid flow scenarios considered in this thesis, a pre-processing algorithm is run on the observational set relating it to the state space - this negates the need for an observation operator  $\mathbb{H}$  which is set to an identity matrix.

## 5.2 Data assimilation - a review

This section aims at providing a literature review dedicated to the development and performance of data assimilation methodologies rather than focusing on the mathematical aspects of DA. For a detailed mathematical description of the different DA models, please refer to Yang (2014); Reich and Cotter (2015). The mathematical principles required for the variational approach, which is studied in this thesis, are provided in §5.3.

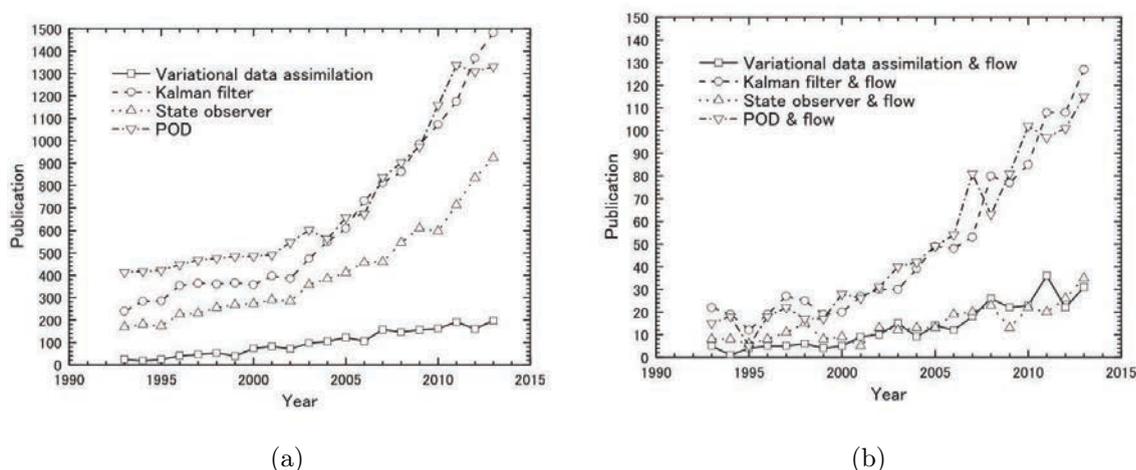


Figure 5.1: Number of publications in the field of DA from 1993-2013 in (a) all fields and (b) fluid flow problems (Hayase, 2015).

The growing importance of DA as a field is well captured by Hayase (2015) in his figures showing the variation of number of publications as a function of year in DA. In addition, he also graphically represented the number of flow related DA studies published over the years 1993-2013. The figures have been reproduced here in figure 5.1a and figure 5.1b. These clearly attest for a rising interest in DA among researchers especially within the domain of fluid mechanics. The importance of this field is further characterised by the emergence of dedicated books to data-assimilation such as Leeuwen et al. (2015); Reich and Cotter (2015); Law et al. (2015); Evensen (2009).

Sequential data-assimilation methods, including the Kalman filter, ensemble Kalman filter, particle filter, and Gauss-Markov estimate methods among others, are based on the concepts of least-square estimates, Bayesian formulation, or maximum likelihood techniques. The Kalman filter, based on stochastic principles developed by Kalman (1960), is one of the most common state-estimation methods within this DA paradigm. The method aims at asymptotically minimising the least-square error for Gaussian process of a discrete-time linear dynamical system. Many variations of this technique have since been developed to deal with unknown noise variance (adaptive Kalman filter), non-linear discrete systems (extended Kalman filter, unscented Kalman filter), for providing pdfs of stochastic variables (ensemble Kalman filter (Eknes and Evensen, 2002; Evensen, 2009)), etc. A similar approach is provided by the particle filter which lacks the Kalman gain making it computationally cheaper and easier to implement. A modification of the particle filter was suggested by Papadakis et al. (2010) called the ensemble Kalman filter to bridge particle filters and ensemble methods.

The Kalman filters are a common methodology in the fields of meteorology and oceanography during the mid 1900's (Sorenson, 1985). However, its application to fluid flow problems happened much later in early 1980's by Uchiyama and Hakomori (1983) who applied the filter to a circular pipe flow to measure instantaneous velocity profiles. Since then further applications have been successfully explored by Takehara et al. (2000); Yagoh et al. (1993). Another field to benefit from these filters is flow control where the filters are generally combined with some form of reduced-order model. Park and Lee (2001) dealt with a feedback control on velocity and temperature in a conducting fluid while Högberg et al. (2003) used an extended Kalman filter for controlling a boundary layer. Control of self-excited modes in jets and wakes were studied by Lauga and Bewley (2004), the

backward-facing step by Favennec et al. (2010), high-speed axisymmetric jets by Sinha (2011) and flat-plate boundary layer by Bagheri et al. (2009); Ahuja and Rowley (2010).

Hœpffner et al. (2005) and Chevalier et al. (2006) focused on the application of the Kalman filter to the DA problem of wall-bounded laminar and turbulent flows using linearised conservation equations. A non-linear estimation was done using the local ensemble transform Kalman filter for a Rayleigh-Benard flow by Cornick et al. (2009). A combination of Kalman filter with LES was suggested by Cahuzac et al. (2010) for obtaining mean flow-fields. The ensemble Kalman filter, common to the field of geophysics, was applied to fluid flows by Colburn et al. (2011) for estimate near-wall structures. Cuzol and Memin (2009) and Chin and Mariano (2010) are a few studies on the particle filter. The former applied the filter for temporal tracking of a fluid flow while the latter focuses on estimating trajectories of drifter particles in a turbulent flow.

Although first introduced by Sasaki (1958) in the 1950's, the variational methods are a much more recent addition to the field of DA with works in meteorology starting only around the late 1980's. Sugiura et al. (2008); Thépaut et al. (1993); Županski and Mesinger (1995) are a few samples of the extensive application of the variational methods in NWP. The considerable use of the variational methods was made feasible by the works of Le Dimet and Talagrand (1986) and Courtier and Talagrand (1987) who developed the adjoint formulation which facilitated a computationally achievable procedure for calculating the functional gradient. The aim of the variational methods is to alter the parameters of the model such that it minimises the gradient of the cost functional containing the error between the model trajectory and the observations. Thus it formulates the situation as an optimal control problem as done by Lions (1971). The parameters controlled could be the initial condition, inlet condition, boundary conditions, or model parameters (Artana et al., 2012; Gronskis et al., 2013; Papadakis and Mémin, 2008). The application of the variational methods to fluid-flow problem is an upcoming and on-going research interest with increased number of publications in the last decade. Misaka et al. (2008) used the variational approach to determine the initial condition for clear air turbulence. A combination of 4D-Var approach with reduced order modelling was by Tissot et al. (2013) with POD and Tissot et al. (2014) with DMD or dynamic mode decomposition. DMD is a post-processing methodology developed by Schmid (2010) which extracts dynamically relevant features from fluid flows - for more details of DMD, please refer to Schmid (2010); Leroux and Cordier (2016). The 4D-Var approach has been successfully applied to wake around a circular cylinder by Artana et al. (2012); Gronskis et al. (2013); Tissot et al. (2014) - this thesis is an expansion on these literary works focusing on the transitional wake behind a circular cylinder.

Other approaches of DA include the state observer proposed by Luenberger (1964) which uses a feedback control system to control the model based on the error between the model and the observations run in parallel. An application of such observer systems to non-linear systems was studied by Misawa and Hedrick (1989). The works of Badmus et al. (1996); Yau et al. (1995); Hayase and Hayashi (1997) are a few examples of the state observer system as applied to fluid flow studies. Another class of methods which fall loosely under the domain of DA, are experimental methods which enforce a physical constraint on the measurement data through a CFD scheme. Okuno et al. (2000); Suzuki et al. (2009b,a) are some studies implementing such a scheme with Okuno combining PIV with the NS equations and Suzuki combining DNS with PTV measurements. The VIC and VIC+ methodology developed at TU-Delft by Schneiders et al. (2014, 2016) are another set of methods which fall in this category.

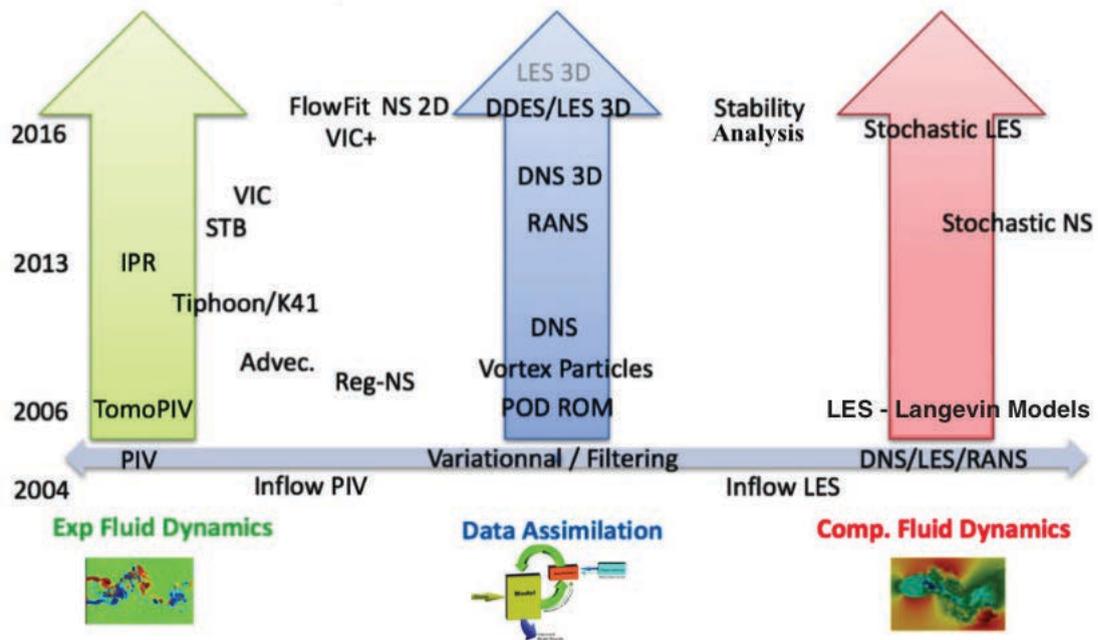


Figure 5.2: Different methodologies developed for fluid flow study over the past decade categorised into EFD, CFD and DA. Legend: IPR - iterative particle reconstruction (Wieneke, 2012), STB - shake-the-box (Schanz et al., 2016), VIC - vortex in cell (Schneiders et al., 2014), VIC+ - vortex in cell plus (Schneiders et al., 2016), Reg-NS - regularised Navier-Stokes (Héas et al., 2012), Advec - Advection based model (Scarano and Moore, 2012).

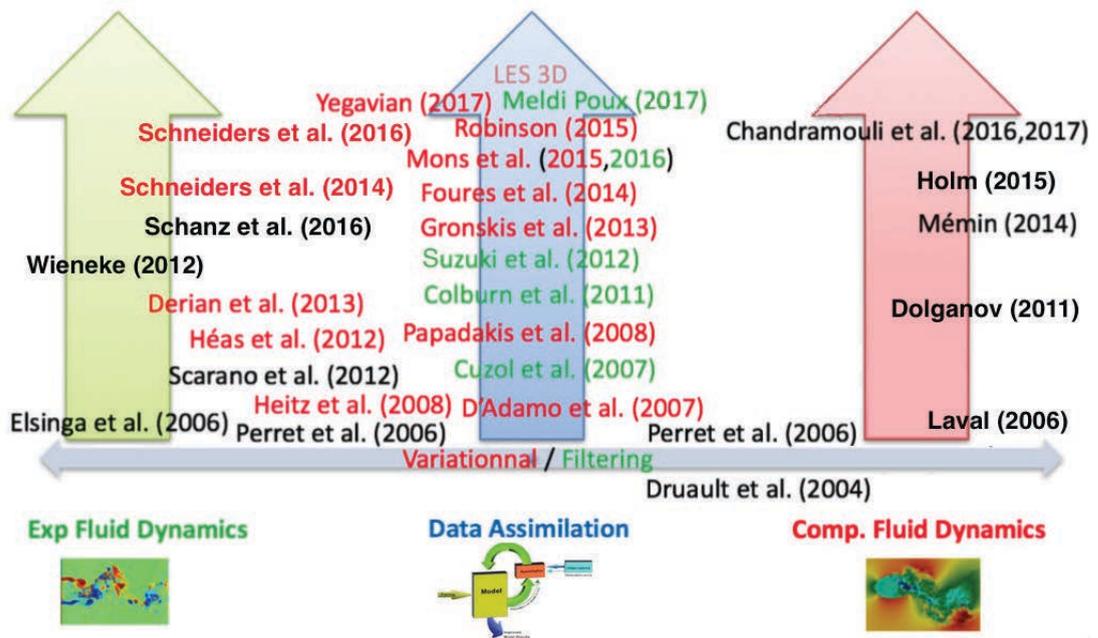


Figure 5.3: State-of-the-art publication in the field of fluid flow study over the past decade classified into variational approaches (in red) and filtering/sequential approaches (in green).

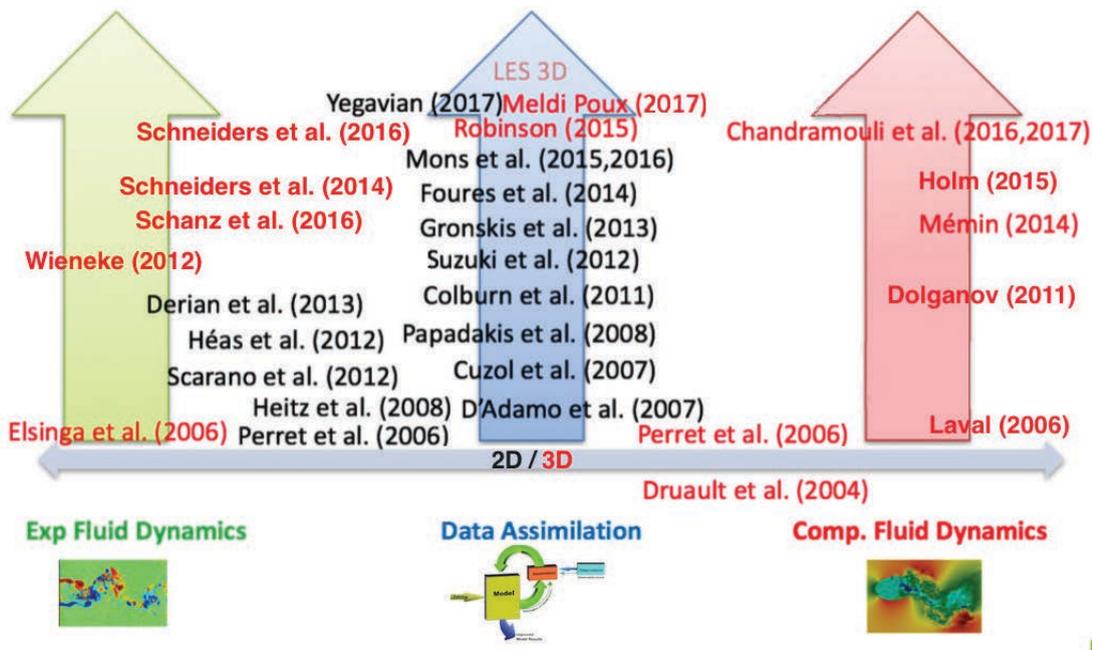


Figure 5.4: State-of-the-art publication in the field of fluid flow study over the past decade classified into 3D studies (in red) and 2D studies (in black).

Clearly, the application of DA to fluid flows has grown exponentially over the last decade. The number of new methodologies developed in the vein of DA can be seen in figure 1.6, which is recalled here in figure 5.2. The importance of DA for fluid flows can be further understood by plotting a non-exhaustive state-of-the-art methods developed in the last decade shown in figure 5.3. The methods have been sorted into variational (in red) and filtering/sequential (in green) methods. Clearly, a trend towards variational methods can be observed with the rapid rise of the publications in red. A further classification can be performed on the basis of dimensionality of the fluid flow and is shown in figure 5.4 where studies in 3D are shown in red and studies in 2D are in black. The restrictive cost of performing DA in 3D is clearly seen by the limited number of studies done in 3D in DA (limited to Robinson (2015); Meldi and Poux (2017)).

This thesis work aims at addressing this issue of restrictive computational cost by combining the newly derived LES model (Mémin, 2014), which has been shown to perform well for the flow under consideration (Chandramouli et al., 2017b, 2018), with the 4D-Var approach of Robinson (2015) at high Reynolds number in 3D. The 4D-Var principles are recalled in the following section.

### 5.3 4D-Var approach

The principle of calculus of variation states that the extrema of functional is obtained by setting the first variation of the function to zero. This is exploited by the variational data assimilation approach where a cost function  $J(\mathcal{X}_0)$  defined as,

$$J(\mathcal{X}_0) = \frac{1}{2}(\mathcal{X}_0 - \mathcal{X}_0^b)^T \mathbf{B}^{-1}(\mathcal{X}_0 - \mathcal{X}_0^b) + \frac{1}{2} \int_{t_0}^{t_f} (\mathbb{H}(\mathcal{X}_t) - \mathcal{Y}_t)^T \mathbf{R}^{-1}(\mathbb{H}(\mathcal{X}_t) - \mathcal{Y}_t), \quad (5.3.1)$$

where  $t_0$  and  $t_f$  are initial and final time of the assimilation algorithm and the evolution of the state of the system  $\mathcal{X}_t$  is given as,

$$\mathcal{X}_t = \mathcal{X}_0 + \int_{t_0}^t \mathbb{M}(\mathcal{X}(x))dx = \psi_t(\mathcal{X}_0), \quad (5.3.2)$$

where  $\psi_t(\mathcal{X}_0)$  is the flow map. The cost function is minimised by a gradient descent method aimed at achieving,

$$\frac{\partial J}{\partial \mathcal{X}} = 0. \quad (5.3.3)$$

The first term in the cost function stands for the error between the initial condition  $\mathcal{X}_0$  and the *a priori* known background condition  $\mathcal{X}_0^b$  weighted by the inverse of the background error covariance matrix  $\mathbf{B}^{-1}$ . The second term corresponds to the error between the state trajectory  $\mathcal{X}_t$  as obtained using the dynamical model ( $\mathbb{M}$ ) and the observations ( $\mathcal{Y}$ ) using an appropriate observation operator ( $\mathbb{H}$ ) and weighted with the inverse of the observation error covariance matrix  $\mathbf{R}^{-1}$ .

In order to optimise this cost function, the gradient needs to be calculated using an appropriate method. The simplest method is to use a finite difference scheme to approximate the gradient for variations of the dependent variables expressed, for the  $i^{th}$  component as ,

$$\frac{\partial J}{\partial \mathcal{X}_{0,i}} = \frac{J(\mathcal{X}_0 + \alpha \mathbf{e}_i) - J(\mathcal{X}_0)}{\alpha}, \quad (5.3.4)$$

where  $\mathbf{e}_i$  represents a vector of zeros except for the  $i^{th}$  element which is 1, and  $\alpha$  is a small scalar. For DA studies of fluid flows, the size of the state space could be of the order of  $10^7-10^8$ . To calculate by a brute force method, such as finite difference, the gradient at each iterative step of the minimisation algorithm is clearly not feasible. This fact, more than anything else, restricted the application of variational methods in DA studies until late 1980's when an elegant alternative solution was provided by Le Dimet and Talagrand (1986).

Their adjoint model calculates the gradient of the cost function with a single integration of the adjoint model backwards in time. Consider an adjoint variable  $\boldsymbol{\lambda}$  belonging to the same state space and integrate over the time range of the assimilation, the inner product of this variable with tangent linear model of the non-linear dynamical model,

$$\int_{t_0}^{t_f} \left\langle \frac{\partial d\mathcal{X}_t}{\partial t}, \boldsymbol{\lambda}_t \right\rangle dt + \int_{t_0}^{t_f} \left\langle \partial_{\mathcal{X}} \mathbb{M} d\mathcal{X}_t, \boldsymbol{\lambda}_t \right\rangle dt = 0 \quad (5.3.5)$$

where,

$$\partial_t d\mathcal{X}(t, x) + \partial_{\mathcal{X}} \mathbb{M}(\mathcal{X}(t, x)) d\mathcal{X}(t, x) = 0, \quad (5.3.6)$$

is the tangent linear model of the non-linear model with  $\partial_{\mathcal{X}} \mathbb{M}$  denoting the tangent linear operator associated to the non-linear model operator  $\mathbb{M}$ , and  $\langle \cdot, \cdot \rangle$  stands for the  $\mathbb{L}^2$  inner product. The corresponding gradient of the tangent linear model along a direction  $\delta \mathbf{n}$  is,

$$\left\langle \frac{\partial J}{\partial \mathbf{n}}, \delta \mathbf{n} \right\rangle = \left\langle \mathbf{B}^{-1}(\mathcal{X}_0 - \mathcal{X}_0^b), \delta \mathbf{n} \right\rangle - \int_{t_0}^{t_f} \left\langle (\partial_{\mathcal{X}} \mathbb{H})^* \mathbf{R}^{-1}(\mathcal{Y}_t - \mathbb{H}(\mathcal{X}_t)), \frac{\partial \mathcal{X}}{\partial \mathbf{n}} \delta \mathbf{n} \right\rangle dt, \quad (5.3.7)$$

where the adjoint operator has been introduced for the linearised observation operator as,

$$\langle (\partial_{\mathcal{X}} \mathbb{H}(\mathcal{X}))x, y \rangle = \langle x, (\partial_{\mathcal{X}} \mathbb{H}(\mathcal{X}))^* y \rangle. \quad (5.3.8)$$

Applying integration by parts to the first term in the inner product of the adjoint variable with the tangent linear model gives,

$$-\int_{t_0}^{t_f} \left\langle -\frac{\partial \boldsymbol{\lambda}_t}{\partial t} + (\partial_{\mathcal{X}} \mathbb{M})^* \boldsymbol{\lambda}_t, d\mathcal{X}_t \right\rangle dt = \langle \boldsymbol{\lambda}_{t_f}, d\mathcal{X}_{t_f} \rangle - \langle \boldsymbol{\lambda}_{t_0}, d\mathcal{X}_{t_0} \rangle. \quad (5.3.9)$$

Here, we have introduced the adjoint operator  $(\partial_{\mathcal{X}} \mathbb{M})^*$  for the tangent linear operator  $(\partial_{\mathcal{X}} \mathbb{M})$ , similar to that introduced for the observation operator. By imposing the adjoint variable to be a solution of the adjoint equation set,

$$\begin{cases} -\partial_t \boldsymbol{\lambda}_t + (\partial_{\mathcal{X}} \mathbb{M})^* \boldsymbol{\lambda}_t = (\partial_{\mathcal{X}} \mathbb{H})^* \mathbf{R}^{-1}(\mathcal{Y} - \mathbb{H}(\mathcal{X}_t)), \\ \boldsymbol{\lambda}_{t_f} = 0, \end{cases} \quad (5.3.10)$$

the gradient of the cost function can be evaluated by inserting this relation in to eq. (5.3.9) and equating  $d\mathcal{X}_{t_0} = \partial \mathbf{n}$  and  $d\mathcal{X} = (\partial \mathcal{X} / \partial \mathbf{n}) \delta \mathbf{n}$  as,

$$\frac{\partial J}{\partial \mathbf{n}} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}^{-1}(\delta \mathcal{X}_0^b - \delta \mathcal{X}_0). \quad (5.3.11)$$

Thus to obtain the gradient of the cost function, only a solution of the adjoint equation is necessary, performed backwards in time for a time-dependent problem set.

### 5.3.1 4D-Var Algorithm Variants

The **standard 4D-Var algorithm** aims at minimising,

$$J(\mathcal{X}_0) = \frac{1}{2} \|\mathcal{X}_0^b - \mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \int_{t_0}^{t_f} \|\mathcal{Y}_t - \mathbb{H}(\mathcal{X}_t)\|_{\mathbf{R}^{-1}}^2 dt, \quad (5.3.12)$$

with  $\|\cdot\|_A$  denoting the  $\mathbb{L}^2$  norm with the associated covariance function  $A$ . The aim of the assimilation approach is to find the optimal initial condition  $\mathcal{X}_0^a$  given a background initial condition  $\mathcal{X}_0^b$  subject to a dynamical model and observations. The optimal initial condition is then used to evaluate the analysis trajectory which is an ideal compromise between the background condition and the observations. The process is graphically represented in figure 5.5.

The covariance matrices  $\mathbf{B}$  and  $\mathbf{R}$  are used to normalise and weight the corresponding error functions at each point with respect to every other point depending on the confidence on the background and/or observation available at every point of the mesh. A solution of the minimisation process can be obtained using an iterative optimisation method. In the framework of this thesis, the limited storage gradient-based LBFGS optimisation method proposed by Nocedal (1980) is used.

The standard 4D-Var approach minimises a non-linear model whose non-linearity could result in an analysis trajectory initiated from a local optima rather than the global optima. This is a major limitation of the standard approach and was addressed by Courtier et al. (1994) by the introduction of the **incremental 4D-Var approach** where a linearised version of the dynamical model  $(\partial_{\mathcal{X}} \mathbb{M})$  is used to optimise an increment  $(\delta \mathcal{X}_0)$  to the initial condition around the current trajectory (see figure 5.6) giving the following cost function,

$$J(\delta \mathcal{X}_0) = \frac{1}{2} \|\delta \mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \int_{t_0}^{t_f} \|\mathbb{H}(\mathcal{X}_t) - \mathcal{Y}_t\|_{\mathbf{R}^{-1}}^2 dt, \quad (5.3.13)$$

with respect to the tangent linearised equations,

$$\partial_t \delta \mathcal{X}_t + (\partial_{\mathcal{X}} \mathbb{M}) \delta \mathcal{X}_t = 0, \quad (5.3.14)$$

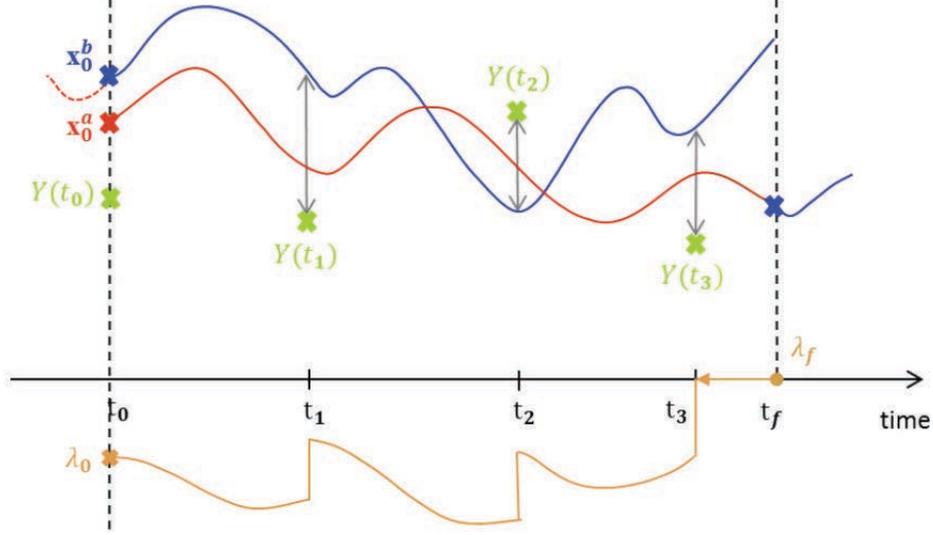


Figure 5.5: Standard 4D-Var approach to DA with the analysis trajectory (in red) obtained by an optimisation procedure from the background trajectory (in blue) subjected to discrete observations (in green) using the adjoint optimisation procedure (in orange).

$$\delta\mathcal{X}_0 = (\mathcal{X}_0^b - \mathcal{X}_0) + \eta(x), \quad (5.3.15)$$

where the flow trajectory ( $\mathcal{X}_t$ ) can be expressed as a function of the background and the incremental term,

$$\mathcal{X}_t = \psi_t(\mathcal{X}_0^b) + \partial_{\mathcal{X}}\psi_t(\mathcal{X}_0^b)\delta\mathcal{X}_0. \quad (5.3.16)$$

An alternate and simpler formulation can be obtained for the cost function as,

$$J(\mathcal{X}_0) = \frac{1}{2}\|\delta\mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\partial_{\mathcal{X}}\mathbb{H}(\delta\mathcal{X}_t) - \mathcal{D}_t\|_{\mathbf{R}^{-1}}^2 dt, \quad (5.3.17)$$

for the incremental flow trajectory,

$$\mathcal{X}_t = \partial_{\mathcal{X}}\psi_t(\mathcal{X}_0^b)\delta\mathcal{X}_0, \quad (5.3.18)$$

where  $\mathcal{D} = \mathcal{Y}_t - \mathbb{H}(\psi_t(\mathcal{X}_t^b))$  is the innovation vector.

This set of equations can be solved by the adjoint method by imposing the adjoint variable to be a solution of a modified adjoint equation set given as,

$$\begin{cases} -\partial_t \boldsymbol{\lambda}_t + (\partial_{\mathcal{X}}\mathbb{M})^* \boldsymbol{\lambda}_t = (\partial_{\mathcal{X}}\mathbb{H})^* \mathbf{R}^{-1} (\partial_{\mathcal{X}}\mathbb{H} \partial_{\mathcal{X}}\psi(\mathcal{X}_0) \mathcal{D}_t), \\ \boldsymbol{\lambda}_{t_f} = 0, \end{cases} \quad (5.3.19)$$

the gradient of the cost function can then be evaluated as, T

Such a methodology requires multiple looping structures with the inner loop optimising the increment while the outer loops optimises the non-linear trajectory. This can be computationally more expensive than the standard algorithm requiring a linearised version of the non-linear model but the likelihood of the incremental approach reaching the global optima is much higher than for the standard approach.

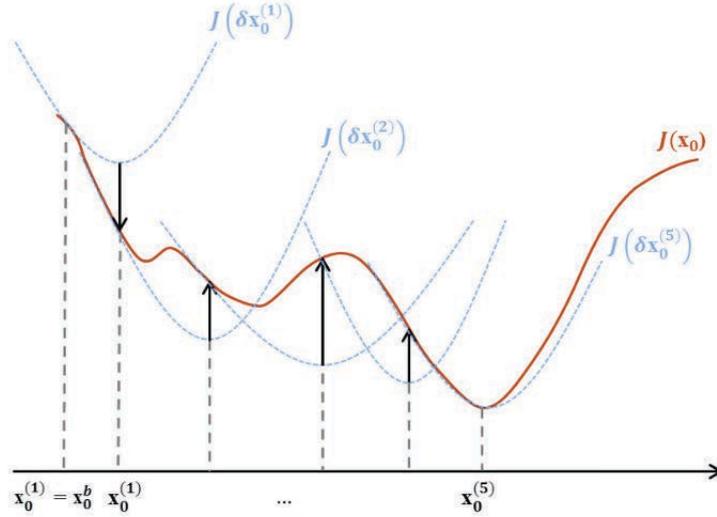


Figure 5.6: Incremental 4D-Var approach to DA with the non-linear cost function in red and the linearised convex cost function around a given  $\mathcal{X}$  shown in blue.

### Additional control

In the above enumerated versions of the 4D-Var approach, the optimisation procedure focuses only on the initial condition as a control parameter for obtaining the analysis trajectory. Additional control parameters can be easily envisaged which could improve the prediction of the analysis trajectory - a few examples of additional control parameters could be the inlet condition, model coefficient, etc. The inlet condition control was explored by Gronskis et al. (2013); Robinson (2015) and will be used as a control parameter in this thesis as well. The control on the LES model coefficient is an additional control that arises from the modelling approach and will be assessed in the chapter on DA applications. In this section, the mathematical formulation of the cost function for additional control variables  $\boldsymbol{\vartheta}_t$  is enumerated along with the modified adjoint equation set accounting for  $\boldsymbol{\vartheta}_t$ .

Consider an additional control parameter  $\boldsymbol{\vartheta}_t$  in the dynamical model giving,

$$\partial_t \mathcal{X}_t + \mathbb{M}(\mathcal{X}_t, \boldsymbol{\vartheta}_t) = q_t = 0, \quad (5.3.20)$$

$$\mathcal{X}_0 = \mathcal{X}_0^b(x) + \eta. \quad (5.3.21)$$

The cost function now has to minimise, alongside the background error  $\eta$  and the observation error  $\epsilon$ , the deviation of the control parameter from its *a priori* value ( $\boldsymbol{\vartheta}_t^c$ ) subject to a covariance  $\mathbf{B}_c$  as well,

$$J(\mathcal{X}_0) = \frac{1}{2} \|\mathcal{X}_0^b - \mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \int_{t_0}^{t_f} \|\mathcal{Y}_t - \mathbb{H}(\mathcal{X}_t)\|_{\mathbf{R}^{-1}}^2 dt + \frac{1}{2} \int_{t_0}^{t_f} \|\boldsymbol{\vartheta}_t - \boldsymbol{\vartheta}_t^c\|_{\mathbf{B}_c^{-1}}^2 dt. \quad (5.3.22)$$

The spatial definition of the control parameter strictly depends on the parameter itself. For an inlet condition control, the parameter is defined on the inlet plane while for the model coefficient it is defined over the entire spatial domain. The control parameter can also be temporally varying, thus requiring an individual optimisation at each time-step - this can provide, for example, a gradually changing inflow condition capable of better capturing the optimal analysis trajectory.

The perturbed linearised evolution equation, assuming a perturbation of the form  $d\mathcal{X} = \frac{\partial \mathcal{X}}{\partial \boldsymbol{\vartheta}} \delta \boldsymbol{\vartheta}_t + \frac{\partial \mathcal{X}}{\partial \mathbf{n}} \delta \mathbf{n}$ , is,

$$\partial_x \delta \mathcal{X}_t + \partial_x \mathbb{M}(\mathcal{X}_t, \boldsymbol{\vartheta}_t) \delta \mathcal{X}_t + \partial_u \mathbb{M}(\mathcal{X}_t, \boldsymbol{\vartheta}_t) \delta \boldsymbol{\vartheta}_t = 0, \quad (5.3.23)$$

Taking into account the adjoint observation operator  $(\partial_{\mathcal{X}} \mathbb{H}(\mathcal{X}))^*$ , the gradient of the cost function with the addition control parameter can be evaluated as,

$$\begin{aligned} \left\langle \frac{\partial J}{\partial \mathbf{n}}, \delta \mathbf{n} \right\rangle &= \left\langle \mathbf{B}^{-1}(\mathcal{X}_0 - \mathcal{X}_0^b), \delta \mathbf{n} \right\rangle \\ &\quad - \int_{t_0}^{t_f} \left\langle (\partial_{\mathcal{X}} \mathbb{H})^* \mathbf{R}^{-1}(\mathcal{Y}_t - \mathbb{H}(\mathcal{X}_t)), \frac{\partial \mathcal{X}}{\partial \mathbf{n}} \delta \mathbf{n} \right\rangle dt, \end{aligned} \quad (5.3.24)$$

$$\begin{aligned} \left\langle \frac{\partial J}{\partial \boldsymbol{\vartheta}}, \delta \boldsymbol{\vartheta} \right\rangle &= \int_{t_0}^{t_f} \left\langle \mathbf{B}_c^{-1}(\boldsymbol{\vartheta}_t - \boldsymbol{\vartheta}_t^c), \delta \boldsymbol{\vartheta} \right\rangle dt \\ &\quad - \int_{t_0}^{t_f} \left\langle (\partial_{\mathcal{X}} \mathbb{H})^* \mathbf{R}^{-1}(\mathcal{Y}_t - \mathbb{H}(\mathcal{X}_t)), \frac{\partial \mathcal{X}}{\partial \boldsymbol{\vartheta}} \delta \boldsymbol{\vartheta} \right\rangle dt, \end{aligned} \quad (5.3.25)$$

Applying the inner product of the adjoint variable ( $\boldsymbol{\lambda}$ ) to the evolution equation and performing an integration by parts of the first term gives,

$$\begin{aligned} - \int_{t_0}^{t_f} \left\langle - \frac{\partial \boldsymbol{\lambda}}{\partial t} + (\partial_{\mathcal{X}} \mathbb{M})^* \boldsymbol{\lambda}_t, d\mathcal{X}_t \right\rangle dt &= \langle \boldsymbol{\lambda}_{t_f}, d\mathcal{X}_{t_f} \rangle - \langle \boldsymbol{\lambda}_{t_0}, d\mathcal{X}_{t_0} \rangle \\ &\quad + \int_{t_0}^{t_f} \left\langle \delta \boldsymbol{\vartheta}_t, (\partial_{\boldsymbol{\vartheta}} \mathbb{M})^* \boldsymbol{\lambda}_t \right\rangle dt, \end{aligned} \quad (5.3.26)$$

where we now introduce the adjoint model operator with respect to the control variable  $(\partial_{\boldsymbol{\vartheta}} \mathbb{M})^*$ . By forcing this adjoint variable to be a solution to the adjoint equation set (eq. (5.3.10)), the gradient along  $(\delta \mathbf{n}, \delta \boldsymbol{\vartheta})$  can be evaluated as,

$$\frac{\partial J}{\partial \mathbf{n}} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}^{-1}(\mathcal{X}_0^b - \mathcal{X}_0), \quad (5.3.27)$$

$$\frac{\partial J}{\partial \boldsymbol{\vartheta}} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}_c^{-1}(\boldsymbol{\vartheta} - \boldsymbol{\vartheta}_c) + (\partial_{\boldsymbol{\vartheta}} \mathbb{M})^* \boldsymbol{\lambda}, \quad (5.3.28)$$

and the optimal can be identified using the iterative LBFGS or other similar algorithms.

Thus, the adjoint methodology combined with the 4D-Var approach allows easy addition of control parameters provided the adjoint operator with respect to the control parameter  $((\partial_{\boldsymbol{\vartheta}} \mathbb{M})^*)$  can be constructed.

The incremental form of the cost function inclusive of the control variables can be formulated as,

$$J(\delta \mathcal{X}_0, \delta \boldsymbol{\vartheta}) = \frac{1}{2} \|\delta \mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2} \int_{t_0}^{t_f} \|\delta \boldsymbol{\vartheta}_t\|_{\mathbf{B}_c^{-1}}^2 dt + \frac{1}{2} \int_{t_0}^{t_f} \|\mathbb{H}(\mathcal{X}_t) - \mathcal{Y}_t\|_{\mathbf{R}^{-1}}^2 dt, \quad (5.3.29)$$

with the corresponding gradient, using the adjoint approach, defined as,

$$\frac{\partial J}{\partial (\delta \mathcal{X}_0)} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}^{-1} \delta \mathcal{X}_0, \quad (5.3.30)$$

$$\frac{\partial J}{\partial (\delta \boldsymbol{\vartheta})} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}_c^{-1} \delta \boldsymbol{\vartheta} + (\partial_{\boldsymbol{\vartheta}} \mathbb{M})^* \boldsymbol{\lambda}. \quad (5.3.31)$$

The incremental approach including the additional control variable, as applied by Robinson (2015), is termed, henceforth, as the 4D-Var approach and is used in this thesis for performing Data Assimilation. The algorithm for the 4D-Var approach is given in algorithm 1. The 4D-Var technique is modified to give the 4D-Var<sub>les</sub> approach with the inclusion of the spatial variance based model (StSp) defined in §2.2.4. The StSp model has been validated for the flow under consideration in §3.3.3. The numerical formulation of the tangent and adjoint code for the model is elaborated in chapter 7 along with its application and results.

---

**Algorithm 1** Incremental approach to variation assimilation with additional control variables

---

Initialisation:  $\mathcal{X}_0^{(1)} = \mathcal{X}_0^b$

**repeat**

Set initial increment:  $\delta\mathcal{X}_0^{(k)} = 0$ .

**repeat**

Forward evolution with the tangent linearised model

Calculate cost functional  $J(\delta\mathcal{X}_0, \delta\boldsymbol{\vartheta})$

Backward evolution with the adjoint model - calculate  $\boldsymbol{\lambda}_{t_0}$

Apply LBFGS algorithm

Update increment  $\delta\mathcal{X}_0^{(k)} = \mathbf{B}\boldsymbol{\lambda}_{t_0}$

**until** ( $\|\delta\mathcal{X}_0^{(k),curr} - \delta\mathcal{X}_0^{(k),prev}\|$ )  $< \delta tol$

Update initial condition:  $\mathcal{X}_0^{(k+1)} = \mathcal{X}_0^{(k)} + \delta\mathcal{X}_0^{(k)}$

Update control variable:  $\boldsymbol{\vartheta}^{(k+1)} = \boldsymbol{\vartheta}^{(k)} + \mathbf{B}_c(\partial_{\boldsymbol{\vartheta}}\mathbb{M})^*\boldsymbol{\lambda}_{t_0}$

**until** ( $\|\mathcal{X}_0^{(k+1)} - \mathcal{X}_0^{(k)}\|$ )  $< tol$

Forward integration of the non-linear dynamical model with  $\mathcal{X}_0^a = \mathcal{X}_0^{(k+1)}$  and  $\boldsymbol{\vartheta}^a = \boldsymbol{\vartheta}^{(k+1)}$  to get analysis trajectory  $\mathcal{X}_t^a$

---

The next chapter focuses on a new methodology for very fast and efficient reconstruction of 3D flow-fields from 2D cross-plane observation data. Such a method is very useful in DA for prescribing the background condition as well as providing observations with higher confidence in the data-set.

## Chapter 6

# Flow reconstruction - snapshot optimisation method

### 6.1 Introduction

With the advent of particle image velocimetry (PIV) in the 1980's (Adrian, 1984), the field of experimental fluid dynamics grew exponentially with the ability now to measure data over large 2D domains. Further improvements saw the measurement of velocity fields in 3D albeit at limited spatial resolutions using tomographic-PIV (tomo-PIV) techniques (Scarano, 2012) or more recently using 3D+time Particle Tracking Velocimetry (4D-PTV) techniques (Schanz et al., 2016). However, these methods are still quite restricted in terms of their spatial and temporal resolutions as well as the ability to observe the full state-space vectors. To improve 3D spatiotemporal PTV measurements, Navier-Stokes constraints have been proposed in an optimisation framework. Schneiders et al. (2016) use the vortex in cell plus (VIC+) technique to reconstruct instantaneous fields from time-resolved measurements using the velocities and the material derivatives. In a similar approach, Gesemann et al. (2016) proposed FlowFit, a similar approach to interpolate Lagrangian measured velocities to a regular grid. The aim of all these methods has been to reconstruct 3D full-scale observations from sparse partial data-sets.

An upcoming methodology to complete this partial state-space vector obtained from an experiment is to perform Data Assimilation (DA), wherein a combination of computational methods (CFD) and experimental data (EFD) is used to complete the missing data. However, DA is limited in performance by the numerical methods employed and more importantly, the quality of the experimental data used to guide the numerical simulation.

PIV data-sets provide a good spatial and temporal resolution of data in 2D. Many studies exist in literature that use this data to perform 2D DA to estimate an improved field of interest (see the works of Gronskis et al. (2013), Fujisawa et al. (2005), and D'adamo et al. (2007)). Until recently, the extent of DA studies have been limited to 2D due to inhibitive costs of performing DA - both the variational and the sequential approach to DA are computationally expensive. The variational approach, formulated as an optimal control problem, seeks to minimise the error between the simulation and the observation data by estimating an optimal trajectory bounded by physical laws. Such a minimisation procedure relies on the adjoint formulation of the associated dynamics for calculating the gradient based on the works of Lions (1971), and Le Dimet and Talagrand (1986). To formulate, implement and use this adjoint formulation can be tedious and numerically difficult. In the sequential approach, based on stochastic filtering techniques, multiple realisations of the state vector are propagated in time taking into account available measurements at each

time. Such a formulations suffers from the ‘curse of dimensionality’ (Yang et al., 2015) as the few samples considered are unlikely to represent the large state-space. In addition, the ease of adaptability of many sequential approaches to high-dimensional, non-linear problems is questionable (Mons et al., 2016).

In the field of variational DA in fluid dynamics, due to the use of turbulence modelling such as Large Eddy Simulations (LES), and Reynolds-averaged Navier-Stokes (RANS) models, an expansion into three dimensional assimilation studies have recently been achieved (Chandramouli et al., 2017a). A DNS based 3D DA study was performed by Robinson (2015) at a very low Reynolds number (Re) of 300 for flow around circular cylinder. As we expand into higher Re flows, the need for 3D data-sets describing the flow becomes imperative. Existing 3D DA studies either rely on non-physical mirror-imaging technique, or interpolation/extrapolation techniques to obtain 3D observation data from existing 2D PIV data or involve complex time-consuming methodologies such as the ‘Empty Box’ technique of Robinson (2015). Tomo-PIV and PTV techniques are capable of producing 3D flow fields but are still limited in their spatial temporal extent to sparse observations. Thus, a quick and effective method to reconstruct the velocity field over a volume could be very useful for 3D DA studies until advances in experimental techniques become capable of capturing spatially well-resolved observations in a large 3D domain.

Simple methods for 3D complex flow reconstruction from 2D PIV data have been proposed in literature. The scanning PIV technique proposed by Brücker (1995) for the analysis of 3D effects in a cylinder wake flow involves classical PIV technique with the scanning of a volume via a light sheet. Since then this approach has been carried out for single (Zhang et al., 2008; Brücker et al., 2012) or stereo (Casey et al., 2013) camera arrangements. The technique assumes a frozen velocity field during the scanning process that thus need to be sufficiently quick. With the assumption of frozen turbulence via Taylor hypothesis, Ganapathisubramani et al. (2008) used cinematographic stereoscopic PIV to transform time into space and reconstruct a quasi-instantaneous volume of data in a turbulent jet. The PIV plane was perpendicular to the streamwise flow direction. Steinberg et al. (2009) developed an orthogonal-plane cinematographic stereoscopic PIV to study turbulence-flame interactions. The plane perpendicular to the streamwise flow was used for the 3D reconstruction of the flow and the plane parallel to the streamwise flow direction to observe the interactions between the turbulent structures and the flame. In order to obtain simultaneous spatiotemporal velocity information for 3D flow structure investigation, Kähler and Kompenhans (2000) studied multiple plane stereo PIV. Braud et al. (2004) carried out such an approach in a dual plane stereo PIV configuration combined with 3D POD to reconstruct and analyse the flow structures of the circular cylinder near wake-mixing layer interaction. Kit et al. (2005) proposed another approach coupling hot-wire measurements, to obtain detailed time-series, and a limited number of phase-locked stereo PIV measurements to reconstruct time variations of large coherent structures in a forced mixing layer. More recently in the same spirit, Hamdi et al. (2018) synchronised with a trigger signal parallel planes stereo PIV and POD to reconstruct a volume of an impinging jet.

This work proposes a quick and effective algorithm for reconstructing 3D velocity fields from 2D observations obtained on two perpendicular planes, easily available from established experimental methods such as stereo PIV. The snapshot optimisation (SO) methodology is formulated in the next section. This is followed by a section analysing the performance of the algorithm using varied sets of observational data - both synthetic and experimental. A final section of concluding remarks follows.

## 6.2 Snapshot optimisation method

The concept of the SO method and the algorithm inputs are explained in section 6.2.1. The mathematical formulation of the optimisation problem is provided in section 6.2.2. A reduced order version of the model using proper orthogonal decomposition (POD) is discussed in section 6.2.3. An averaging variant of the model is proposed in section 6.2.4.

### 6.2.1 SO model concept

Consider a flow exhibiting a direction of homogeneity, for example, wake flow around a circular cylinder. Figure 6.1 depicts the geometry for this flow with the two planes of measurements. The plane in red, parallel to the cylinder axis, is referred to as the inlet plane (IP) while the plane in green, perpendicular to the cylinder, is referred to as the observation plane (OP). While the OP has been depicted in the middle of the domain, this is not a necessary condition of the reconstruction - any plane perpendicular to the spanwise  $z$ -direction can be utilised as the OP. The objective is to reconstruct the 3D domain encompassing the IP and the OP, from only one 3C vector field in the IP and an ensemble of 3C vectors fields in the OP.

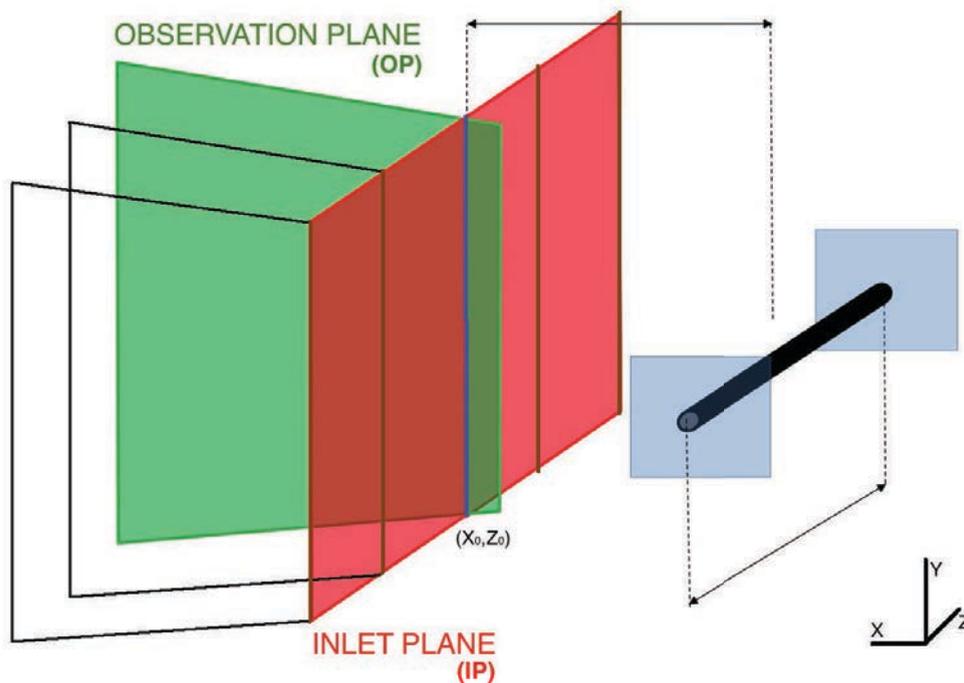
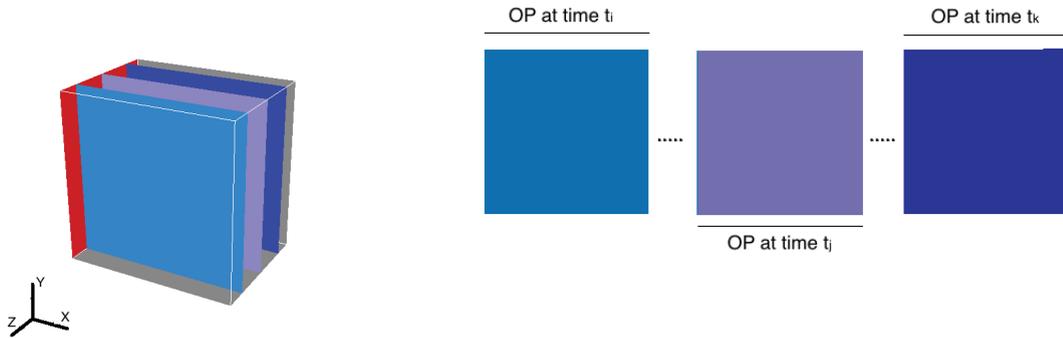


Figure 6.1: Geometry for wake flow around a circular cylinder with planes of measurement. The observation plane (OP) is shown in green and the inlet plane (IP) in red. The line of intersection between the two planes at  $(x_0, z_0)$  is shown in blue. Spanwise planes that need to be reconstructed are outlined in black with shared edges with the inlet plane outlined in brown.

This 3D domain can be considered to be made up of a set of 2D parallel planes (a sample decomposition containing three such planes is shown in figure 6.2a). Each parallel plane is assumed to begin at the IP thus sharing a common edge (along  $y$ ) with the inlet

plane. Now, we make the critical assumption: considering the turbulence homogeneity of the flow along the spanwise  $z$ -direction, a sufficiently long time-sequence of snapshots of the observation plane (in green in figure 6.1) contains at some point within that time-sequence a representation of each of the 2D parallel planes of figure 6.2a. This is graphically shown in figure 6.2b. In other words, it is assumed that one event constituted by one planar velocity field on a given plane, is observed at a given time of the observations sequence. Such a reconstruction is along the lines of the “pouring time in to space” techniques of Schneiders et al. (2016).



(a) 3D domain split into multiple parallel 2D plane observations.

(b) 2D planes represented within the time-sequence of the observation plane measurements.

Figure 6.2: 3D domain of interest shown within the original volume and within the observation measurements.

Thus, to reconstruct the 3D domain of interest from the 2D measurements, we need to identify the snapshots within the OP time-sequence which correspond to each of the parallel planes comprising the volume. Since all of these planes share an edge with the inlet plane where measurements are available, an optimisation algorithm can be formulated minimising the error between the OP and the IP measurements along the intersecting edge for each plane. The inputs needed for such a reconstruction is the inlet plane velocity field at a given time instance at which the volumetric measurement is required and sufficiently long time-sequence of snapshots of the OP. These, upon reconstruction, would then provide us with the much needed volumetric measurements from easily obtainable planar measurements. The minimisation algorithm for identifying the optimal snapshot time instance is formulated in the next section.

## 6.2.2 Mathematical formulation

Let the velocity field in 3D be defined as  $\mathbf{U}(x, y, z, t)$ , where  $\mathbf{U}$  is a vector incorporating the three components of velocity at each point in space defined by  $(x, y, z)$  and  $t$  defines the instant in time of the flow field. The velocity field in the OP can be expressed in this format as  $\mathbf{U}(x, y, z_0, t)$ , where  $z_0$  stands for the point of intersection between the OP and the IP. Similarly, the velocity field in the IP can be expressed as  $\mathbf{U}(x_0, y, z, t_0)$ , where  $x_0$  stands for the first point in the OP and  $t_0$  stands for the time instance at which the volumetric measurement is required. For ease of understanding, the velocity on the OP common with the IP (i.e. at  $x_0, y, z_0$ ) is expressed as  $U_{OP}(y, t)$  and the velocity on the IP at initial time is referred to as  $U_{IP}(y, z)$ . In order to find the optimal instance of match between the OP

observations in the plane  $z = z_0$  and the IP observation in the plane  $x = x_0$ , a simple least square based cost function can be formed for each vertical line (along  $y$ ) in the IP as a function of the spanwise direction  $z$ :

$$J(z) = \sum_y (\mathbf{U}_{OP}(y, t_z) - \mathbf{U}_{IP}(y, z))^2, \quad (6.2.1)$$

where  $J(z)$  is the cost function for a given  $z$ , and  $t_z$ , which is the parameter to optimise, is the time of observation, as a function of  $z$ , that matches best with the IP measurements for a given  $z$ . The optimisation function here accumulates the error of all three velocity components. In practice, individual optimisation procedures may be carried out for each velocity component and the results can be collated to provide the volumetric field.

The gradient of the cost function  $J(z)$  along the variable of interest  $t_z$  is then given as:

$$\partial_{t_z} J(z) = \sum_y 2[\mathbf{U}_{OP}(y, t_z) - \mathbf{U}_{IP}(y, z)] \partial_{t_z} \mathbf{U}_{OP}(y, t_z), \quad (6.2.2)$$

where the velocity derivative is calculated using a first order forward finite difference scheme.

A simple time stepping gradient descent method can then be envisaged as:

$$t_z^{n+1} = t_z^n - \epsilon \partial_{t_z} J(z), \quad (6.2.3)$$

where  $\epsilon$  is a suitable weight.

In order to perform a gradient descent (GD) optimisation, an appropriate starting guess  $t_{st}$  for  $t_z$  needs to be provided. While a random  $t_{st}$  can be provided, better methods can be envisaged. Two methods were implemented to identify a good  $t_{st}$ . The first method identifies a fixed  $t$  such that the summation of the total square error between OP at time  $t$  and IP for all  $z$  was minimal:

$$t_{st} = \min_z \left( \sum_y (\mathbf{U}_{OP}(y, t) - \mathbf{U}_{IP}(y, z))^2 \right) \quad (6.2.4)$$

Such a method could help in reducing computational cost by providing a starting guess closer to the optimal solution - for an observation dataset ranging over a long period of time, such a method could reduce computational cost considerably.

The above defined method provides a good estimate as a starting guess however, for a large spatial domain, the cost of computing this would be of the order of  $n_y \times n_z \times n_t$ , where  $n_i$  is the number of points along the  $i$  spatial or temporal direction. The second method, which is computationally less expensive, involves an averaging along the spanwise  $z$  direction of the IP ( $\bar{\mathbf{U}}_{IP}(y)$ ) and minimising the error of the OP measurements with this averaged lateral profile of the IP:

$$t_{st} = \min_y (\mathbf{U}_{OP}(y, t) - \bar{\mathbf{U}}_{IP}(y))^2. \quad (6.2.5)$$

The cost for such a computation is only of the order of  $(n_y \times n_z + n_y \times n_t)$ .

It is important to note that the use of a GD methodology could result in a fractional  $t(z)$  for which observations are not available. For such cases, an interpolation of the velocity profile is performed to obtain the field at the fractional  $t_z$  - linear interpolation has been used in this manuscript when required.

### 6.2.3 Reduced order formulation

A full scale 3D reconstruction of a flow field can be computationally expensive while inducing strong discontinuities and introducing divergence within the flow field. A Proper Orthogonal Decomposition (POD) based reconstruction would be computationally inexpensive while providing a smoothed flow field. Here, a POD method of reconstruction is constructed similar to the full scale reconstruction explained in the previous section. Two different POD can be envisaged on the given flow field. One is along the 1D vertical ( $y$ ) direction for identifying the best match - this would reduce the computational cost of optimisation. The second is a 2D POD of the OP velocity facilitating a large scale reconstruction resulting in a smoothed flow field. The two POD formulations are described further in the following sections.

#### 1D-3C POD - matching

Let the flow field along  $x_0$  in the OP be decomposed as:

$$\mathbf{U}'_{OP}(y, t) = \sum_{i=1}^{N_{1D}} \phi_i^{1D}(y) a_i^{1D}(t) \quad (6.2.6)$$

where  $\mathbf{U}'_{OP}(y, t) = \mathbf{U}_{OP}(y, t) - \bar{\mathbf{U}}_{OP}(y)$ , with  $\bar{\mathbf{U}}_{OP}(y)$  being the time averaged quantity,  $N_{1D}$  is the total number of modes,  $\phi_i^{1D}$  is the POD spatial modes, and  $a_i^{1D}$  are the time coefficients for a given  $t$ .

Considering the turbulence homogeneity assumption along the spanwise direction, the IP flow field can also be expressed on the same spatial modes obtained from the OP decomposition.

$$\mathbf{U}'_{IP}(y, z) = \sum_{i=1}^{N_{1D}} \phi_i^{1D}(y) a_i^{1D}(z), \quad (6.2.7)$$

where  $a_i^{1D}(z)$  represents the IP coefficient for the flow field along the lateral direction for each spanwise  $z$  coordinate and thus is expressed as a function of  $z$  rather than time  $t$ . With the two sets of time coefficients, based on the same set of spatial modes, a GD minimisation algorithm can be formulated, similar to the full scale reconstruction, comparing the time coefficients instead of the flow field:

$$J(z) = \sum_{i=1}^{n_{1D}} (a_i^{1D}(t_z) - a_i^{1D}(z))^2, \quad (6.2.8)$$

where  $J(z)$  is the cost function, and  $n_{1D}$  stands for the subset of modes considered from the  $N_{1D}$  modes global set. A gradient of this cost function is then given as:

$$\partial_{t_z} J(z) = \sum_{i=1}^{n_{1D}} 2[a_i^{1D}(t_z) - a_i^{1D}(z)] \partial_{t_z} a_i^{1D}(t_z). \quad (6.2.9)$$

Similar methods can be used for the starting guess  $t_z$ , and for calculating the gradient as enumerated in the full scale reconstruction. Such a decomposition will provide a decrease in computation cost of the minimisation algorithm, however, it would still lead to a full scale reconstruction of the flow field. This reconstruction could have strong discontinuities as well as non-negligible divergence. To attenuate this, a large-scale reconstruction of the flow can be done by performing a POD on the 2D OP measurements and using only the high-energy modes.

### 2D-3C Planar POD - reconstruction

With the optimal time  $t_z^{\text{opt}}$  obtained from the 1D POD matching algorithm or from the full scale algorithm, the reconstruction of the 3D domain can be done using 2D POD. Let the 2D OP flow field be decomposed as:

$$\mathbf{U}'_{OP}(x, y, t) = \sum_{i=1}^{N_{2D}} \phi_i^{2D}(x, y) a_i^{2D}(t), \quad (6.2.10)$$

where  $\mathbf{U}'_{OP}(x, y, t) = \mathbf{U}_{OP}(x, y, t) - \bar{\mathbf{U}}_{OP}(x, y)$ ,  $\bar{\mathbf{U}}_{OP}(x, y)$  is the average along time, and  $N_{2D}$  is the total number of modes in 2D.

The reconstruction of the 3D velocity ( $\mathbf{U}_{rec}$ ) can then be obtained used the time coefficients of the snapshot at the optimal time in the sequence.

$$\mathbf{U}_{rec}(x, y, z, t_0) = \bar{\mathbf{U}}_{OP}(x, y) + \sum_{i=1}^{n_{2D}} \phi_i^{2D}(x, y) a_i^{2D}(t_z^{\text{opt}}) \quad \forall z \quad (6.2.11)$$

Depending on the number of modes considered ( $n_{2D}$ ), a smooth field or a full scale representation can be obtained for  $\mathbf{U}$ .

#### 6.2.4 Averaging SO method

For a given line on the IP, there may be multiple optimal time-instances of OP snapshots whose corresponding error varies only marginally from one another. By performing the optimisation as described above, we identify only one such optimal time-instance which is used to recreate the 3D velocity field. Such a reconstruction will realistically provide a velocity field that has minimal error but with significant noise/discontinuities due to the violation of continuity between subsequent spanwise velocity fields. This could be possibly redressed by identifying and averaging multiple time-instances depending on a user-defined error threshold as a percentage of the minimal error obtained from the optimisation procedure. For such a methodology (termed as ‘Averaging SO method’), the reconstructed velocity field can be expressed as,

$$\mathbf{U}_{rec}(x, y, z, t_0) = \frac{\sum_{i=1}^{n_z} \mathbf{U}_{OP}(x, y, t_z^i)}{n_z} \quad \forall z, \quad (6.2.12)$$

where  $n_z$  is the number of optimal snapshots identified by the algorithm for each  $z$  with the user-defined tolerance, and  $t_z^i$  is the time of each optimal snapshot from 1 -  $n_z$  for each  $z$ .

## 6.3 Application and results

The full-scale reconstruction algorithm is applied first in §6.3.1 to two synthetic cases at moderate Re: (i) Wake flow around a cylinder at Re 3900 (section 6.3.1), and (ii) channel flow at friction velocity based  $\text{Re}_\tau$  of 590. This is followed by §6.3.3 where the method is applied to three experimental data sets: (i) Wake flow around a cylinder at Re 300, (ii) Wake flow around a cylinder at Re 3900, and (iii) mixing layer with  $\lambda$  of 0.33. This section concludes with an analysis of variations of the SO algorithm and post-processing techniques on the reconstructed volumetric data. Details of the five cases are provided in table 6.1.

Table 6.1: Summary of synthetic and real configurations.

| Case                 | $Re$ | Snapshots | $\Delta t U/L^a$ | $(L_y \times L_z)/L$ | $(L_x \times L_y)/L$ | $n_y \times n_z$ | $n_x \times n_y$ | $t_{\text{rec}} [s]$ |
|----------------------|------|-----------|------------------|----------------------|----------------------|------------------|------------------|----------------------|
| 1 – LES wake         | 3900 | 2 500     | 0.003            | $4 \times \pi$       | $5 \times 4$         | $130 \times 48$  | $60 \times 130$  | 14.96                |
| 2 – DNS channel      | 590  | 1 000     | 0.00125          | $2 \times \pi$       | $2\pi \times 2$      | $257 \times 384$ | $384 \times 257$ | 257.12               |
| 3 – PIV wake 300     | 300  | 4 000     | 0.12             | $4.3 \times 6.4$     | $8.4 \times 4.3$     | $72 \times 108$  | $145 \times 72$  | 4.85                 |
| 4 – PIV wake 3900    | 3900 | 20 540    | 0.258            | $7 \times 5.6$       | $4.3 \times 7$       | $93 \times 72$   | $64 \times 93$   | 9.16                 |
| 5 – PIV mixing layer | 1540 | 4 000     | 0.015            | $2.4 \times 2.7$     | $2.4 \times 2.4$     | $100 \times 115$ | $100 \times 100$ | 6.85                 |

<sup>a</sup> $L$  stands for the cylinder diameter  $D$ , for the channel height  $H$  or for the mixing layer vorticity thickness  $\delta_\omega$ .

### 6.3.1 Synthetic case-studies

#### Wake flow around a circular cylinder at $Re$ 3900

A first application of the optimisation algorithm is performed with the synthetic data for wake flow around a circular cylinder at  $Re$  3900. This flow is of specific importance as DA is carried out in further chapters on this flow for this Reynolds number. Thus, a good reconstruction from the algorithm will provide importance information, namely background initial condition and/or volumetric observations, for carrying out 4D-Var. The input data are obtained from an LES simulation of the flow using the flow solver Incompact3d with the local spatial variance model in the MULU framework (Laizet and Li, 2011, Chandramouli et al., 2018). A domain measuring  $20D \times 20D \times 3.14D$ , where  $D$  stands for the diameter of the cylinder, is simulated with  $241 \times 241 \times 48$  mesh points with a time-step of 0.003 non-dimensional time. The mesh along the  $y$  direction is stretched with more points near the centre of the domain. A total of 2500 3D velocity snapshots are stored every 50 time-steps corresponding to  $\sim 75$  vortex sheddings after an initialisation period of 15 vortex sheddings with no data-collection.

The 3D snapshots are pre-processed to extract the 2D information on the IP and the OP shown in figure 6.1. The IP, measuring  $L_y \times L_z = 4D \times 3.14D$ , is located  $1.66D$  behind the trailing edge of the cylinder. The OP, measuring  $L_x \times L_y = 5D \times 4D$ , intersects the IP at the mid-point of the spanwise direction ( $z = 1.57D$ ). The IP is obtained corresponding to the first instance in the OP time-sequence. A sample streamwise velocity field on the IP and OP is shown in figure 6.3.

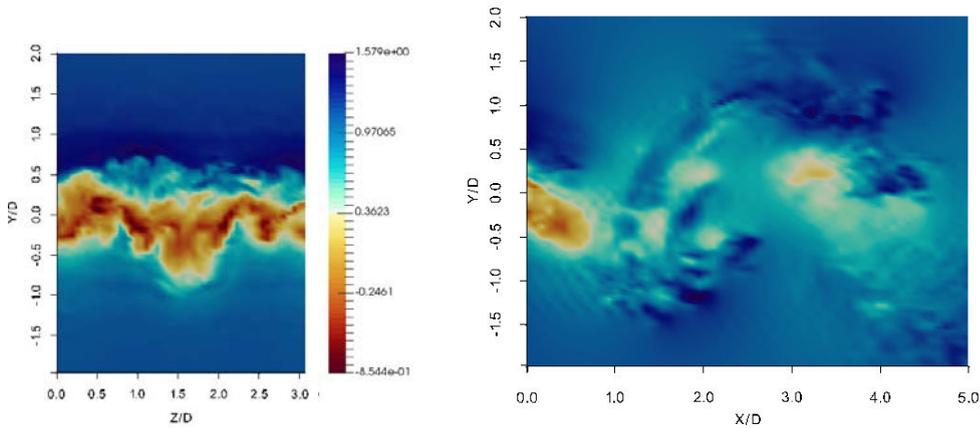


Figure 6.3: Instantaneous streamwise velocity fields extracted on the IP and the OP from the LES of wake flow.

The IP reconstruction from the algorithm is shown in figure 6.4. The biggest advantage of such a reconstruction algorithm is the time taken for the reconstruction ( $t_{\text{rec}}$ ) which for this case was  $t_{\text{rec}} = 14.96$  s. This time is independent of the number of OP snapshots considered or the type of flow. It is only dependent on the discretisation of the intersection line between the OP and the IP. A finer discretisation of this line would increase the computational cost of the algorithm and vice-versa. The ability of the algorithm to capture the main structures of the flow is seen in figure 6.4. Just from a single plane observation time-sequence, the IP has been reconstructed capturing all large scale structures of the flow in all three velocity directions. The small scale structures are also captured, however, this is attenuated by a noise and discontinuities in the reconstruction. Such a reconstruction provides a much better estimate of the 3D flow domain as compared to simplistic techniques such as mirror-imaging or interpolation.

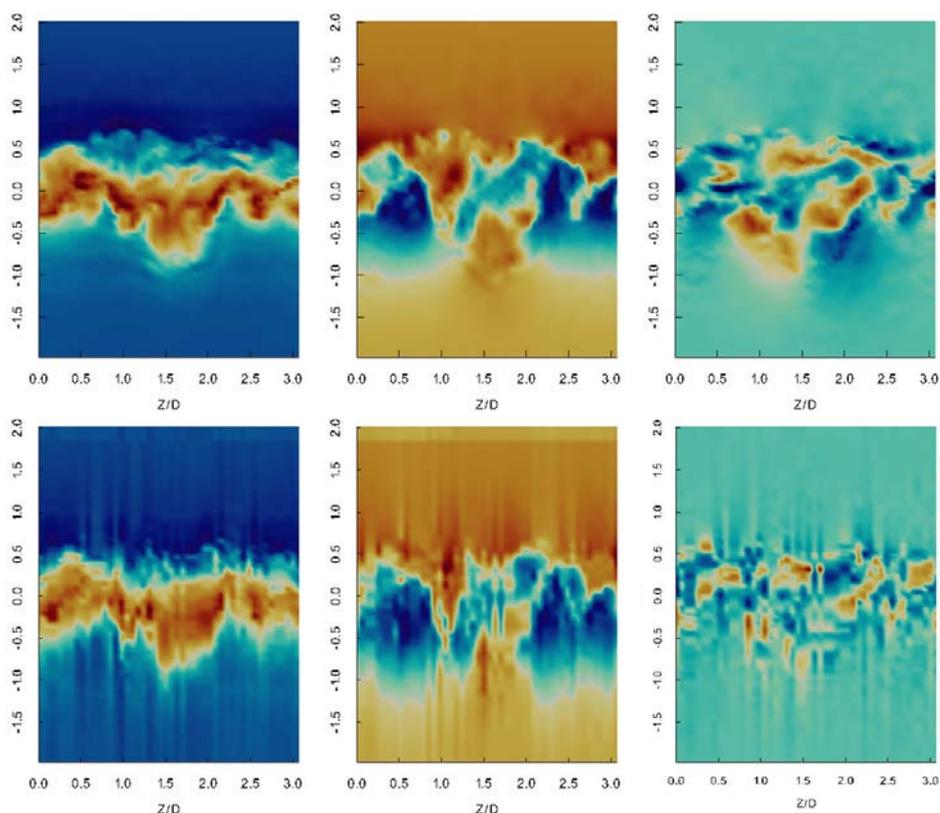


Figure 6.4: LES wake flow instantaneous velocity fields extracted on the IP. From top to bottom: reference and reconstruction. From the left to the right: streamwise, vertical and spanwise velocity components.

The availability of 3D reference velocity field allows us to study the 3D reconstruction efficiency of the algorithm by comparing the velocity iso-contours (see figure 6.5). The reconstructed 3D velocity accurately captures the large-scale motions of the flow. The location of the vortex street is captured well in the reconstruction iso-contours. There is significant introduction of noise especially in the spanwise velocity component but the large scale-correlation of vortex structures is efficiently represented. Considering the computational efficiency and simplicity of the algorithm, the results obtained advocate towards the efficacy and applicability of the algorithm as a preliminary reconstruction tool to obtain 3D motion fields from cross-plane observations. The case study corresponding to the averaging

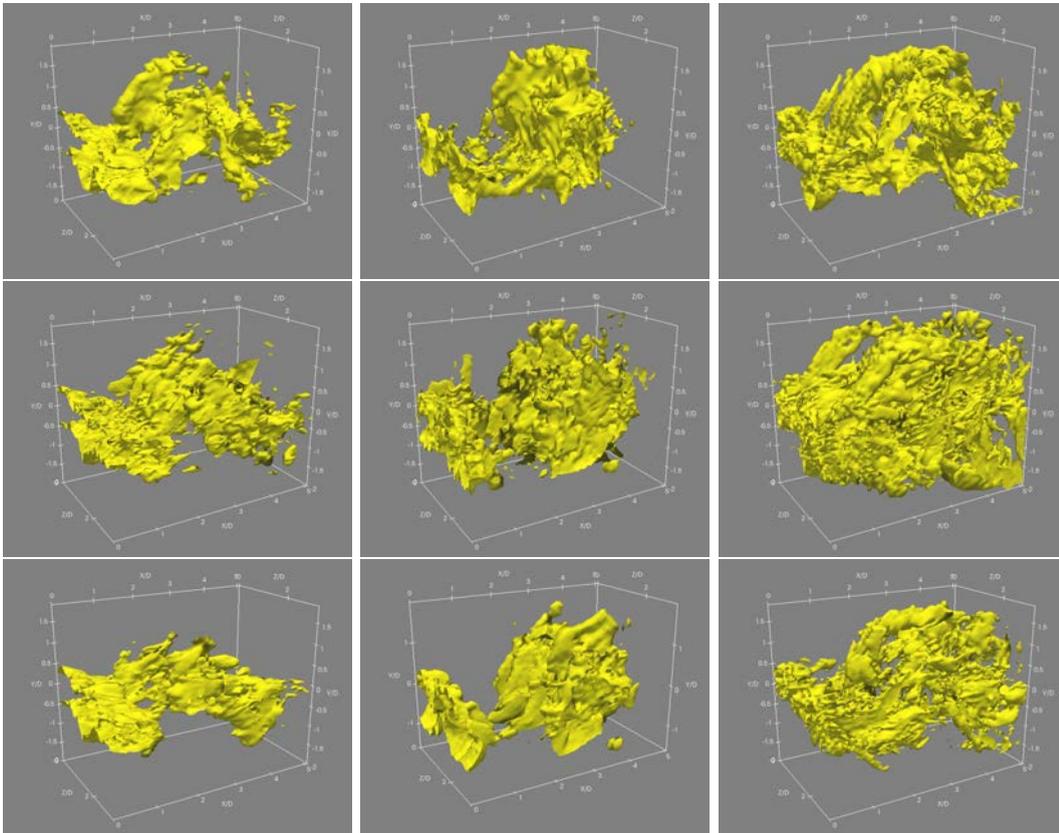


Figure 6.5: LES wake flow instantaneous velocity iso-contours extracted on the 3D domain. From top to bottom: reference, SO reconstruction and averaging SO reconstruction. From the left to the right: streamwise, vertical and spanwise velocity components.

SO reconstruction method is explained in §6.3.2. The iso-contours are presented here for ease of comparison and to avoid repetition.

### Channel flow at $Re_\tau$ 590

The previous case study analysed algorithm capability with respect to a large-scale data-set obtained from an LES. In this case, the algorithm capabilities are analysed for a highly resolved numerical data set - the ability of the algorithm to handle high resolution, small scale information is analysed. The data are obtained from a direct numerical simulation (DNS) of channel flow at  $Re_\tau$  590 using Incompact3d. The DNS is performed on a domain measuring  $12.56H \times 2H \times \pi H$  discretised into  $768 \times 257 \times 384$  mesh-points with a time-step of 0.00125 non-dimensional time. Considering the computational cost and memory requirements of the DNS, only a limited set of 1 000 snapshots could be obtained spaced 50 time-steps apart. The collection of the snapshots begins after the flow is converged. This case-study is a good example to analyse the performance of the algorithm under limited observation snapshots. However, considering the availability of 3D data-set, multiple OP can be defined and velocity fields from parallel OPs can be extracted. This alleviates the issue of insufficient snapshots and provides an opportunity to study the performance of the algorithm with respect to the size of the OP snapshots. Three sub-cases are studied with OP data extracted from 1(1p), 5(5p), and 10(10p) parallel planes - 10 planes consist of only 2.5% of the total spanwise set of planes.

The IP is placed in the centre of the channel measuring  $L_y \times L_z = 2H \times \pi H$  discretised with the same number of points as the DNS, i.e.  $257 \times 384$ . The OP begins at the IP and extends to the end of the channel with the intersection at  $z = [1.57H]$  for 1p,  $z = [0.52H, 1.04H, 1.57H, 2.09H, 2.62H]$  for 5p, and for 10p  $z = [0.26H, 0.52H, 0.78H, 1.04H, 1.3H, 1.57H, 1.83H, 2.09H, 2.35H, 2.62H]$ . The OP is discretised into  $N_x \times N_y = 384 \times 257$  mesh-points. The geometry of the flow and the position of the measurement planes are shown in figure 6.6. The reference inlet streamwise velocity and a sample OP measurement for the streamwise velocity is shown in figure 6.7.

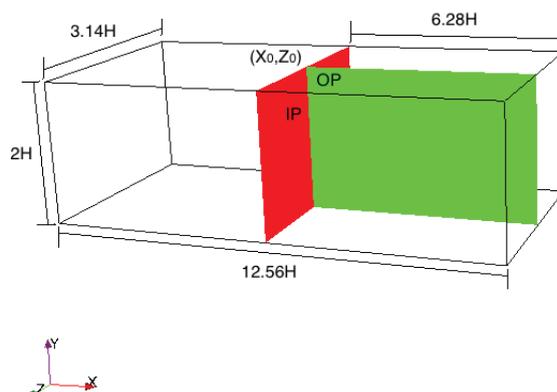


Figure 6.6: Geometry for channel flow at  $Re_\tau$  590 with the OP in green and IP in red placed in the middle of the channel at  $x = 6.28H$ , where  $H$  is the half-height of the channel.

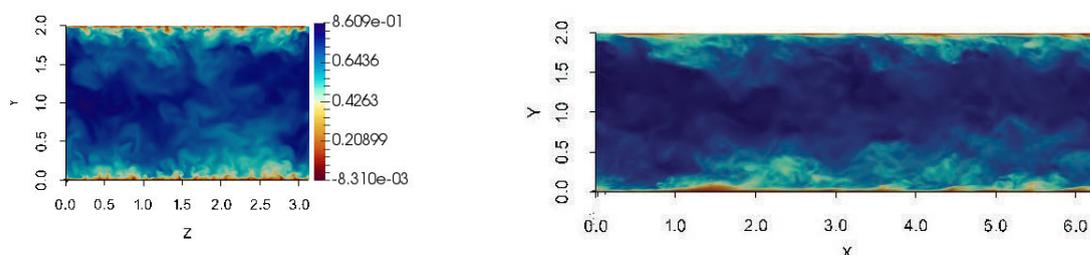


Figure 6.7: Instantaneous streamwise velocity fields extracted on the IP and the OP from the DNS of channel flow.

The reconstructed velocity fields on the IP are shown in figure 6.8. The reconstruction time for 1p case, as shown in table 6.1, is  $t_{rec} = 257.12s$  - the higher time is due to the finer resolution of the DNS. From the reconstructed images, it is clear that small scale structures of the DNS are not entirely reconstructed by the algorithm for the vertical ( $v$ ) and spanwise ( $w$ ) velocity profiles. However, for all velocity components, near the wall where important turbulent phenomena occur, the algorithm reconstructs with good accuracy the flow features. A significant improvement in the reconstructed velocity fields can be seen with 5p reconstruction as compared to the 1p. A further increase to 10p does not provide a drastic improvement but better representation of the small scale structures can be observed, mainly near the wall. This suggests that a minimum set of OP snapshots are required in order to obtain a good reconstruction and any further increase tends to improve the output further albeit at a reduced rate.

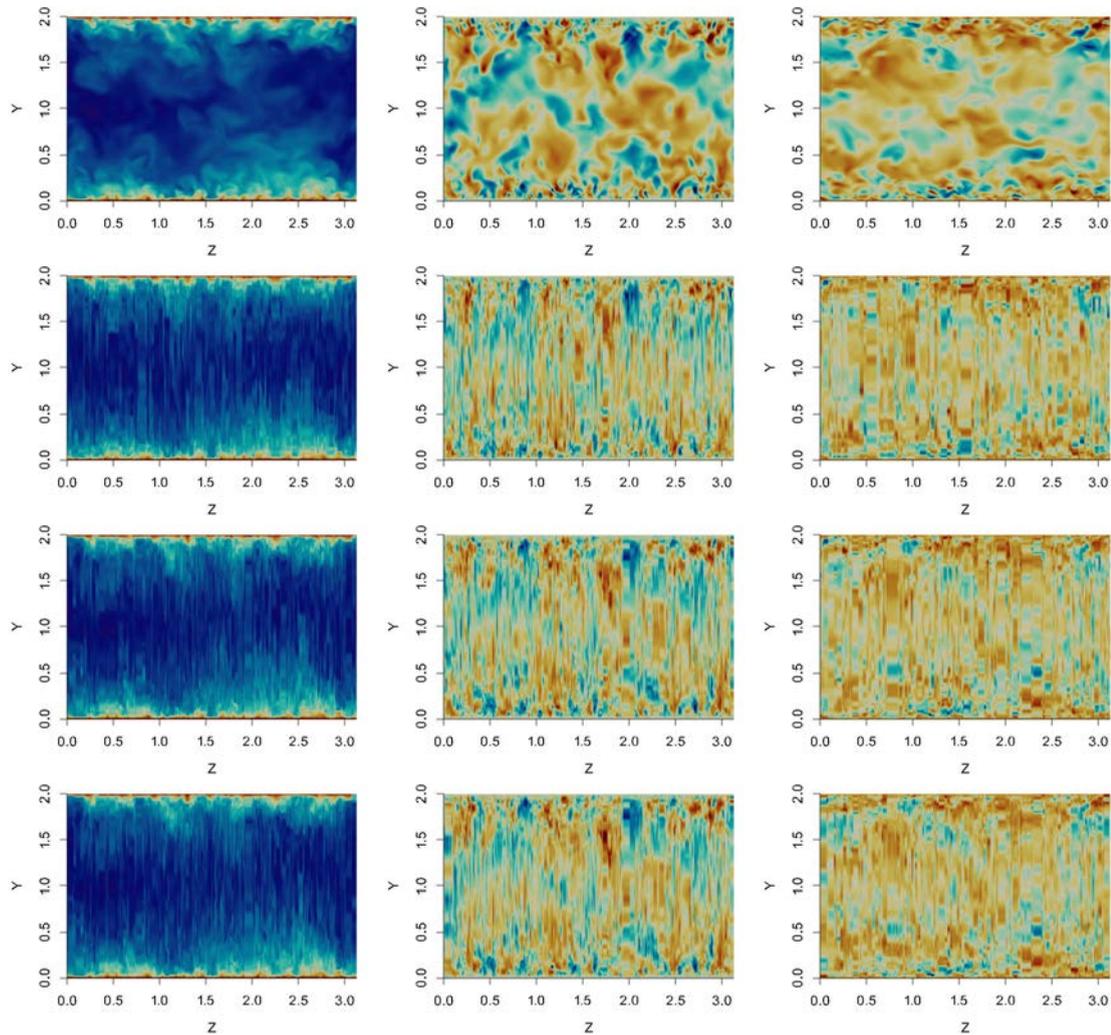


Figure 6.8: DNS channel flow instantaneous velocity components on the IP. From top to bottom: reference, reconstruction with 1 plane, reconstruction with 5 planes, and reconstruction with 10 planes. From the left to the right: streamwise, vertical and spanwise velocity components.

### 6.3.2 Algorithm Enhancement Techniques

Before diving into algorithm performance on experimental data-sets, the algorithm variants are studied here with respect to the ideal case of synthetic data obtained for wake flow around the cylinder at  $Re$  3900. The POD and averaging SO algorithms are explored towards their capabilities to smoothen the discontinuities in the flow and a simulation-based technique is analysed for removing divergence from the reconstructed volume.

#### POD based SO method

The mathematical formulation for the POD based SO method, provided in section 6.2.3, is applied here to wake flow around a circular cylinder at  $Re$  3900 (case-study 1 in table 6.1). Both a 1D POD based matching algorithm and a 2D POD based reconstruction is performed and the reconstructed volume is presented in figure 6.9. The first 72 modes are used for the POD matching while only the first 20 modes are used in the reconstruction phase - the reconstruction process is faster taking only  $t_{rec} = 1.6s$  for the reconstruction. The POD based reconstruction provides a smooth representation of the velocity fields thus reducing discontinuities. However, this is associated to a reduced recovery of the small-scale structures in the POD reconstruction. Such a reconstructed field could be used as a reference, so called background initial condition, in variational data-assimilation procedures as it provides a smooth 3D velocity field based on the observations from which the optimisation procedure for the DA can begin.

#### Averaging SO method

The results obtained from the POD reconstruction (in figure 6.9) validates our assumption that there maybe multiple optimal snapshots for a given  $z$  plane. The time instances obtained from the full-scale reconstruction and the POD based reconstructed vary slightly for certain  $z$  planes. The availability of multiple optimal time-instances is the basis behind the averaging SO method which is applied here. The utility of this method lies in reduction of spurious oscillations and removal of spanwise discontinuities due to averaging procedures similar to the POD. The error threshold for each plane is set as optimal error plus a 10% bias. The results of the averaging SO method is shown in figure 6.9. The averaging method performs similar to the POD based SO method, however, a higher recovery of the small scale structures are obtained with the averaging method compared to POD as well as a significant improvement in spanwise continuity. The reconstruction of the averaging SO method provides a balance between the POD based SO method and the simple SO algorithm - the results by averaging are smoother with moderate retainment of small-scale structures.

Both the averaging SO method and the POD based SO method help reduce discontinuities within the reconstruction. The POD method is an ideal alternative for fast computations ( $\sim 1.6$  seconds) as the optimisation is done on a smaller set of modes as compared to the full-scale reconstruction. The averaging SO method, however, appears to give better results compared to the POD at the cost of higher computational time ( $\sim 8.1$  seconds) as well as needing, ideally, more number of OP snapshots to find multiple optimal snapshots. While for this specific case the additional time taken may not be high, for larger dimensional systems searching over a larger time-sequence of data, the difference between the two method may be high.

A 3D comparison is performed with the velocity iso-contours of the averaging SO method, and the simple SO method from §6.3.1 with the reference field in figure 6.5.

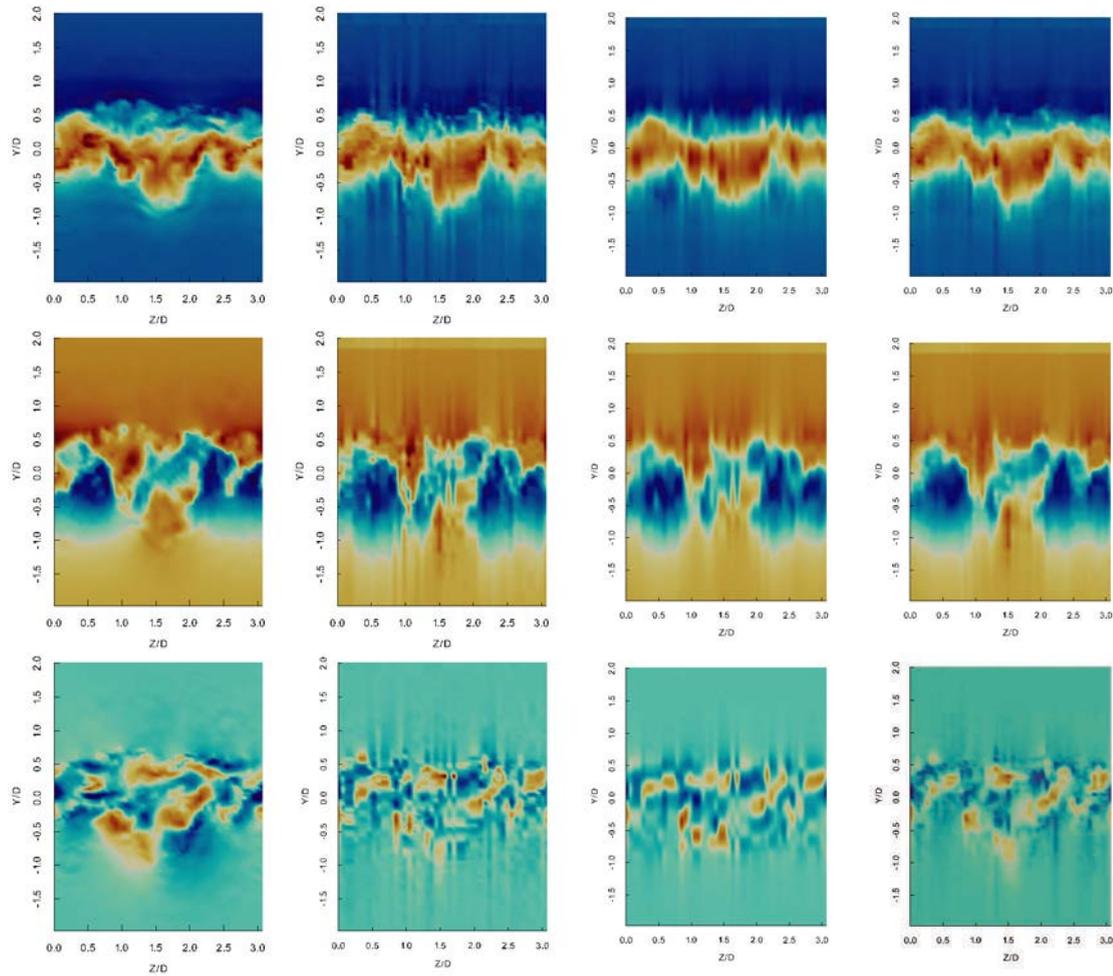


Figure 6.9: Reconstructed velocity fields using POD based SO method and averaging SO method. From top to bottom: streamwise, vertical, and spanwise velocity component. From left to right: reference, SO reconstruction, POD based SO reconstruction, averaging SO reconstruction.

The reduction of noise can be seen with the use of the averaging procedure as compared to the simple SO algorithm. In addition, a better correlation between structures is also seen with the averaging method. However, coherent small-scale structures are lost with the application of the averaging method that are seen in the reference - this is clearly seen in the streamwise velocity iso-contours where the rising vortex structure of the reference is seen in the simple reconstruction but not the averaged.

### Divergence-free reconstruction

The flows considered in this work are incompressible fluid flows and thus are divergence free. However, the process of reconstruction provides a final volumetric data-set which is clearly not divergence free. In this sub-section a few possible methods for producing divergence free data-sets by post-processing are discussed.

An incompressible fluid flow solver, such as `incompact3d`, maintains a divergence free flow by the action of the pressure-poisson equation where the divergence present in the flow is used to modify the pressure which successively modifies the velocity field to be divergence free using a fractional step method (Laizet and Lamballais, 2009). Thus, performing a simulation, using such a flow solver with the reconstructed velocity field as the initial condition, for a single time-step would provide a divergence free velocity field with an associated pressure containing the divergence.

The reconstruction for the synthetic wake flow at  $Re$  3900 has been simulated with `incompact3d` to produce such a divergence-free field - the final divergence of the flow was of the order of  $10^{-14}$  i.e. numerical accuracy. The divergence field before and after simulation with `Incompact3d` is shown in figure 6.10. The stability of the reconstruction was also analysed by simulating further the flow for a limited number of time-steps and the flow was indeed stable when initialised with the reconstructed velocity fields. This is necessary in order to use the results of the SO method in variational DA as the background initial condition. An added advantage of combining SO reconstructed volume with `Incompact3d` for DA is the automatic enforcement of divergence free condition as it is evolved forwards in time by the flow solver. Thus explicit methods to remove divergence from the flow are not required.

An alternative to produce a divergence free reconstruction is to use the Helmholtz decomposition (Van Bladel, 1958). Helmholtz theorem states that "A continuous vector field can be decomposed into the sum of a gradient and a curl term", i.e. a velocity field can be split as:

$$\bar{U} = -\nabla\phi + \nabla \times \psi \quad (6.3.1)$$

where  $\bar{U}$  is the velocity field,  $\phi$  is the scalar potential and  $\psi$  is the vector potential. The solenoidal  $U_{sol}$  and irrotational  $U_{irr}$  part of velocity can hence be expressed as:

$$U_{sol} = \nabla \times \psi; \quad U_{irr} = -\nabla\phi \quad (6.3.2)$$

Using simple algebraic manipulations, the scalar potential can be calculated following which the irrotational field, which contains the divergence, can be calculated and removed from the true field. Such a manipulation can be easily done in the Fourier space. The resultant vector field would contain minimal divergence as compared to the original field. However, as a part of the original field was removed, the energy of the resultant field would be lower than the original and this difference needs to be compensated for.

The main application of this algorithm is towards experimental stereoscopic PIV data-sets. These data-sets tend to be a large-scale representation of the flow due to the use of interrogation windows i.e. similar to the LES case-study presented and the algorithm could

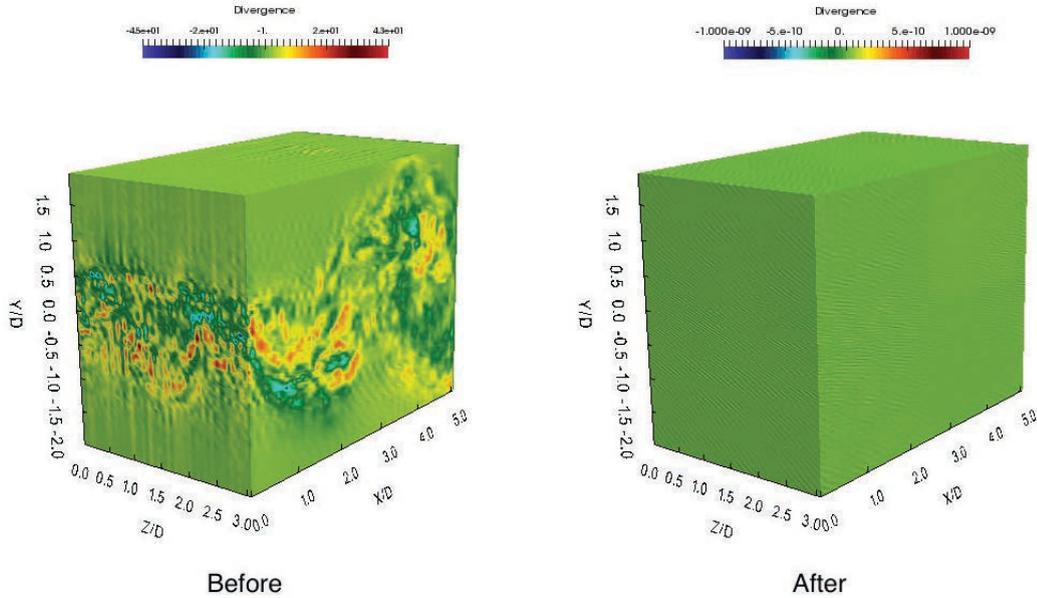


Figure 6.10: Divergence of the flow field before and after simulation with Incompact3d.

be expected to reconstruct accurately the large-scale volumetric data from such data-sets. The performance of the algorithm with such experimental data-sets is the focus of the next section.

### 6.3.3 Experimental case-studies

The experiments were conducted in two wind tunnels located at the Irstea regional centre in Rennes, France. One open wind tunnel was used to simulate both low Re wake flow and mixing layer flow (case 3 and 5 respectively in table 6.1) in a test section measuring  $3 \text{ m} \times 1 \text{ m} \times 1 \text{ m}$ . The other closed-loop wind tunnel was used to simulate a wake flow (case 4 in table 6.1) in a test section measuring  $2.4 \text{ m} \times 0.28 \text{ m} \times 0.28 \text{ m}$ .

Two perpendicular plane 2D3C PIV measurements are taken in the test sections as shown in figure 6.1. The origin of the coordinate system is defined at the intersection between the cross planes, with  $x$ ,  $y$  and  $z$  directed towards the streamwise, transverse and spanwise direction, respectively. A stereo PIV system was carried out to measure the three velocity components in both planes with a Litron laser system and two high speed cameras Phantom M310 (CMOS size of  $800 \times 1280 \text{ px}$ ). The flows were seeded by smoke generated from oil. The resulting image pairs were analysed by Davis software from LaVision. The specifics of the setup for each case-study is explained further in the corresponding sections.

#### Wake flow around a circular cylinder at Re 300

This case study analyses the performance of the algorithm for a low speed cylinder wake using experimental data-set. Data-sets obtained from PIV provide a significant challenge to reconstruction algorithms as the algorithms need to contend with measurement errors and noisy fields. In addition, there is a limitation on the temporal resolution for data measurement along with a limited ability to measure only large-scale flow-fields in 2D which is spatially restricted. Considering all the added difficulties, a first application is done for a laminar, low speed case study for wake flow around a cylinder.

A cylinder of diameter  $D = 10$  mm and length 300 mm was placed in the air flow. Two thin end-plates were used with specifications provided by Stansby (1974) to avoid the boundary layer effects of the walls of the test-sections. The end plates were separated by a distance of 300 mm with the clearance between the plates and the walls of the wind tunnel kept at 350 mm. The free stream velocity was set at  $U = 0.48$  m/s corresponding to a  $Re$  of 300. On both the IP and the OP, 4 000 image pairs were acquired at a frequency of 500 Hz. The IP was placed  $19D$  (190 mm) downstream of the cylinder. The OP intersects the IP through the middle with a 10 mm protrusion behind the IP.

The IP is discretised into  $N_y \times N_z = 349 \times 268$  mesh-points while the OP is discretised into  $N_x \times N_y = 280 \times 347$  mesh-points. Due to the presence of end-plates the data measured on the borders of the planes can be erroneous. In addition, due to the interaction of the laser sheets, the final measurement of the OP is performed not on a rectangular domain but on a trapezoidal domain thus increasing error on the borders and requires clipping of the domain. The presence of extensive laminar regions in a snapshot reduces the effectiveness of the algorithm as it reduces the weight attributed to the turbulent region. Thus only a sub-set of the original snapshot was considered for the optimisation procedure. The IP sub-set measures  $L_y \times L_z = 4.3D \times 6.4D$  containing  $N_y \times N_z = 72 \times 108$  mesh-points. The OP sub-set measures  $L_x \times L_y = 8.4D \times 4.3D$  discretised into  $N_x \times N_y = 145 \times 72$  points. A sample streamwise velocity field on the IP and the OP is shown in figure 6.11.

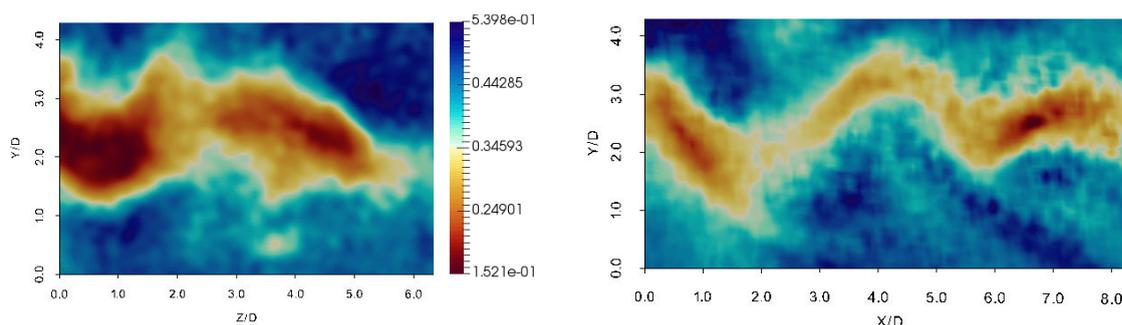


Figure 6.11: Instantaneous streamwise velocity fields extracted on the IP and the OP from PIV data-set of wake flow around a cylinder at  $Re$  of 300.

The reconstructed velocity fields on the IP are shown in figure 6.12 along with the reference IP velocity fields. A run-time of  $t_{rec} = 4.85$  s was clocked by the algorithm for this case-study. Contrary to the synthetic data-sets, the experimental observations provide a noisy velocity field. However, the main structures of the flow are still captured well by the algorithm in all three principal directions. The main large scale structures in the vortex wake are recovered by the algorithm. Significant discontinuities are observed along the spanwise direction due to the temporal discontinuity of snapshots between each neighbouring spanwise plane - methods to smoothen the spanwise spurious variations are presented in the section on post-processing techniques. This case-study is positive towards the ability of the algorithm to reconstruct the 3D domain and these results are an improvement on the mirror-imaging technique used by Robinson (2015) for the same data-set in 3D variational data-assimilation. It is also an improvement on the ‘empty box’ technique where the a numerical simulation, with a superimposed inlet condition at every time-step, is used to create a background condition. Such a methodology is time-consuming and requires high temporal resolution of the data. The SO method is extremely quick ( $\sim 8$  s) and requires only a long sequence with no limit on the temporal resolution.

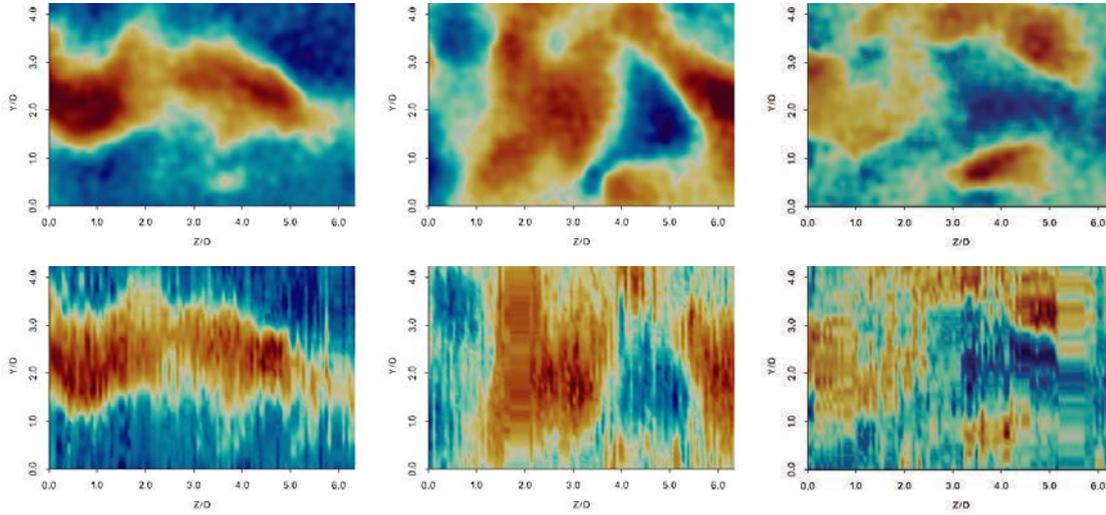


Figure 6.12: PIV wake flow instantaneous velocity fields extracted on the IP for  $Re$  300. From the left to the right: streamwise, vertical and spanwise velocity components. From top to bottom: reference and reconstruction.

### Wake flow around a circular cylinder at $Re$ 3900

With the encouraging results obtained for the laminar case study, we now expand the analysis to the more complex case study of wake flow around a cylinder in the transition regime at  $Re$  3900. The increased turbulence is associated to more noisy measurements as well as decreased spatial and temporal resolutions. A similar schematic to the previous experimental setup was used for data generation. A cylinder of diameter  $D = 12$  mm and length 260 mm was placed in the air flow. Rectangular end-plates (100 mm  $\times$  85 mm) are fixed 15 mm from the ends of the cylinder. The free stream velocity  $U$  was increased to 5 m/s in order to obtain a Reynolds in the range of 3900. On the IP 4108 image pairs are recorded and processed by Davis 10 software from LaVision. In order to account for increased turbulence combined with the noise levels generally associated with PIV data-sets, significantly higher number of snapshots (20 540) were recorded on the OP. The time-difference between two frames for one correlation field was set at 20  $\mu$ s for the IP and at 30  $\mu$ s for the OP. All images were recorded at the highest frequency supported by the camera at 1.617 kHz. The OP intersects the IP in the middle at  $6.5D$  from the centre of the cylinder.

The IP is discretised into  $N_y \times N_z = 110 \times 92$  mesh-points while the OP is discretised into  $N_x \times N_y = 69 \times 108$  mesh-points. The issue related to erroneous data at the border of the planes persists and thus the boundary points are excluded in the reconstruction. Thus, only a sub-set of the original snapshot measuring  $L_y \times L_z = 7D \times 5.6D$  on the IP and  $L_x \times L_y = 4.3D \times 7D$  on the OP are extracted. These are discretised into  $N_y \times N_z = 93 \times 72$  mesh-points for the IP and  $N_x \times N_y = 64 \times 93$  points for the OP. A sample streamwise velocity field on the IP and the OP is shown in figure 6.13

The reconstruction capabilities of the SO method and the averaging SO method are analysed with this flow configuration. The two-way reconstructed velocity fields on the IP are shown in figure 6.14 along with the reference IP velocity fields. A run-time of  $t_{rec} = 9.16$  s was clocked by the SO algorithm and  $t_{rec} = 12.14$  s by the averaging SO algorithm. Since the experimental data-set is noisy, the SO reconstructed velocity field also contains comparable noise. Thus, spanwise discontinuities, which were clearly visible for all

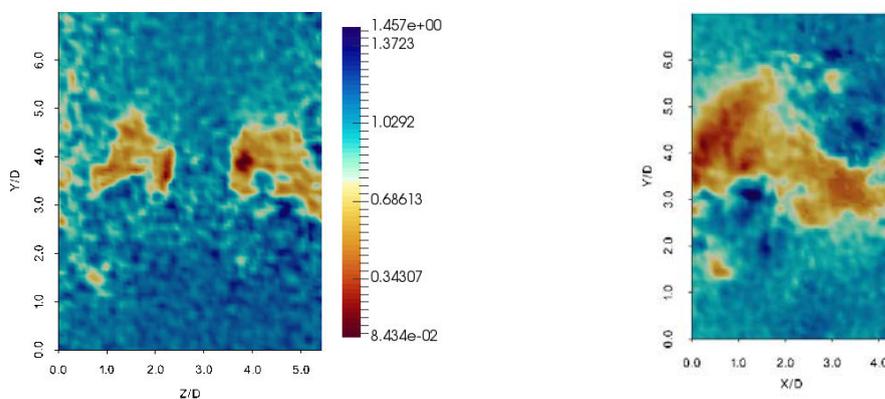


Figure 6.13: Instantaneous streamwise velocity fields extracted on the IP and the OP from PIV data-set of wake flow around a cylinder at  $Re$  3900.

previous cases reconstructed with the SO method, appear attenuated in the reconstruction. Such an attenuation is more a result of the data-deterioration from perfect synthetic or low turbulence data to noisy, highly turbulent data-set than an algorithm improvement. For the averaging SO reconstruction, a dilution of the noise is observed with minimal discontinuities - a slight attenuation of the kinetic energy can be observed with this reconstruction. Both the algorithms manage to reconstruct all the main turbulent structures in the flow as well as capturing significant small-scale structures.

The availability of a long sequence of OP snapshots allows for the study of reconstruction efficiency as a function of OP snapshots. Table 6.2 tabulates the relative error percentage between the reconstructed IP and the reference IP as a function of the number of OP snapshots used for the reconstruction. The case associated with maximum OP snapshots of 20480 is used as the base for relative error calculation. The improvement by using 20480 snapshots instead of 2048 is best highlighted in the spanwise component where we obtain a significant 57.86% improvement. A drastic error reduction is seen for all components by doubling the number of snapshots from 2048 to 4096. An inverse exponential like curve is overall observed with further increase in OP snapshots producing smaller and smaller improvements. By using 14336 snapshots, a 17% improvement is observed as compared to 2048 snapshots. However, a further increase from 14336 to 20480 produces only 1.5% improvement on the reconstruction.

These results are promising but in order to adapt the reconstruction methodology to data assimilation studies for creation of 3D backgrounds and/or 3D observations, the volumetric data-set needs to be analysed. Figure 6.15 plots the 3D volumetric iso-contours for each velocity component reconstructed from the SO and the averaging SO reconstruction techniques. Due to the 2D restriction of the experimental data-set, a comparison is not possible, however, large-scale structures can be observed in the iso-contours especially in the averaging SO case. Significant noise reduction in the averaging SO reconstruction is also seen in the iso-contours especially for the spanwise velocity in comparison with the simple case. Figures 6.14 and 6.15 are a good indication that meaningful preliminary volumetric reconstructions can indeed be obtained using the SO methods for use in data assimilation studies. The SO family of methods does not provide a true reconstruction of the field but rather a good representation of the turbulence contained in the flow for use in data assimilation.

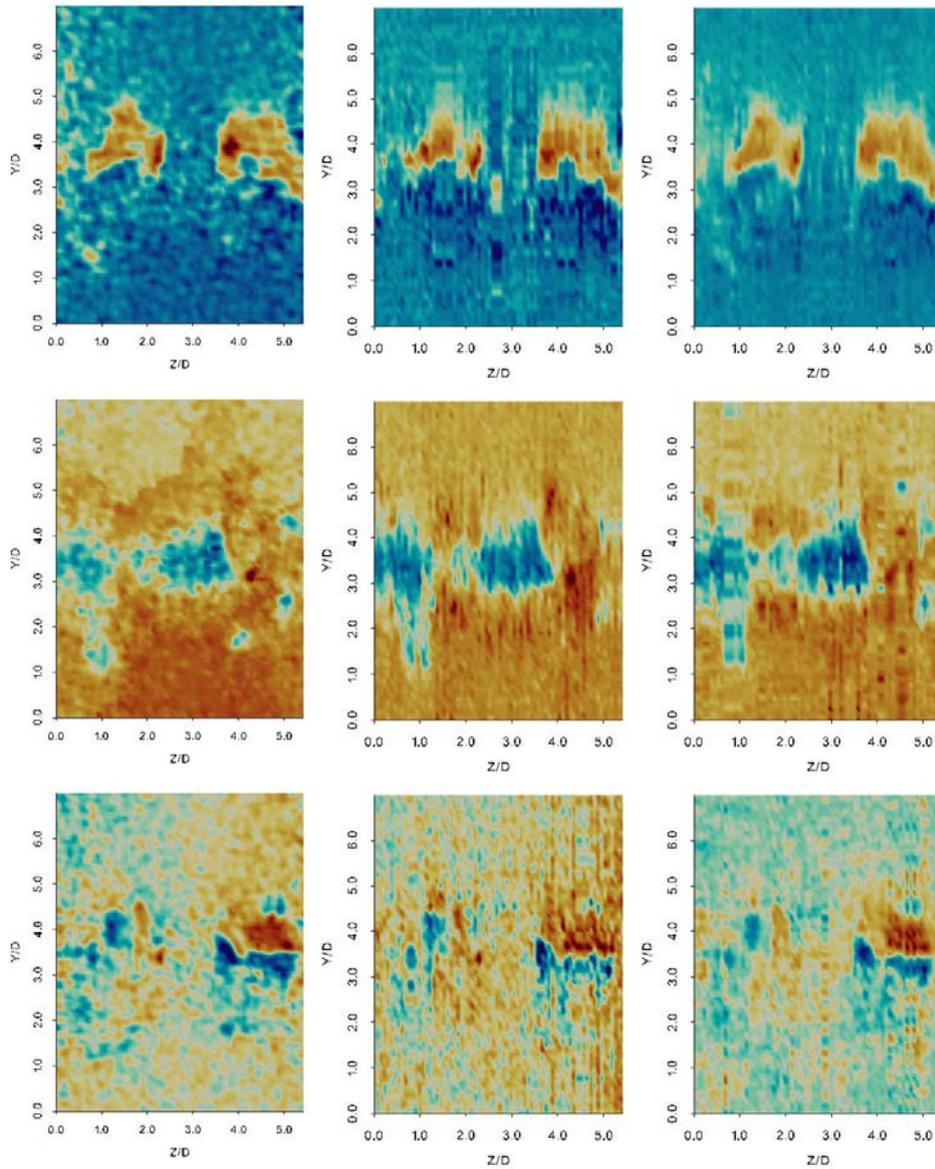


Figure 6.14: PIV wake flow instantaneous velocity fields extracted on the IP at Re of 3900. From top to bottom: streamwise, vertical and spanwise velocity components. From the left to the right: reference, SO reconstruction, and averaging SO reconstruction.

Table 6.2: Relative error percentage with respect to maximum OP snapshot case (20480) as a function of number of OP snapshots

| OP snapshots | Error %(u) | Error %(v) | Error %(w) |
|--------------|------------|------------|------------|
| 2048         | 9.34%      | 16.26%     | 57.86%     |
| 4096         | 5.42%      | 6.50%      | 12.38%     |
| 6144         | 3.17%      | 5.58%      | 9.47%      |
| 8192         | 3.17%      | 5.12%      | 8.09%      |
| 10240        | 3.17%      | 3.98%      | 5.48%      |
| 12288        | 2.00%      | 3.38%      | 4.97%      |
| 14336        | 0.00%      | 2.43%      | 4.55%      |
| 16384        | 0.00%      | 1.87%      | 3.72%      |
| 18432        | 0.00%      | 0.18%      | 1.11%      |
| 20480        | 0.00%      | 0.00%      | 0.00%      |

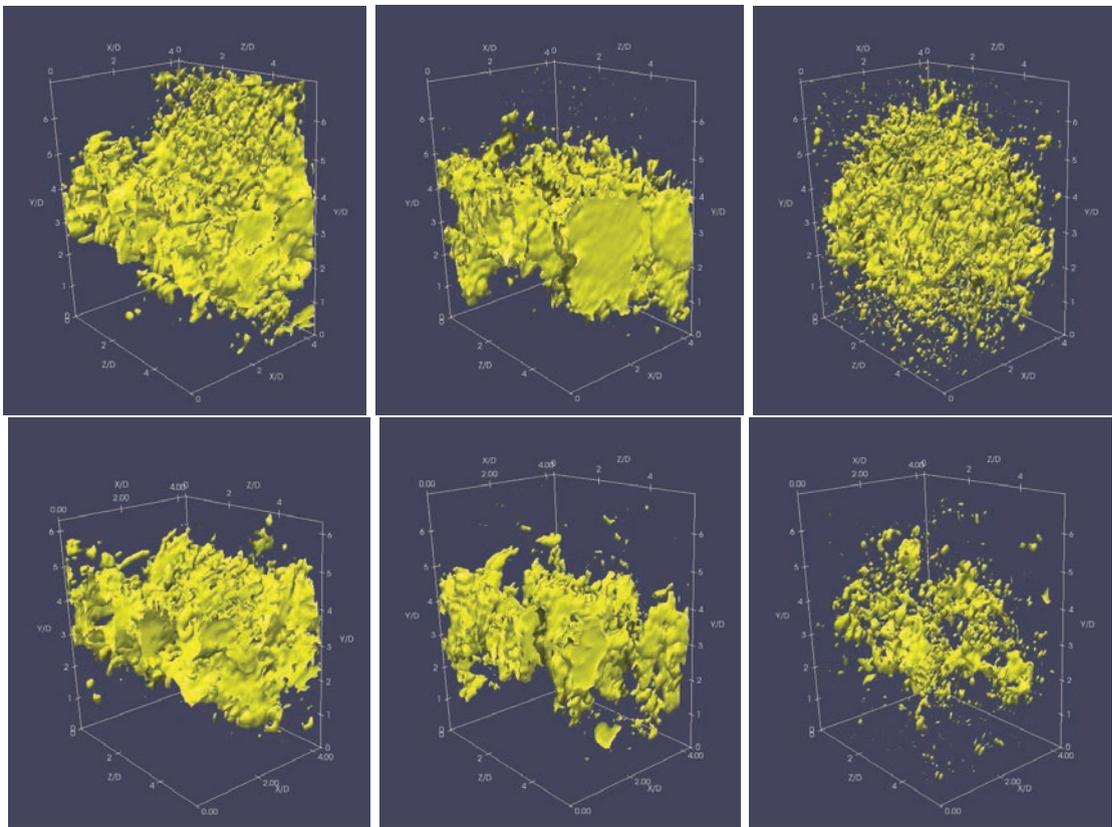


Figure 6.15: PIV wake flow instantaneous velocity iso-contours extracted on the 3D domain. From top to bottom: SO reconstruction and averaging SO reconstruction. From the left to the right: streamwise, vertical and spanwise velocity components.

### Mixing layer at $\lambda$ of 0.33

The final case study is aimed at investigating performance for a different flow configuration from those studied previously, namely the mixing layer. Another stereoscopic PIV

experiment was performed for a plane mixing layer with a modified velocity ratio  $\lambda$  of 0.33 where  $\lambda = (1 - r)/(1 + r)$  with  $r = U_{\min}/U_{\max}$ . A  $\lambda$  approaching 0 implies a stable mixing layer while a  $\lambda$  approaching 1 implies an unstable mixing layer. The modified velocity ratio was maintained with  $U_{\max} = 1$  m/s and  $U_{\min} = 0.5$  m/s. Both air flows were maintained at the same temperature. The vorticity thickness  $\delta_\omega$  of the mixing layer at the intersection between IP and OP was estimated to be equal to 51 mm. The Reynolds based on this vorticity thickness and on the velocity difference was equal to 1540. With a similar imaging setup as the wake flow case 3, 4 000 image pairs were acquired at a frequency of 600 Hz. The domain size for the measurement planes were identical to the previous case-study. The OP intersects the IP in the middle and protrudes for 10 mm behind the IP. The image pairs were analysed by the Davis 7.2 software from LaVision providing 4 000 vector fields for both OP and IP. This PIV data-set was consistent with hot-wire measurements carried out by Sodjavi and Carrier (2013) in the same wind tunnel.

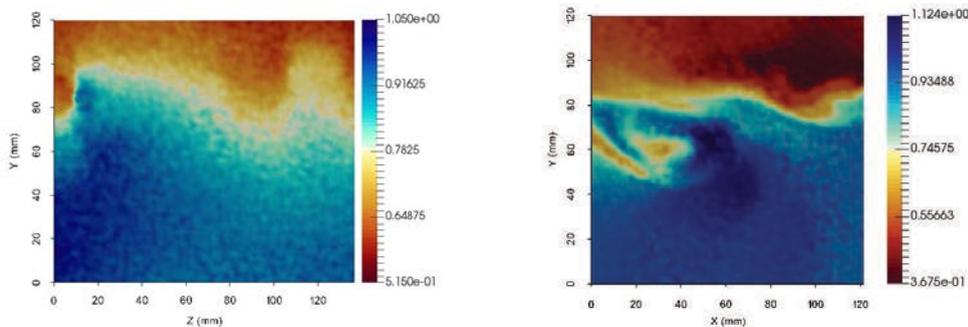


Figure 6.16: Instantaneous streamwise velocity fields extracted on the IP and the OP from PIVs data-set of the mixing layer.

The IP is discretised into  $N_y \times N_z = 175 \times 134$  mesh-points and the OP into  $N_x \times N_y = 140 \times 174$  mesh-points. Ignoring the borders and the region far away from the mixing layer, a subset of the IP measuring  $L_y \times L_z = 120 \text{ mm} \times 138 \text{ mm}$  and  $L_x \times L_y = 122 \text{ mm} \times 120 \text{ mm}$  for the OP was extracted. This sub-set contains  $N_x \times N_y = 100 \times 100$  points in the OP and  $N_y \times N_z = 100 \times 115$  points in the IP. The sample streamwise velocity field on this sub-set is provided in figure 6.16

Figure 6.17 depicts the original and reconstructed velocity fields on the inlet. The SO method required  $t_{\text{rec}} = 6.85s$  to recover the velocity field for this case-study. All three components are reconstructed quite well by the algorithm. One can see that the interface between the two fields is very accurately reconstructed by the SO method. For the vertical and spanwise components, a noisy but well-represented reconstruction is obtained. It is interesting to note here that the amount of discontinuities is quite less as compared to previous reconstructions due to smoothness of the reference velocity.

These studies clearly indicate the ability of the SO algorithm to reconstruct the flow using 2D planar data. An important advantage of such a reconstruction method is the time taken which is in the order of seconds. Most DA methods and other complex reconstruction algorithms take much longer to provide results.

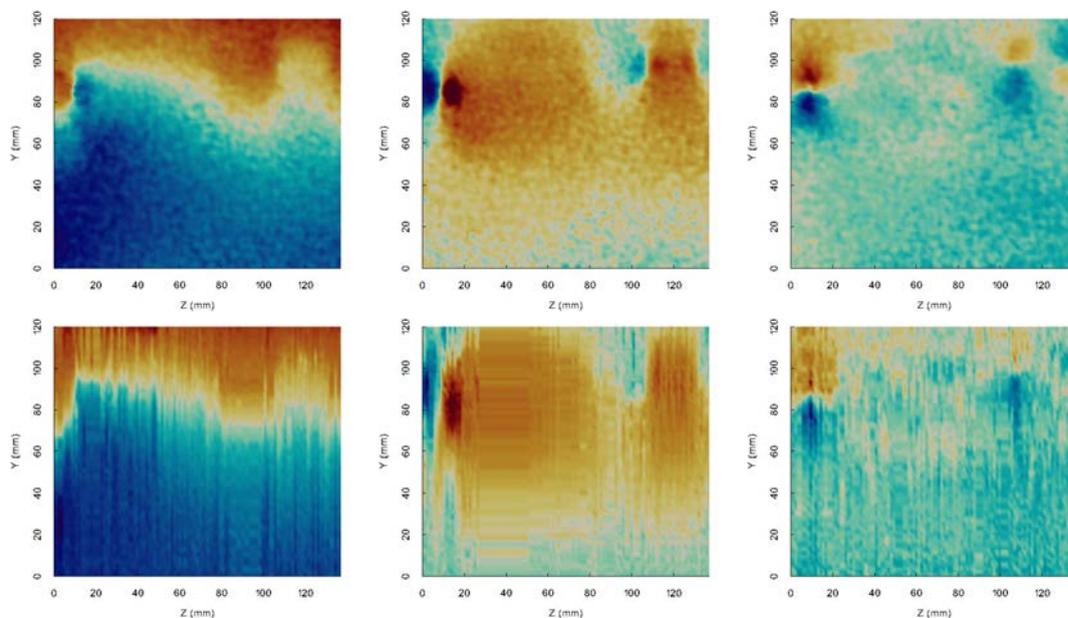


Figure 6.17: PIV mixing layer instantaneous velocity fields extracted on the IP. From top to bottom: reference and reconstruction. From the left to the right: streamwise, vertical and spanwise velocity components.

## 6.4 Conclusion

This chapter investigated a novel flow reconstruction method to obtain 3D volumetric velocity measurements using 2D planar measurements. The reconstruction algorithm is based on the assumption that considering the periodicity of the flow perpendicular to the OP, a sufficiently long time-sequence of snapshots on the OP contains at some instance within that time-sequence a representation of each of the parallel 2D planes comprising the 3D volumetric domain encompassed by the OP and the IP. Three versions of the minimisation algorithm - SO method, POD based SO method, and averaging SO method - were formulated and tested on two synthetic case studies and three experimental case studies. The algorithm is shown to reconstruct the main turbulent structures of the flow for all cases. Good performance is shown for the three wake flow studies while for the mixing layer, small scale structures are absent in the reconstruction. For a higher-dimension system such as the channel DNS data-set, the algorithm shows significant discontinuities while marginal oscillations are found in the other cases. Methods for addressing these issues are also analysed, and applied to the synthetic wake flow, showing significant reduction when POD based SO method or averaging SO methods are employed - the 3D turbulent structures are shown to match well with the reference field. The averaging SO method, which provided the best reconstruction for the synthetic case, was also applied to the experimental data-set on wake flow at Re of 3900. The method was shown to reduce noise and smooth the reconstructed velocity. The volumetric iso-contours showed meaningful large-scale structures are reconstructed using this algorithm. Techniques to ensure a divergence free reconstruction are also explored. The combination of the reconstructed velocity with Incompact3d flow solver ensures, implicitly, a divergence free field after one iteration of the time-step. This is important for 4D-Var which is based on Incompact3d and will use the SO reconstructed fields to define the background/observation conditions. Given the successful application of the SO methodology, especially with PIV sparse data-

sets, it could be an useful tool for performing variational DA studies where in an accurate background, constructed using the SO method, can help converge to the optimal solution with less iterations and thus lower cost. In addition, the use of the POD based SO can help calculate the so called background covariance matrix for variational DA which is an important weighting factor in the cost functional. The algorithm can also be expanded to reconstruct axisymmetric flows such as flow through a pipe or jet flows.

## Chapter 7

# Data assimilation - wake flow at $Re = 3900$

The varied tools for performing data assimilation incorporating LES have been established in the previous chapters. The concepts for variational data assimilation with incremental assimilation and additional control parameters have been established in chapter 5. The LES models were enumerated in chapter 2 and analysed in chapter 3 - the StSp model was identified to be an optimal yet easy to implement LES model for use with DA. An easy, and fast reconstruction method was developed and analysed in chapter 6 which can be used to create an improved background condition for the DA procedure. In this chapter, all these tools and techniques are combined together to perform data assimilation of wake flow around a circular cylinder at a transitional  $Re$  of 3900. The following section delves briefly into the 4D-Var code of Robinson (2015) along with the automatic differentiation (AD) principles behind the code formulation. The section also enumerates the inclusion of the LES model into the code giving the 4D-Var<sub>les</sub> code. This is followed by a section on the application of the code to wake flow using synthetic 3D observations as a validation test-case. The final section assesses the capabilities of the code to assimilate synthetic cross-plane observations emulating experimental data-sets.

## 7.1 Code Formulation

### 7.1.1 4D-Var

The 4D-Var code used in this thesis was built on the foundations of the work done by Gronskis et al. (2013) who performed the tangent and adjoint code formulation for the 2D version of Incompact3d (the details of the DNS flow solver Incompact3d is given in §3.1). Robinson (2015) expanded the tangent-adjoint formulation to incorporate the 3D parallel version of the flow solver constructing the 4D-Var code. In order to perform variational assimilation using the adjoint method, the backward integration of the adjoint model is necessary to calculate the gradient.

The construction of the adjoint model can be envisaged in two ways: the differentiate-then-discretise method involves the formulation of mathematical adjoint which is then discretised on the numerical mesh while the discretise-then-differentiate method constructs the discrete adjoint directly on the discretised non-linear model. The two methods are also referred to as continuous adjoint and discrete adjoint depending on the formulation of the adjoint model (see figure 7.1). Both methods have been studied and compared in literature in various scenarios (Nadarajah and Jameson, 2000; Griesse and Walther, 2004).

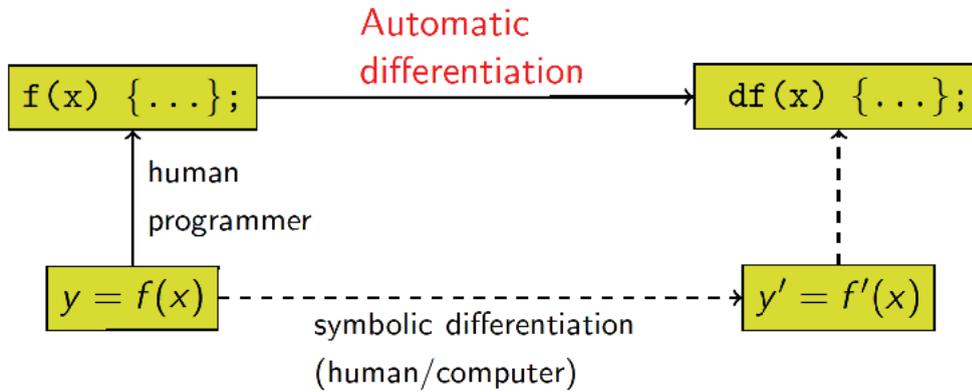


Figure 7.1: Adjoint model construction methods for the analytical function  $y = f(x)$ . Discretise-then-differentiate expressed as a solid line and differentiate-then-discretise as a dotted line.

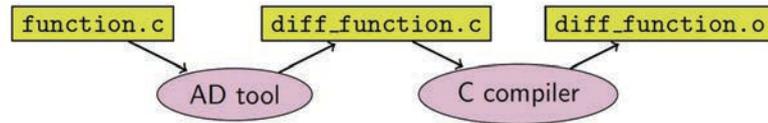
Clearly, a choice has to be made on which approach is to be taken for formulating the adjoint model. The continuous adjoint poses certain restrictions: the accuracy of this adjoint is of the order of discretisation errors, boundary terms which arise through integration by parts need to be handled separately using appropriate assumptions. The discrete adjoint, however, is limited in accuracy by the original dynamical model code itself. Thus for a highly accurate numerical scheme, the discrete adjoint is capable of providing an accurate adjoint up to machine accuracy while not requiring any external assumption for treating terms. However, the discrete adjoint formulation can be tricky and can require extensive and careful validation.

In literature, variational DA studies have preferred the continuous adjoint for 2D (Lemke and Sesterhenn, 2016; Mons et al., 2014). A combination of the discrete adjoint with the flow solver Incompact3d was done by Gronsksis et al. (2013) - the discrete strategy was shown to provide the exact adjoint for the discrete code up to machine accuracy. In addition, the discrete adjoint enables to construct the gradient descent direction of the ‘true’ energy that needs to be minimised, i.e. the one constructed from the numerical code. Continuous adjoint, on the other hand, provides only an approximation of it through eventually different numerical choices for the direct and adjoint dynamics. Thus, the discrete adjoint technique was also used by Robinson for formulating the adjoint model code of the parallelised 3D version of incompact3d using automatic differentiation (AD) techniques.

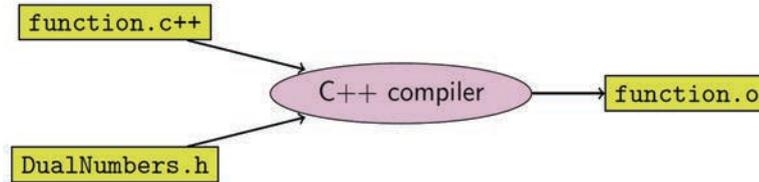
### Automatic Differentiation

AD techniques, also called algorithmic differentiation or computational differentiation, are used to numerically evaluate the derivative of a function specified by a computer program. The derivative is evaluated by exploiting the sequential nature of computers where each function is evaluated by performing a series of elementary arithmetic functions. Thus, the derivative of this function can be obtained by repeated application of the chain rule to these arithmetic operations. The numerical application of AD principles can be done in one of two ways (see figure 7.2):

**Source Code Transformation (SCT)** replaces the original source code for a function by an automatically generated code that simultaneously calculates the derivative of each step of the original code alongside the original instructions. Additional variables, arrays and matrices are automatically generated to hold the derivatives and any tempo-



(a) Source Code Transformation: original function (function.c) is converted by AD tool to give diff\_function.c including the differential and further compiled,



(b) Operator Overloading: original function (function.c++) is passed to the compiler along with operator overloading function (DualNumbers.h) to give compiled output,

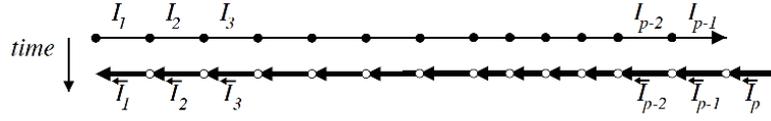
Figure 7.2: Flow chart of Automatic Differentiation methodologies.

rary storage required. Using SCT, the reverse mode, also called the adjoint code, can also be implemented easily. However, the complexity of the source code has a direct parallel to the complexity of the AD code. Highly complex codes generally need to be split into smaller set of instructions before being input in to an AD tool for performing SCT and the resultant programs can be combined into an efficient code. The resultant code is generally of much larger size compared to the original and is not possible to perform by hand and require sophisticated AD tools to perform SCT accurately and rapidly.

**Operator Overloading (OO)** replaces every real number and elementary mathematical operators with a dual set with the secondary variable containing the differential and the new operator working on the pair of reals computing the function value and its differential. Thus, there is very little change to original source code but for a change of data-types for the variables which is also performed at compile time - this can lead to increased compiler time. However, OO is only supported by specific languages and is thus restricted in application to such languages. Within OO, the forward or linear trajectory is easily implemented while the reverse mode is more complex to implement due to current compiler restrictions.

The AD tool TAPENADE, developed at INRIA Sophia-Antipolis by the Tropic and Ecuador teams, is based on the SCT methodology and is used to perform the tangent and adjoint code formulation for this thesis. This is along the same lines as followed by Cordelia and by Gronskis before that. Regarding the development and principles behind the automatic differentiation engine TAPENADE, please refer to Hascoët and Pascual (2004, 2013). TAPENADE, a successor to Odyssee 1.7, works with FORTRAN or C codes providing a resultant code that calculates the tangent or the adjoint derivatives depending on user requirement. A web-interface is available for simple applications while the AD tool can also be downloaded for use from command line - the latter is preferred for complex scenarios. The tool is implemented 90% in JAVA and the remaining in C and its latest version is suitable for codes written in FORTRAN95, such as Incompact3d. Applications of the tool to optimal control problems, oceanography, and DA can be found in Ferron and Hascoët (1996); Tber et al. (2007); Gronskis et al. (2013).

Within an AD tool, the forward or tangent mode is easily implemented but the adjoint or reverse mode is more complex due to issues arising from the need for values of all intermediary variables in the forward trajectory. Consider a non-linear dynamical model



(a) Store all strategy for adjoint code formulation using AD tools,



(b) Recompute all strategy for adjoint code formulation using AD tools,

Figure 7.3: Classical strategies for computing the adjoint of a code (Tber et al., 2007).

expressed as,

$$\mathcal{X}_0 \xrightarrow{I_1} \dots \xrightarrow{I_j} \mathcal{X}_j = I_j(\mathcal{X}_{j-1}) \xrightarrow{I_{j+1}} \dots \xrightarrow{I_f} \mathcal{X}_f, \quad (7.1.1)$$

its linearised version (tangent procedure) can be written as,

$$\delta \mathcal{X}_0 \xrightarrow{I'_1} \dots \xrightarrow{I'_j} \delta \mathcal{X}_j = I'_j(\mathcal{X}_{j-1}) \delta \mathcal{X}_{j-1} \xrightarrow{I'_{j+1}} \dots \xrightarrow{I'_f} \delta \mathcal{X}_f, \quad (7.1.2)$$

and the corresponding adjoint mode can be computed using,

$$\lambda_0 \xleftarrow{I_1^*} \dots \xleftarrow{I_j^*} \lambda_j = I_{j+1}^*(\mathcal{X}_{j-1}) \lambda_{j+1} \xleftarrow{I_{j+1}^*} \dots \xleftarrow{I_p^*} \lambda_p. \quad (7.1.3)$$

Thus to calculate  $\lambda_j$ , the value of  $\mathcal{X}_{j-1}$  is required along with  $\lambda_{j+1}$ . For a temporal model which is performed over sufficient number of time-steps the storing of these variables can be clearly problematic in terms of memory requirements. Such a method of storing all the intermediary values is termed as the ‘store-all’ strategy (see figure 7.3a). During a forward preliminary run of the code called the forward sweep, all the intermediary values indicated by black filled in dots are stored in a stack. Then, during the backward sweep for the adjoint, the stack variables are recalled, indicated by the white dots in figure 7.3a, and used as required by the code. A linear memory requirement increase is observed with increasing  $p$ .

An alternative option of minimal energy storage is to recompute from the initial condition  $\mathcal{X}_0$  for every intermediary step  $j$ , the value of the variable  $\mathcal{X}_{j-1}$ . This strategy, termed as ‘recompute-all’ solves the memory overload issue, however, at the cost of high computational cost (see figure 7.3b). The black dots indicate the initial stored state of the state-space variable from which the re-computation is performed for all intermediary steps  $p$ . A brute-force strategy involving recompute-all has a quadratic relation between the computational cost and number of intermediary steps.

For a large code, such as Incompact3d, neither of these strategies are ideal. A combined strategy involving a trade-off between the two classical strategies is required for efficiency.

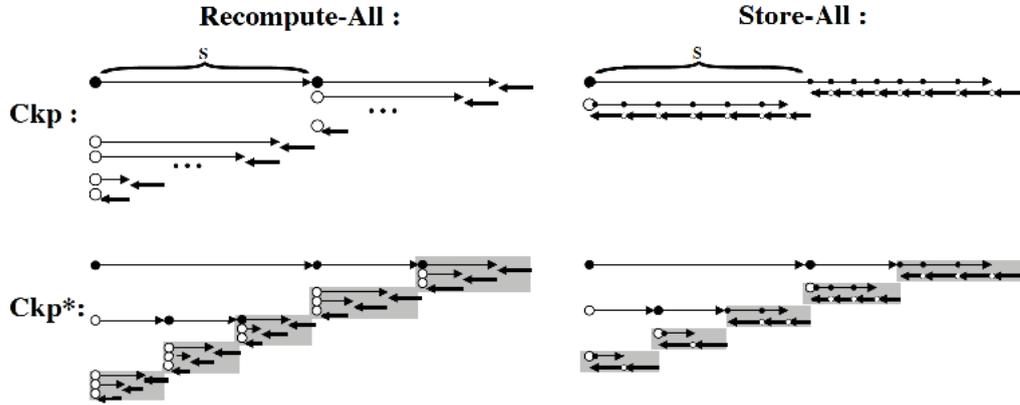


Figure 7.4: Checkpointing strategies for adjoint code formulation with both classical strategies.

Such a trade-off is referred to as the ‘checkpointing’ strategy where the original code is check-pointed at specific intermediary points where the value are stored. All intermediary values between two checkpoints can then be easily obtained when required by the adjoint program by using either the classical strategies depending on the importance of memory/computational time. The variations of the checkpointing strategy with a single checkpoint (**Ckp**) and multiple checkpoints (**Ckp\***) is shown in figure 7.4.

TAPENADE employs a store-all strategy with checkpointing for procedure calls. However, due to the complex nature of the dynamical code, the inherent strategy of TAPENADE cannot be directly applied for the 4D-Var case. User defined check-pointing is performed (generally at each time-step where observations are assimilated) and the recompute-all strategy is used from one time-step to the next due to memory restrictions while within time-steps the store-all strategy is implemented to reduce computational cost. As input to TAPENADE, a function code can rarely be provided entirely as the numerous function calls within cannot be processed aptly by the AD tools. Comprehensive pre-processing and post-processing of the code is required to obtain an accurate, efficient and correct tangent/adjoint procedure.

## Validation

The process of creating a tangent or adjoint formulation is not simple and is extremely sensitive to numerical and human errors. Thus a stringent validation process is required to ensure the output of the AD tool is accurate before it can be incorporated into the 4D-Var code. The tangent procedures are validated first and the adjoint validation is done using the verified tangent code.

### Tangent Validation

We consider Taylor’s expansion of a function  $f$  around a state  $\mathcal{X}_0$  in a given direction  $\delta$ ,

$$f(\mathcal{X}_0 + \alpha\delta) \approx f(\mathcal{X}_0) + \alpha \langle \nabla f(\mathcal{X}_0), \delta \rangle . \quad (7.1.4)$$

Thus, the tangent procedures can be validated by monitoring the ratio,

$$\Pi_1 = \frac{f(\mathcal{X}_0 + \alpha\boldsymbol{\delta}) - f(\mathcal{X}_0)}{\alpha \langle \nabla f(\mathcal{X}_0), \boldsymbol{\delta} \rangle} \quad (7.1.5)$$

A value of  $\Pi_1$  around 1 for orders of  $\alpha$  ranging from  $[10^{-3} : 10^{-9}]$  is indicative of a good tangent code created from the AD tool. The limited range for  $\alpha$  is due to the contributions of round-off errors at lower values below  $10^{-9}$  and errors introduced due to model non-linearity at large values of  $\alpha$ . All the tangent formulations used in this thesis have been verified strictly with this validation test.

### Adjoint Validation

The adjoint model created using an AD tool in reverse mode can be validated with the help of verified tangent model for the same non-linear function. The validation for the adjoint procedure is performed by considering the definition of an adjoint operator  $(\partial_{\mathcal{X}}\mathbb{M})^*$  with respect to two variables  $\mathbf{X}$  and  $\mathbf{Y}$ ,

$$\langle (\partial_{\mathcal{X}}\mathbb{M})\mathbf{X}, \mathbf{Y} \rangle = \langle \mathbf{X}, (\partial_{\mathcal{X}}\mathbb{M})^*\mathbf{Y} \rangle \quad (7.1.6)$$

If we enforce the variable  $\mathbf{Y}$  to be the output of the tangent linearised model  $(\partial_{\mathcal{X}}\mathbb{M})$  applied to  $\mathbf{X}$  (i.e.  $\mathbf{Y} = (\partial_{\mathcal{X}}\mathbb{M})\mathbf{X}$ ), define  $\bar{\mathbf{X}} = (\partial_{\mathcal{X}}\mathbb{M})^*(\partial_{\mathcal{X}}\mathbb{M})\mathbf{X} = (\partial_{\mathcal{X}}\mathbb{M})^*\mathbf{Y}$ , and take the dot product of  $\bar{\mathbf{X}}$  and  $\mathbf{X}$ , we get,

$$\langle \mathbf{X}, \bar{\mathbf{X}} \rangle = \langle \mathbf{X}, (\partial_{\mathcal{X}}\mathbb{M})^*\mathbf{Y} \rangle = \langle (\partial_{\mathcal{X}}\mathbb{M})\mathbf{X}, \mathbf{Y} \rangle = \langle \mathbf{Y}, \mathbf{Y} \rangle. \quad (7.1.7)$$

Thus, the validity of the adjoint procedure can be analysed by the ratio,

$$\Pi_2 = \frac{\langle \mathbf{Y}, \mathbf{Y} \rangle}{\langle \mathbf{X}, \bar{\mathbf{X}} \rangle} = \frac{\langle (\partial_{\mathcal{X}}\mathbb{M})\mathbf{X}, (\partial_{\mathcal{X}}\mathbb{M})\mathbf{X} \rangle}{\langle \mathbf{X}, (\partial_{\mathcal{X}}\mathbb{M})^*\mathbf{Y} \rangle} \approx 1. \quad (7.1.8)$$

In order to calculate  $\Pi_2$ ,  $\mathbf{Y}$  is calculated by application of the tangent differentiated version of the function to  $\mathbf{X}$  while  $\bar{\mathbf{X}}$  is obtained from the adjoint differentiated function applied to  $\mathbf{Y}$ . Unlike the tangent validation, the adjoint validation is strict with admissible error only of the order of machine accuracy due to the higher accuracy standards required for the adjoint model.

The 4D-Var code has been stringently validated by Robinson (2015). For this thesis, all additional/new tangent-adjoint procedures formulated using TAPENADE have been validated using the above described validation procedures up to admissible errors. All functions have been individually validated followed by a final validation of the entirety of the tangent-adjoint model code.

### Remarks

**Discrete adjoint of parallel-MPI function:** Incompact3d utilises numerous MPI routines for its parallel procedures. For the adjoint model, the adjoint routines for the MPI functions need to be formed. The numerical adjoint/gradient for MPI routines has been studied in literature by Faure and Dutto (1999); Utke et al. (2009, 2012); Schanen et al. (2010). Figure 7.5 tabulates MPI routines and their corresponding adjoint routines.  $x$  denotes a normal variable while  $\bar{x}$  denotes the adjoint variable, and  $P$  stands for a specific processor to which the MPI routine relates to.

| Original                                       | Adjoint   |
|--|---|
| <code>MPI_bcast(x, ... ,P, ...)</code>         | <code>MPI_reduce (<math>\bar{x}, \bar{t}, \dots</math>, SUM, P, ...)</code><br>$\bar{x} = 0.0$ ; on P : $\bar{x} = \bar{x} + \bar{t}$               |
| <code>MPI_reduce(x,y, ... ,SUM, P, ...)</code> | on P : $\bar{t} = \bar{y}$ ; on P : $\bar{y} = 0.0$<br><code>MPI_bcast (<math>\bar{t}, \dots</math>, P, ...)</code> ; $\bar{x} = \bar{x} + \bar{t}$ |
| <code>MPI_allreduce(x,y, ... ,SUM, ...)</code> | <code>MPI_allreduce (<math>\bar{y}, \bar{t}, \dots</math>, SUM, ...)</code><br>$\bar{y} = 0.0$ ; $\bar{x} = \bar{x} + \bar{t}$                      |
| <code>MPI_gather(x,y, ... ,P, ...)</code>      | <code>MPI_scatter (<math>\bar{y}, \dots, \bar{t}, P, \dots</math>)</code><br>on P : $\bar{y} = 0.0$ ; $\bar{x} = \bar{x} + \bar{t}$                 |
| <code>MPI_scatter(y, ... ,t, P, ...)</code>    | <code>MPI_gather (<math>\bar{y}, \dots, \bar{t}(:), P, \dots</math>)</code><br>on $\bar{y} = 0.0$ ; on P : $\bar{x} = \bar{x} + \bar{t}(:)$         |

Figure 7.5: Discrete adjoint for parallel-MPI functions

**Discrete adjoint of pencil-swap functions:** Due to the `decomp2d` library of `Incompact3d`, each processor work with a pencil grid depending on the principle direction considered. In order to switch between different pencils (see figure 3.1), a transpose function is called depending on the required swap. The existence of a duality between forward and adjoint MPI routines, shown by Faure and Dutto (1999); Cheng (2006), can be extended here to the transpose routines of `Incompact3d` (Robinson, 2015). Thus, the adjoint for a transpose routine call from  $x$  to  $y$  or `transpose_x_to_y` is the transpose routine call from  $y$  to  $x$  or `transpose_y_to_x` and likewise for all the other transpose sub-routines.

**Incompact3d adjoint algorithm:** For a given function, `func`, its tangent is denoted as `func_d` and its adjoint as `func_b` in line with TAPENADE notations. The same is followed for variables and their tangent/adjoint form as well. Consider a velocity field  $\mathbf{u}$  with a corresponding pressure  $p$  simulated for  $N$  time-steps with `Incompact3d`, its backward adjoint trajectory is summarised by algorithm 2.  $u^*$  and  $u^{**}$  stand for velocity values at fractional time-steps used by the temporal scheme while  $F^i$  stands for the convection + diffusion contributions at time-step  $i$ . Table 7.1 summarises each function and its objective in the numerical code.

---

**Algorithm 2** `Incompact3d_b` ( $\mathbf{u}^0, \mathbf{u}_b^0$ )

---

```

Initialisation:  $i = N, \mathbf{u}_b^N = 0$ 
while  $i > 0$  do
  corgp_b( $\mathbf{u}^i, \mathbf{u}_b^i, \nabla p, \nabla p_b$ )
  gradp_b( $\nabla p, \nabla p_b, p, p_b$ )
  poisson_b( $p, p_b, \mathbf{u}^{**}, \mathbf{u}_b^{**}$ )
  pre_correc_b( $p, p_b$ )
  intt_b( $\mathbf{u}^*, \mathbf{u}_b^*, F^{i-1}, F_b^{i-1}, F^{i-2}, F_b^{i-2}$ )
  convdif_b( $\mathbf{u}^{t-1}, \mathbf{u}_b^{t-1}, F^{i-1}, F_b^{i-1}, F^{i-2}, F_b^{i-2}$ )  $i = i - 1$ 
end while

```

---

For the discrete adjoint of each individual function, please refer to Robinson (2015).

Table 7.1: Incompact3d functions and their objectives

| Function    | Objective   |
|-------------|---|
| convdiff    | Calculate convection and diffusion terms                    |
| intt        | Explicit time advancement of the flow field                 |
| pre_correc  | Enforce boundary condition through pressure gradient        |
| poisson     | Poisson equation solver                                     |
| gradp       | Calculate pressure gradient                                 |
| corgp       | Correct velocity by pressure gradient for incompressibility |
| Incompact3d | Parent function calling all modules                         |

### 7.1.2 4D-Var<sub>les</sub>

The various sub-functions of the DNS flow solver - Incompact3d, as shown in figure 3.2, is recalled here in figure 7.6. In order to perform 4D-Var DA, the adjoint mode for each of these functions need to be formulated using an AD tool like TAPENADE. The corresponding adjoint mode for the 4D-Var algorithm is shown as a flow-chart in figure 7.7 - the simulation direction, as required, is backwards in time and order.

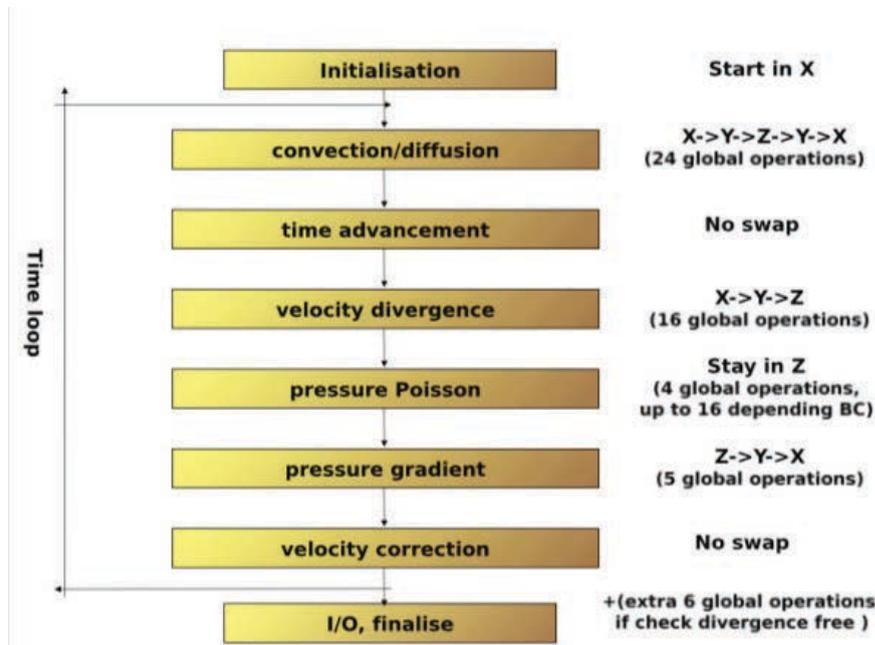


Figure 7.6: General structure of the DNS flow solver - Incompact3d

In order to incorporate an LES model into 4D-Var, the numerical model needs to be first implemented in Incompact3d. This model can then be fed as an input to TAPENADE obtaining the tangent and adjoint version of the model as the output to be included into 4D-Var. The model chosen, as iterated before, is the stochastic spatial variance version of the models under location uncertainty - despite the name, the model is applied in a deterministic sense (for more details, please refer to §2.2.4). A numerical expression of the

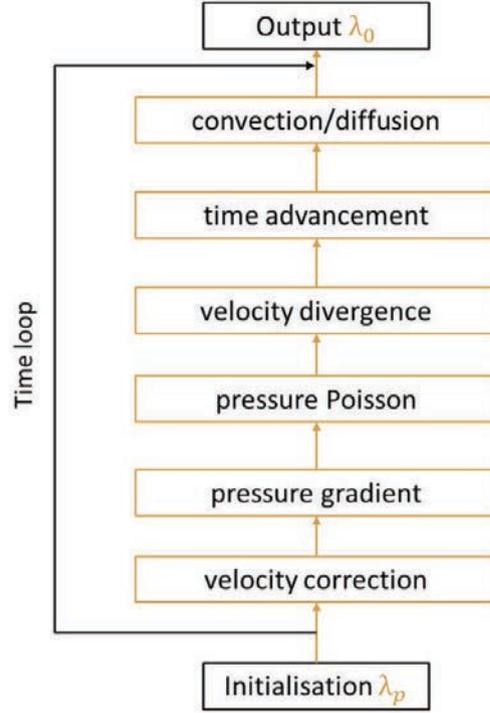


Figure 7.7: Adjoint model structure of 4D-Var for the DNS flow solver - Incompact3d

model requires four distinct steps,

- Numerical determination of the variance tensor  $\mathbf{a}$  from the large scale resolved velocity field,
- Calculation of the sub-grid scale dissipation  $\left(\frac{1}{2} \sum_{ij} \partial_{x_i} (a_{ij} \partial_{x_j} \mathbf{u})\right)$  from the variance tensor,
- Evaluation of the bias contribution to the large-scale advection  $\left(\left(\frac{1}{2} \nabla \cdot \mathbf{a}\right) \cdot \nabla \mathbf{u}\right)$ ,
- Evaluation of the model contribution  $(\nabla \cdot \nabla \cdot \mathbf{a})$  to the divergence free pressure field.

The modification of Incompact3d to incorporate the LES model is graphically represented in figure 7.8. The first three steps are evaluated and incorporated in the code *a posteriori* to the convection/diffusion calculation. The final step for the divergence contribution, while evaluated along with other model contributions, is incorporated in the velocity divergence subroutine to modify the pressure ensuring an incompressible flow. The addition functions needed for the LES model and their corresponding contribution are summarised in table 7.2.

In the following, the algorithms for the forward non-linear, tangent and backward adjoint modes for each LES model function will be enumerated.

**Algorithm notations:**

- Bold letters ( $\mathbf{i}$ ) indicate vectors.
- Superscript  $()^d$  and  $\_d$  denotes tangent variables/modules/functions.

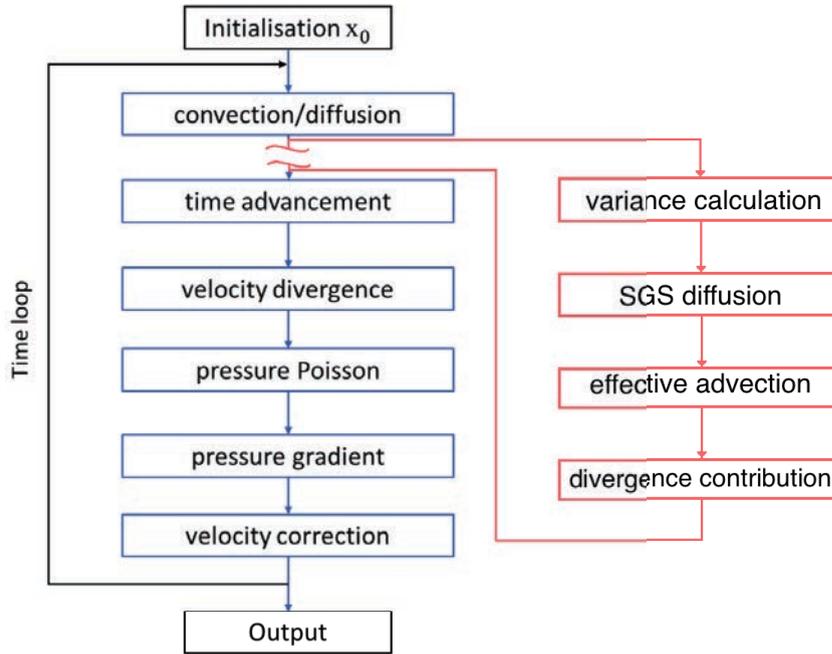


Figure 7.8: Flow chart for the LES version of Incompact3d with the StSp model. The original DNS functions are represented in blue boxes while the new LES model functions are shown in red boxes.

Table 7.2: LES model functions added to Incompact3d and their objectives

| Function          | Objective   |
|-------------------|---|
| variance - stsp   | Evaluate the variance tensor $\mathbf{a}$ from resolved velocity      |
| dissipation - les | Calculate sub-grid dissipation contribution                           |
| advection - les   | Calculate advection bias introduced by $\mathbf{a}$                   |
| divergence - les  | Evaluate $\nabla \cdot \nabla \cdot \mathbf{a}$ for incompressibility |
| LES               | Parent routine calling all LES model functions                        |

- Superscript  $()^b$  and  $\_b$  denotes adjoint variables/modules/functions.
- $x, y, z$  as subscript indicate the corresponding component of the variable.
- $i, j, k$  as subscript or within square bracket  $([i])$  indicate an implicit summation of each index in all three directions.
- $tmp$  stands for temporary variables.
- $der*$  is the intrinsic derivative functions with  $der * \_d$  and  $der * \_b$  standing for the AD derived tangent and adjoint derivative functions..
- LES is the parent program calling all LES model functions.

**Forward Non-linear Mode**

---

**Algorithm 3** variance – stsp ( $\mathbf{u}, \mathbf{a}, C_{sp}$ )

---

 $\mathbf{u}_m = avg_{\Omega}(\mathbf{u})$  - Spatial averaging over domain  $\Omega$  $a_{ij} = [(u_i - u_{m,i})(u_j - u_{m,j})]C_{sp}$ 

---

---

**Algorithm 4** dissipation – les ( $\mathbf{u}, \nabla\mathbf{u}, \mathbf{a}, lesdiss$ )

---

 $tmp_{ij} = (\nabla_i u_x) a_{ij}$  $lesdiss_x = \frac{1}{2}[\sum_j(\text{derx}(tmp_{1j}) + \text{dery}(tmp_{2j}) + \text{derz}(tmp_{3j}))]$  $tmp_{ij} = (\nabla_i u_y) a_{ij}$  $lesdiss_y = \frac{1}{2}[\sum_j(\text{derx}(tmp_{1j}) + \text{dery}(tmp_{2j}) + \text{derz}(tmp_{3j}))]$  $tmp_{ij} = (\nabla_i u_z) a_{ij}$  $lesdiss_z = \frac{1}{2}[\sum_j(\text{derx}(tmp_{1j}) + \text{dery}(tmp_{2j}) + \text{derz}(tmp_{3j}))]$ 

---

---

**Algorithm 5** advection – les ( $\mathbf{u}, \mathbf{a}, lesadv$ )

---

 $tmp_x = \frac{1}{2}[\text{derx}(a_{xx}) + \text{dery}(a_{xy}) + \text{derz}(a_{xz})]$  $tmp_y = \frac{1}{2}[\text{derx}(a_{yx}) + \text{dery}(a_{yy}) + \text{derz}(a_{yz})]$  $tmp_z = \frac{1}{2}[\text{derx}(a_{zx}) + \text{dery}(a_{zy}) + \text{derz}(a_{zz})]$  $lesadv_x = tmp_x \text{derx}(u_x) + tmp_y \text{dery}(u_x) + tmp_z \text{derz}(u_x)$  $lesadv_y = tmp_x \text{derx}(u_y) + tmp_y \text{dery}(u_y) + tmp_z \text{derz}(u_y)$  $lesadv_z = tmp_x \text{derx}(u_z) + tmp_y \text{dery}(u_z) + tmp_z \text{derz}(u_z)$ 

---

---

**Algorithm 6** divergence – les ( $\mathbf{a}, lesdiv$ )

---

 $lesdiv_x = \text{derx}(a_{xx}) + \text{dery}(a_{xy}) + \text{derz}(a_{xz})$  $lesdiv_y = \text{derx}(a_{yx}) + \text{dery}(a_{yy}) + \text{derz}(a_{yz})$  $lesdiv_z = \text{derx}(a_{zx}) + \text{dery}(a_{zy}) + \text{derz}(a_{zz})$ 

---

---

**Algorithm 7** LES ( $\mathbf{u}, C_{sp}, lesdiss, lesadv, lesdiv$ )

---

variance – stsp ( $\mathbf{u}, \mathbf{a}, C_{sp}$ ) $\partial_i \mathbf{u} = \text{der}[\mathbf{i}](\mathbf{u})$ dissipation – les ( $\mathbf{u}, \nabla\mathbf{u}, \mathbf{a}, lesdiss$ )advection – les ( $\mathbf{u}, \mathbf{a}, lesadv$ )divergence – les ( $\mathbf{a}, lesdiv$ )

---

## Forward Tangent Mode

---

**Algorithm 8** variance –  $\text{stsp\_d}(\mathbf{u}, \mathbf{u}^d, \mathbf{a}, \mathbf{a}^d, C_{\text{sp}})$ 


---

$$\mathbf{u}_m = \text{avg}_\Omega(\mathbf{u})$$

$$\mathbf{u}_m^d = \text{avg}_\Omega(\mathbf{u}^d)$$

$$a_{ij}^d = [(u_i - u_{m,i})(u_j^d - u_{m,j}^d) + (u_i^d - u_{m,i}^d)(u_j - u_{m,j})]C_{\text{sp}}$$


---

---

**Algorithm 9** dissipation –  $\text{les\_d}(\mathbf{u}, \mathbf{u}^d, \nabla\mathbf{u}, \nabla\mathbf{u}^d, \mathbf{a}, \mathbf{a}^d, \text{lesdiss}^d)$ 


---

$$\text{tmp}_{ij} = (\nabla_i u_x^d) a_{ij} + (\nabla_i u_x) a_{ij}^d$$

$$\text{lesdiss}_x^d = \frac{1}{2} [\sum_j (\text{derx\_d}(\text{tmp}_{1j}) + \text{dery\_d}(\text{tmp}_{2j}) + \text{derz\_d}(\text{tmp}_{3j}))]$$

$$\text{tmp}_{ij} = (\nabla_i u_y^d) a_{ij} + (\nabla_i u_y) a_{ij}^d$$

$$\text{lesdiss}_y = \frac{1}{2} [\sum_j (\text{derx\_d}(\text{tmp}_{1j}) + \text{dery\_d}(\text{tmp}_{2j}) + \text{derz\_d}(\text{tmp}_{3j}))]$$

$$\text{tmp}_{ij} = (\nabla_i u_z^d) a_{ij} + (\nabla_i u_z) a_{ij}^d$$

$$\text{lesdiss}_z = \frac{1}{2} [\sum_j (\text{derx\_d}(\text{tmp}_{1j}) + \text{dery\_d}(\text{tmp}_{2j}) + \text{derz\_d}(\text{tmp}_{3j}))]$$


---

---

**Algorithm 10** advection –  $\text{les\_d}(\mathbf{u}, \mathbf{u}^d, \mathbf{a}, \mathbf{a}^d, \text{lesadv}^d)$ 


---

$$\text{tmp}_x = \frac{1}{2} [\text{derx}(a_{xx}) + \text{dery}(a_{xy}) + \text{derz}(a_{xz})]$$

$$\text{tmp}_y = \frac{1}{2} [\text{derx}(a_{yx}) + \text{dery}(a_{yy}) + \text{derz}(a_{yz})]$$

$$\text{tmp}_z = \frac{1}{2} [\text{derx}(a_{zx}) + \text{dery}(a_{zy}) + \text{derz}(a_{zz})]$$

$$\text{tmp}_x^d = \frac{1}{2} [\text{derx\_d}(a_{xx}^d) + \text{dery\_d}(a_{xy}^d) + \text{derz\_d}(a_{xz}^d)]$$

$$\text{tmp}_y^d = \frac{1}{2} [\text{derx\_d}(a_{yx}^d) + \text{dery\_d}(a_{yy}^d) + \text{derz\_d}(a_{yz}^d)]$$

$$\text{tmp}_z^d = \frac{1}{2} [\text{derx\_d}(a_{zx}^d) + \text{dery\_d}(a_{zy}^d) + \text{derz\_d}(a_{zz}^d)]$$

$$\text{lesadv}_x^d = \text{tmp}_x^d \text{derx}(u_x) + \text{tmp}_x \text{derx\_d}(u_x^d) + \text{tmp}_y^d \text{dery}(u_x) + \text{tmp}_y \text{dery\_d}(u_x^d) + \text{tmp}_z^d \text{derz}(u_x) + \text{tmp}_z \text{derz\_d}(u_x^d)$$

$$\text{lesadv}_y^d = \text{tmp}_x^d \text{derx}(u_y) + \text{tmp}_x \text{derx\_d}(u_y^d) + \text{tmp}_y^d \text{dery}(u_y) + \text{tmp}_y \text{dery\_d}(u_y^d) + \text{tmp}_z^d \text{derz}(u_y) + \text{tmp}_z \text{derz\_d}(u_y^d)$$

$$\text{lesadv}_z^d = \text{tmp}_x^d \text{derx}(u_z) + \text{tmp}_x \text{derx\_d}(u_z^d) + \text{tmp}_y^d \text{dery}(u_z) + \text{tmp}_y \text{dery\_d}(u_z^d) + \text{tmp}_z^d \text{derz}(u_z) + \text{tmp}_z \text{derz\_d}(u_z^d)$$


---

---

**Algorithm 11** divergence –  $\text{les\_d}(\mathbf{a}, \mathbf{a}^d, \text{lesdiv}^d)$ 


---

$$\text{lesdiv}_i^d = \text{derx\_d}(a_{ix}^d) + \text{dery\_d}(a_{iy}^d) + \text{derz\_d}(a_{iz}^d)$$


---

---

**Algorithm 12**  $\text{LES\_d}(\mathbf{u}, \mathbf{u}^d, C_{\text{sp}}, \text{lesdiss}^d, \text{lesadv}^d, \text{lesdiv}^d)$ 


---

```

variance – stsp_d ( $\mathbf{u}, \mathbf{u}^d, \mathbf{a}, \mathbf{a}^d, C_{\text{sp}}$ )
 $\partial_x \mathbf{u} = \text{derx}(\mathbf{u})$ 
 $\partial_y \mathbf{u} = \text{dery}(\mathbf{u})$ 
 $\partial_z \mathbf{u} = \text{derz}(\mathbf{u})$ 
 $\partial_x \mathbf{u}^d = \text{derx\_d}(\mathbf{u}^d)$ 
 $\partial_y \mathbf{u}^d = \text{dery\_d}(\mathbf{u}^d)$ 
 $\partial_z \mathbf{u}^d = \text{derz\_d}(\mathbf{u}^d)$ 
dissipation – les_d ( $\mathbf{u}, \mathbf{u}^d, \nabla \mathbf{u}, \nabla \mathbf{u}^d, \mathbf{a}, \mathbf{a}^d, \text{lesdiss}^d$ )
advection – les_d ( $\mathbf{u}, \mathbf{u}^d, \mathbf{a}, \mathbf{a}^d, \text{lesadv}^d$ )
divergence – les_d ( $\mathbf{a}, \mathbf{a}^d, \text{lesdiv}^d$ )

```

---

### Backward Adjoint Mode

---

**Algorithm 13**  $\text{variance – stsp\_b}(\mathbf{u}, \mathbf{u}^b, \mathbf{a}, \mathbf{a}^b, C_{\text{sp}})$ 


---

```

 $\mathbf{u}_m = \text{avg}_\Omega(\mathbf{u})$ 
 $\mathbf{a}^b = (\mathbf{a}^b \cdot C_{\text{sp}}) / N_\Omega$ 
for  $R \in \Omega$  do
   $\mathbf{u}_R^b = \mathbf{u}_R^b + (u_{j,R} - u_{m,j})a_{j,j}^b + (\mathbf{u}_R - \mathbf{u}_m)\text{diag}(\mathbf{a}^b)$ 
   $\mathbf{u}_m^b = \mathbf{u}_m^b - (u_{j,R} - u_{m,j})a_{j,j}^b - (\mathbf{u}_R - \mathbf{u}_m)\text{diag}(\mathbf{a}^b)$ 
end for
 $\mathbf{u}_m^b = \mathbf{u}_m^b / N_\Omega$ 
for  $R \in \Omega$  do
   $\mathbf{u}_R^b = \mathbf{u}_R^b + \mathbf{u}_m^b$ 
end for

```

---



---

**Algorithm 14**  $\text{dissipation – les\_b}(\mathbf{u}, \mathbf{u}^b, \nabla \mathbf{u}, (\nabla \mathbf{u})^b, \mathbf{a}, \mathbf{a}^b, \text{lesdiss}^b)$ 


---

```

 $(\partial_j u_i)^b = (\partial_j u_i)^b + a_{kj} \cdot \text{der}[\mathbf{k}]_{\mathbf{b}}(\text{lesdiss}_i^b)$ 
 $a_{jk}^b = a_{jk}^b + \partial_k u_i \cdot \text{der}[\mathbf{j}]_{\mathbf{b}}(\text{lesdiss}_i^b) + (1 - \delta_{jk}) \partial_j u_i \cdot \text{der}[\mathbf{k}]_{\mathbf{b}}(\text{lesdiss}_i^b)$ 

```

---



---

**Algorithm 15**  $\text{advection – les\_b}(\mathbf{u}, \mathbf{a}, \mathbf{a}^b, \text{lesadv}^b)$ 


---

```

advection – les ( $\mathbf{u}, \mathbf{a}, \text{lesadv}$ )
 $\mathbf{u}^b = \mathbf{u}^b + \text{derx\_b}(\text{tmp}_x \cdot \text{lesadv}) + \text{dery\_b}(\text{tmp}_y \cdot \text{lesadv}) + \text{derz\_b}(\text{tmp}_z \cdot \text{lesadv})$ 
 $\text{tmp}_x^b = \frac{1}{2}[\text{derx}(u_x)\text{lesadv}_x^b + \text{derx}(u_y)\text{lesadv}_y^b + \text{derx}(u_z)\text{lesadv}_z^b]$ 
 $\text{tmp}_y^b = \frac{1}{2}[\text{dery}(u_x)\text{lesadv}_x^b + \text{dery}(u_y)\text{lesadv}_y^b + \text{dery}(u_z)\text{lesadv}_z^b]$ 
 $\text{tmp}_z^b = \frac{1}{2}[\text{derz}(u_x)\text{lesadv}_x^b + \text{derz}(u_y)\text{lesadv}_y^b + \text{derz}(u_z)\text{lesadv}_z^b]$ 
 $a_{ij} = a_{ij} + \text{der}[\mathbf{i}]_{\mathbf{b}}(\text{tmp}_j^b) + (1 - \delta_{ij})\text{der}[\mathbf{j}]_{\mathbf{b}}(\text{tmp}_i^b)$ 

```

---

---

**Algorithm 16** divergence – les\_b ( $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesdiv^b$ )

---

$$a_{ij}^b = a_{ij}^b + \text{der}[\mathbf{i}]_b(\frac{1}{2}lesdiv_j^b) + (1 - \delta_{ij})\text{der}[\mathbf{j}]_b(\frac{1}{2}lesdiv_i^b)$$


---

---

**Algorithm 17** LES\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $C_{sp}$ ,  $lesdiss^b$ ,  $lesadv^b$ ,  $lesdiv^b$ )

---

divergence – les\_b ( $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesdiv^b$ )

advection – les\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesadv^b$ )

dissipation – les\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $\nabla\mathbf{u}$ ,  $(\nabla\mathbf{u})^b$ ,  $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesdiss^b$ )

$$\mathbf{u}^b = \mathbf{u}^b + \text{derx}_b((\partial_x\mathbf{u})^b) + \text{dery}_b((\partial_y\mathbf{u})^b) + \text{derz}_b((\partial_z\mathbf{u})^b)$$

variance – stsp\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $C_{sp}$ )

---

## Backward Adjoint Mode - Coefficient Control

---

**Algorithm 18** LES\_Coeff\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $C_{sp}$ ,  $C_{sp}^b$ ,  $lesdiss^b$ ,  $lesadv^b$ ,  $lesdiv^b$ )

---

divergence – les\_b ( $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesdiv^b$ )

advection – les\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesadv^b$ )

dissipation – les\_b ( $\mathbf{u}$ ,  $\mathbf{u}^b$ ,  $\nabla\mathbf{u}$ ,  $(\nabla\mathbf{u})^b$ ,  $\mathbf{a}$ ,  $\mathbf{a}^b$ ,  $lesdiss^b$ )

$$C_{sp}^b = a_{ij} \cdot a_{ij}^b$$


---

## 7.2 Validation - synthetic case study with 3D observations

In this section, the performance of the 4D-Var<sub>les</sub> code is analysed using synthetic 3D observations for transitional wake flow around a circular cylinder at  $Re = 3900$ . A biased background estimate is provided in order to assess the ability of the algorithm to revert to the observations. The details of the flow and its configuration are enumerated in §7.2.1. The synthetic observations and its properties are described in §7.2.2. The creation of the biased background trajectory is explained in §7.2.3 and §7.2.4 concludes this section with results and discussion. The optimisation of the LES coefficient is presented separately in §7.5. In this section, the coefficient is kept fixed.

### 7.2.1 Flow configuration

The base flow configuration is equivalent to the wake flow considered in §3.3. The reference domain measures  $20D \times 20D \times \pi D$  and is discretised into  $241 \times 241 \times 48$  voxels with stretching implemented in the lateral  $y$  direction. For this transitional  $Re$ , DNS would be computationally heavy and hence, the StSp LES model is used to simulate the flow. The model coefficient for this model will be optimised in further sections. The same parameters are used for the simulation as in §3.3.

The cylinder is placed at 5D from the inlet and in the centre of the lateral domain at 10D. The immersed boundary method of Gautier et al. (2014) is used to account for the solid body by introducing an additional force in the momentum equation. For all simulations, inflow/outflow boundary condition is implemented along the streamwise ( $x$ ) direction with free-slip and periodic boundary conditions along the lateral ( $y$ ) and spanwise

(z) directions respectively. The turbulence is initiated in the flow by introducing a white noise in the initial condition. Time advancement is performed using the third order Adams-Bashforth scheme. For the LES model, a spatial neighbourhood spanning  $7 \times 7 \times 7$  points is used for the averaging and variance calculations. The coefficient for the model, in the reference simulation, is defined based on the Kolmogorov length and time scales (see eq. (2.2.47)).

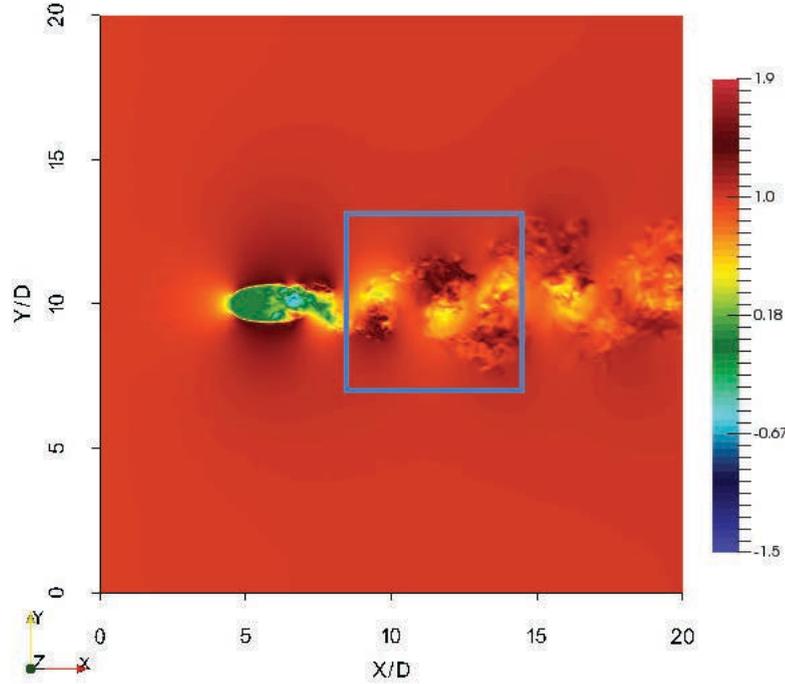


Figure 7.9: Streamwise velocity snapshot in the reference domain. The sub-domain in blue corresponds to the assimilation domain for which the data is extracted.

To attempt a reconstruction over the whole reference domain using data-assimilation is computationally prohibitive. It is also restricted by the presence of solid bodies as well as complexities associated with stretched meshes which are beyond the scope of this thesis. Thus, a smaller domain in the near wake of the cylinder starting at  $3D$  behind the cylinder was chosen for the assimilation procedure. This assimilation sub-domain measures  $6D \times 6D \times \pi D$  and is discretised into  $165 \times 165 \times 48$  (see figure 7.9) - a smaller mesh is essential here in the assimilation run for stability as the mesh is not stretched. All reference fields were preprocessed using a verified interpolation procedure to fit the assimilation mesh. The assimilation is carried out using the StSp model over the same spatial domain with an estimate of the initial coefficient from the reference simulation - an attempt is made at optimising this coefficient using the  $4D\text{-Var}_{les}$  algorithm in §7.5. AB3 scheme is used for time-marching. It is important to note that only the velocity fields are obtained from the reference. Other quantities required for the restart procedure of Incompact3d (pressure field, convection and diffusion terms of previous time-steps, etc) are assumed to be zero. Thus, we begin with an Euler scheme for the first time step of the assimilation before switching to the AB3 scheme.

### 7.2.2 Observations construction

The time-line for the procedure is shown graphically in figure 7.10. The reference simulation is initialised with a constant inlet profile with 10% white-noise seeded in the whole domain. This simulation is allowed to run until convergence - convergence is assumed to be reached when the effect of the initial condition is removed from the simulation. This convergence is ensured by letting the simulation run for an 90 non-dimension time - generally 45 non-dimension time are sufficient for convergence. The observations are then collected every  $10 \Delta t'$  from  $30000 \Delta t'$  to  $30100 \Delta t'$ . This is the period in which data assimilation is carried out corresponding to 1/16th of a vortex shedding. Longer time spans for assimilation can lead to instabilities in the procedure and thus needs to be handled with care - one option for performing long term assimilation is to use time-marching assimilation windows but these are beyond the scope of this thesis.

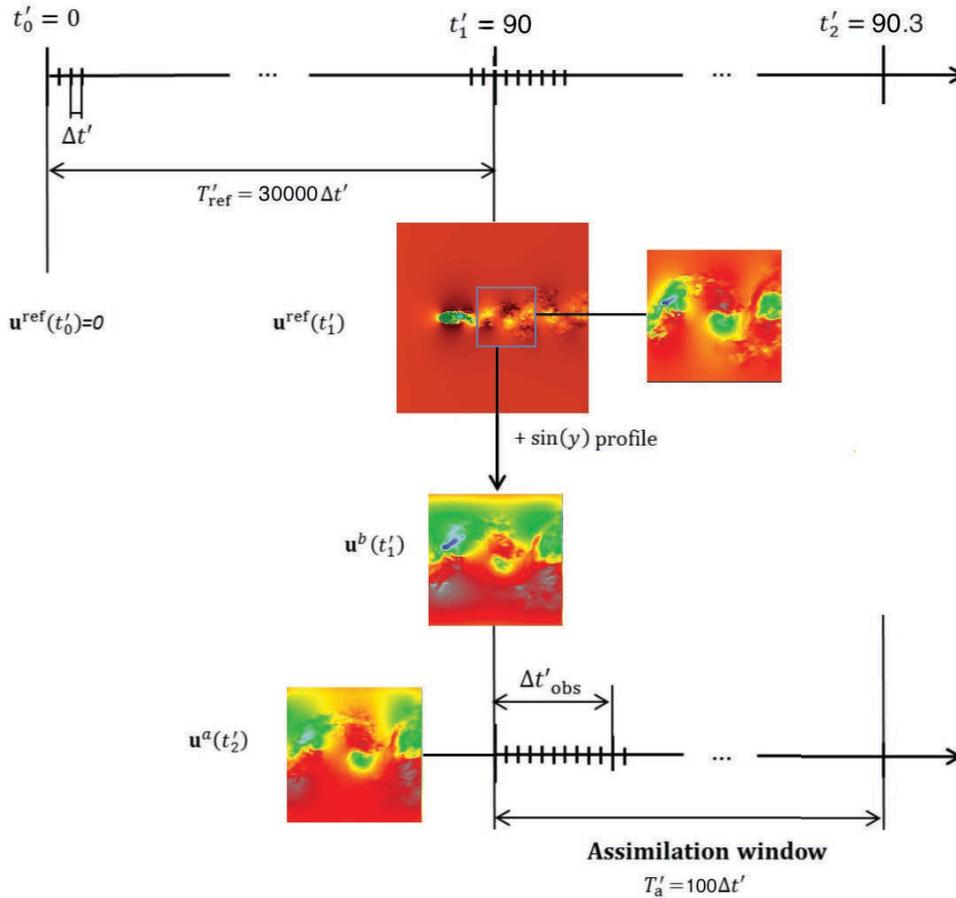


Figure 7.10: Time-line for reference simulation + data assimilation procedure.

The observation at  $30000 \Delta t'$ , which is the initial condition for the assimilation procedure, is intentionally biased to provide a wrong background trajectory. Due to mesh mis-match between the reference and the assimilation, the stretched reference velocity field needs to be interpolated onto a non-stretched mesh. In addition, due to the stream-wise discrepancy in the mesh size, a 2D linear interpolation scheme is implemented in the  $x, y$  directions. The 3D observations, preprocessed, are assimilated every  $10 \Delta t'$ . Table 7.3 concisely depicts configuration parameters. The instantaneous reference velocity field at the beginning of the assimilation window on the streamwise plane is shown in figure 7.11.

Table 7.3: Configuration parameters for the reference and assimilation runs.

|             | Re   | $n_x \times n_y \times n_z$    | $l_x/D \times l_y/D \times l_z/D$ | $U\Delta t/D$ | Duration        |
|-------------|------|--------------------------------|-----------------------------------|---------------|-----------------|
| Full Domain | 3900 | $241 \times 241^s \times 48$   | $20 \times 20 \times 3.14$        | 0.003         | $30100\Delta t$ |
| Observation | 3900 | $165^i \times 165^i \times 48$ | $6 \times 6 \times 3.14$          | 0.003         | $100\Delta t$   |
| 4DVAR       | 3900 | $165 \times 165 \times 48$     | $6 \times 6 \times 3.14$          | 0.003         | $100\Delta t$   |

s - Stretched; i - Interpolated

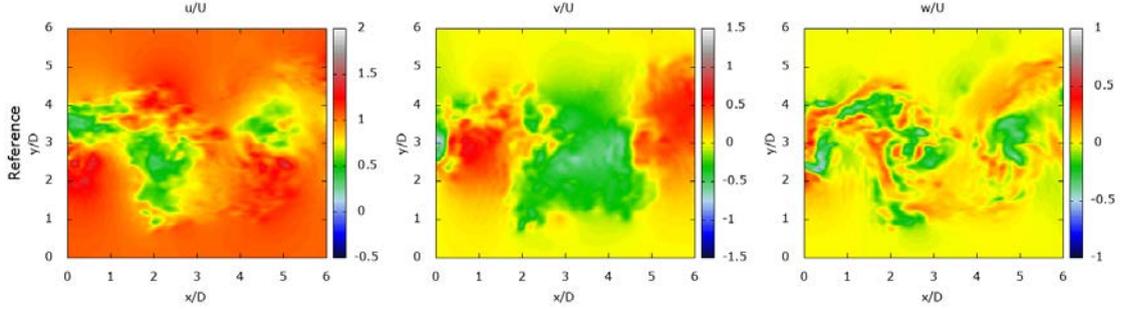


Figure 7.11: Instantaneous observation velocity snapshots at the beginning of the assimilation window ( $t'_1$ ) in the streamwise plane  $z = 1.57D$ . From left to right: Streamwise, lateral, and spanwise velocity field.

The observation covariance matrix  $\mathbf{R}^{-1}$  is implemented to be a function of space in order to account for boundary condition effects. A hyperbolic profile is used to reduce confidence in regions near the boundaries smoothly up to a user-defined minimum value ( $\mathbf{R}_{min}^{-1}$ ). A user-defined maximum ( $\mathbf{R}_{max}^{-1}$ ) value is enforced in the middle of the domain. In the streamwise direction, due to the availability of measurements on the inlet plane, confidence is reduced only near the outlet boundary.

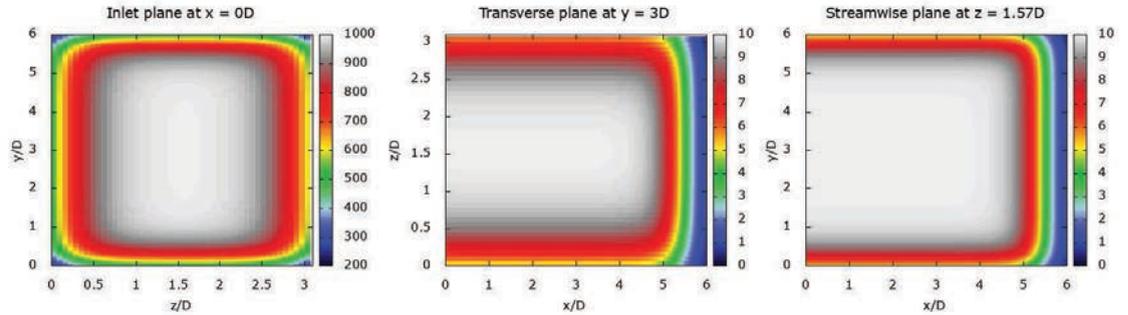


Figure 7.12: Inverted observation covariance matrix  $\mathbf{R}^{-1}$  on planes perpendicular to the three principle directions. From left to right: inlet plane at  $x = 0D$ , transverse plane at  $y = 3D$ , and streamwise plane at  $z = 1.57D$ .

The resultant inverse of the observation covariance matrix  $\mathbf{R}^{-1}$  is shown in figure 7.12 in planes perpendicular to all three principle directions. The value of  $\mathbf{R}^{-1}$  on the inlet plane is higher due to higher confidence on the inlet condition as well as to account for the mismatch between 2D observations of the inlet plane as compared to the 3D reference - this mismatch affects the cost functional which is summed over the whole spatial domain. The observation covariance is scaled to calculate the inlet plane covariance. A temporal

interpolation with gaussian weights is also encoded to calculate inlet covariances for cases with temporally varying observation covariance. This is important as the inlet condition is optimised at every assimilation time-step unlike the observations which are assimilated with a longer time interval between each observation ( $dt_{obs} > dt_{assim}$ ). The algorithm for defining the observation covariance matrix is mathematically expressed as,

$$\mathbf{R}^{-1}(x, y, z) = \frac{1}{2}(\tanh(-2.0(x - 0.9L_x)) + 1) \quad (7.2.1)$$

$$\begin{aligned} \mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \frac{1}{2}(\tanh(-2.0(y - \frac{L_y}{2} - 0.5L_y)) \\ - \tanh(-2.0(y - \frac{L_y}{2} + 0.5L_y))) \end{aligned} \quad (7.2.2)$$

$$\begin{aligned} \mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \frac{1}{2}(\tanh(-2.0(z - \frac{L_z}{2} - 0.5L_z)) \\ - \tanh(-2.0(z - \frac{L_z}{2} + 0.5L_z))) \end{aligned} \quad (7.2.3)$$

$$\mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * (R_{max}^{-1} - R_{min}^{-1}) + R_{min}^{-1} \quad (7.2.4)$$

### 7.2.3 Initial background construction

Ideally, the background trajectory is chosen such that it is an optimal guess of the true trajectory from which the assimilation can optimise further. This reduces the time taken to obtain the analysis trajectory by providing an ideal starting point. However, the aim here is to identify the capabilities of the assimilation technique to capture the true state given an incorrect initial state. This is achieved by biasing the true velocity field at start of the assimilation window with a sinusoidal profile to produce an ‘incorrect’ background trajectory - this trajectory is henceforth also referred to as ‘case (a)’. The ability of the SO reconstruction method, developed in chapter 6, to produce a background trajectory is also analysed in this section: two background trajectories are developed by using the SO method (case (b)) and the averaging SO method (case (c)). The details regarding the reconstruction data-set is provided in §7.3.1 below. All three initial instantaneous background velocity fields on a streamwise plane at  $z = 1.31D$  are shown in figure 7.13. The diagonal inverse background error covariance matrix  $\mathbf{B}^{-1}$  is set to 1 for all cases in this and the next section. A novel method for estimating this covariance matrix is formulated in §7.4 and its effects on the assimilation algorithm is analysed.

### 7.2.4 Results and Discussion

Given the observations and a background initial guess, the assimilation can proceed towards identifying an optimal assimilated trajectory which will henceforth be referred to as the analysis trajectory. The background trajectory is also evolved from the given background initial condition. All simulations were carried out on a single assimilation window with only one external loop iteration due to cost limitations. There is no restriction on the inner loop iterations - LBFGS algorithm converges when no change is observed on the cost function over 10 inner loop iterations. The optimisation characteristics, vis-à-vis cost function and its gradient, are studied in the next section followed by a contour map comparison of the analysis trajectory for all three background trajectories.

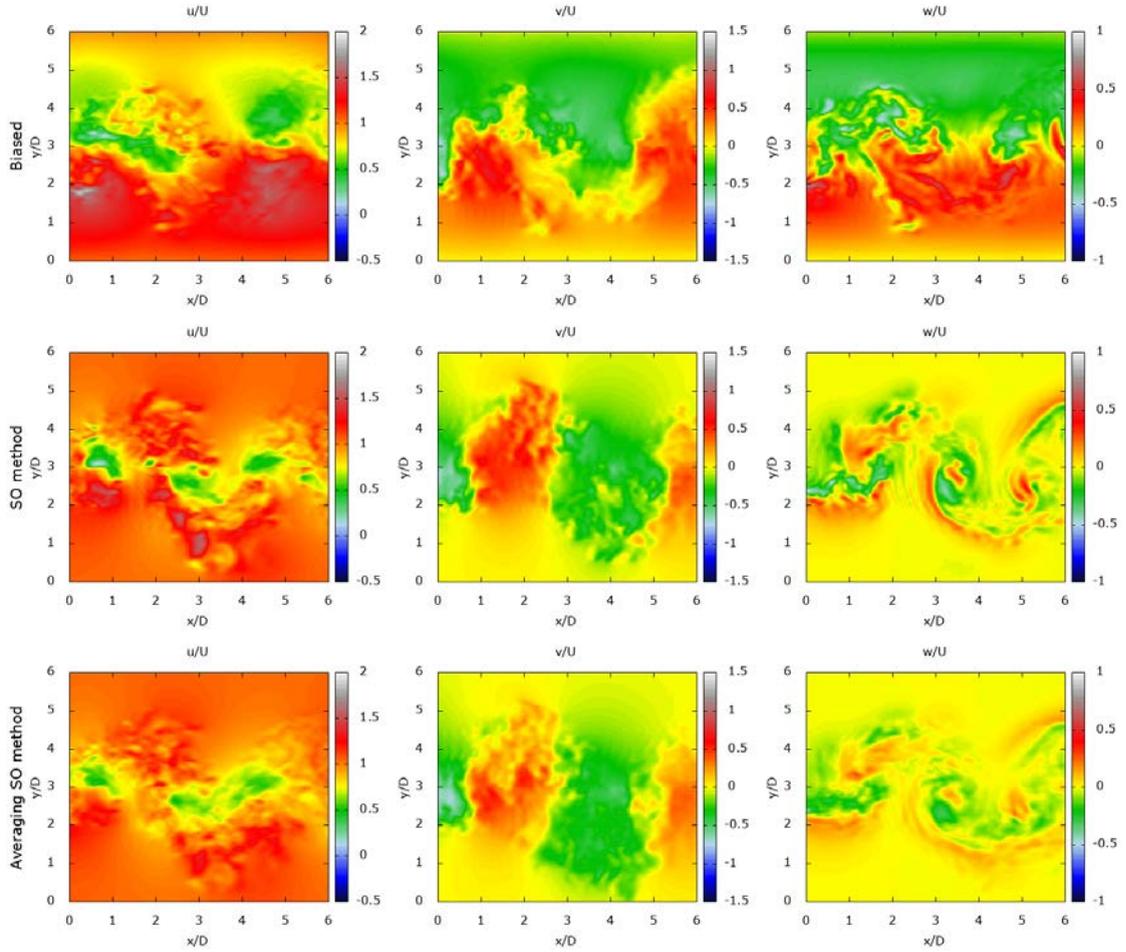


Figure 7.13: Instantaneous background velocity snapshots at the beginning of the assimilation window ( $t'_1$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: Biased (case (a)), SO method (case (b)), and averaging SO method (case (c)) based background.

### Optimisation characteristics

A first measure of the algorithm is the evolution of the cost function  $J(\mathcal{X}_0)$  and norm of its gradient  $\|\nabla J(\mathcal{X}_0)\|$  as a function of the inner loop iterations. The cost function is plotted for each background condition in figure 7.14. A better background estimate reduces significantly the number of inner loop iterations required to optimise the analysis trajectory. The biased background (the worst estimate of the three background conditions) requires maximum iterations to converge with a gentle slope of convergence. The SO method requires the least number of iterations, however, it is interesting to note that the initial cost function for this case is higher than case (a). This could arise either due to a noisy reconstructed field or due to a lack of temporal correlation between the reference and the background - the SO method optimises the match between the two fields only at the inlet and not over the whole 3D domain (for details refer to §6.2). The biased background, on the other hand, is inherently the reference field itself with an overlay of a sinusoidal bias and thus has a lower initial cost function. The averaging SO method which provides a smoother reconstruction appears to provide a better initial estimate and converges with similar number of iterations to the SO method.

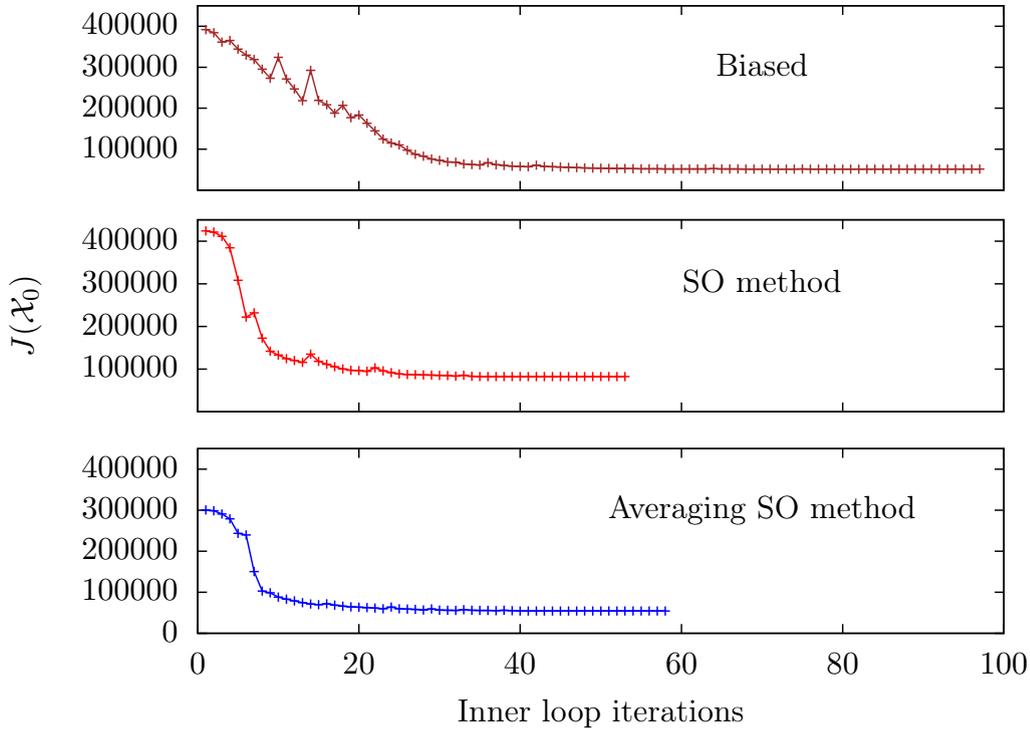


Figure 7.14: Evolution of the cost function  $J(\mathcal{X}_0)$  for the assimilation as a function of inner loop iterations for all three background trajectories.

The gradient of the cost function for each case is plotted in figure 7.15. The gradient estimate for case (b) and case (c) is smoother with minimal peaks as compared to case (a). The points along the strong initial gradient peak can be correlated with the drastic decrease in the cost function for the same iterations. The multiple strong peaks on the gradient function for case (a) and (b) correspond to the miniature peaks on the cost function for the two cases. Despite this deviation, the method shows significant robustness and manages to converge to the optimal for all cases. Case (c) is the only instance which has a smooth functional - the averaging SO method addresses the shortcomings of both case (a) and (b). Namely, it provides a more physically relevant background than case (a) and a smoother option than case (b). Considering the noisy curve of the gradient for the biased case, it suggests a difficulty in finding an optimal analysis trajectory when the background is far away from the reference.

### Optimised analysis trajectory

The analysis velocity maps are analysed on a streamwise plane at  $z = 1.31D$ . Such a translation is necessary as the SO based reconstruction of the background field is identical to the reference field in the mid-plane at  $z = 1.57D$  - the analysis trajectory would thus be minimally modified by the assimilation procedure. Figure 7.16 shows the velocity maps of the 4D-Var<sub>les</sub> optimised initial condition for each case along with the true reference field. A comparison with figure 7.13 clearly depicts an improvement from the background condition provided to the algorithm. The improvement is limited for the biased case while a significant improvement can be observed in case (b) and case (c). This difference can be especially seen for the lateral velocity map ( $v/U$ ). Maximum improvement is seen for  $w/U$  as compared to the other fields. Increased small scale structures can be seen in

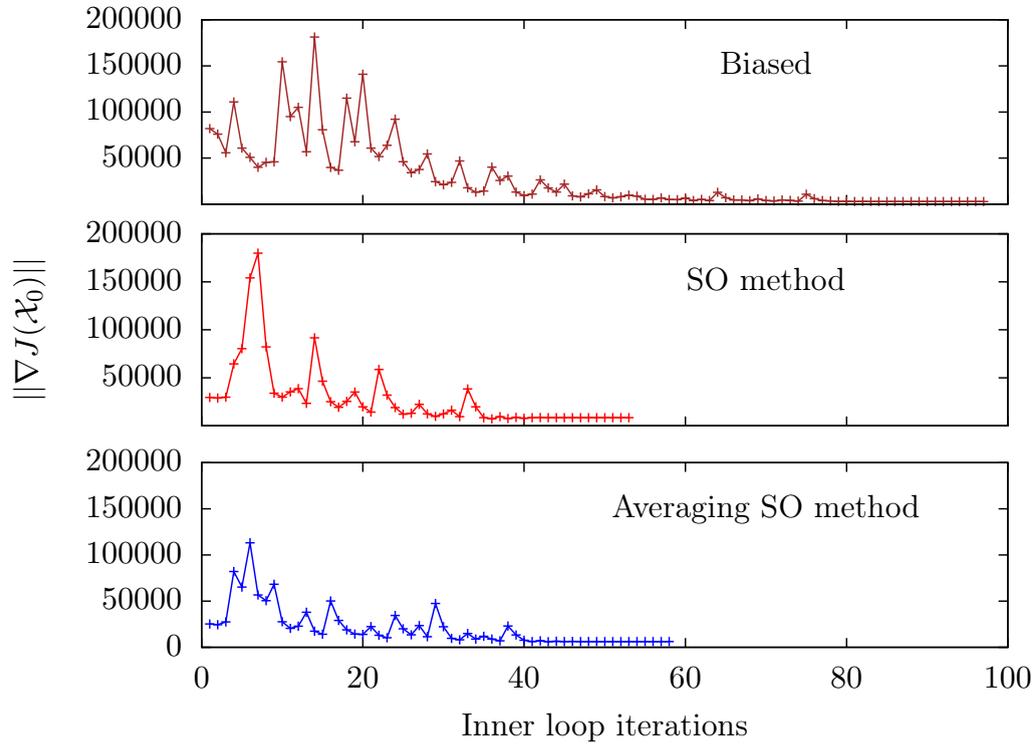


Figure 7.15: Evolution of the gradient of the cost function  $\|\nabla J(\chi_0)\|$  for the assimilation as a function of inner loop iterations for all three background trajectories.

case (b) probably due to the reconstruction methodology which introduces noise into the background - this is reduced by case (c) whose averaging procedure smooths the background velocity fields. Case (c) provides the best initial estimate of the analysis out of the three cases considered.

Considering the nature of the adjoint gradient calculation, the error between the current analysis trajectory and the observations at final observation time (first observation for the adjoint simulation) would contribute the most to the gradient calculation. Thus, it is natural to expect a more significant improvement at the final observation time as compared to the initial observation which has the least contribution (such an argument is valid under the assumption that the error between the observation and the current trajectory is similar at each assimilation of the observation). Figure 7.17 shows the velocity maps along the same streamwise plane at the final observation time  $t'_2$ . A significant improvement on the background is indeed observed for all three cases with the analysis contours comparable to the observations. Case (a) shows maximum improvement from  $t'_1$  to  $t'_2$  with the final trajectory being slightly better than case (b) which appears to advect the noise - this is inline with the findings of Robinson (2015) who observed that the assimilation algorithm tends to advect a noisy background with minimal de-noising. Case (c) could arguably be considered the best analysis trajectory - small scale structures from the reference seen in case (c) are missing in the other two cases.

The results are indicative towards the applicability and accuracy of the 4D-Var<sub>les</sub> algorithm with respect to 3D true observations. However, the main application of DA lies in utilising sparse observation data to guide a full-scale (spatial) simulation and this is analysed in the next section for the 4D-Var<sub>les</sub> algorithm.

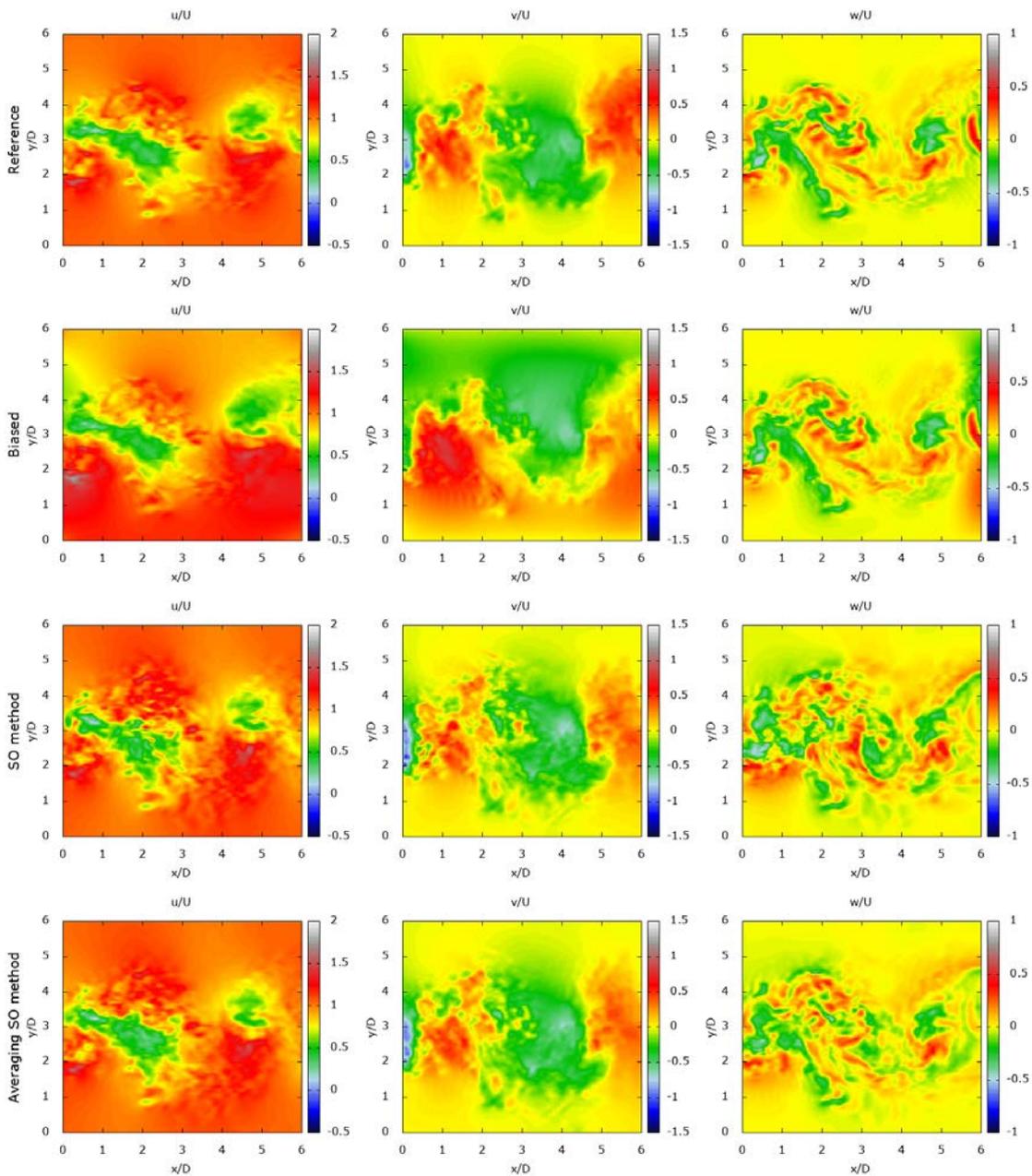


Figure 7.16: Instantaneous analysis and reference velocity snapshots at the beginning of the assimilation window ( $t_1$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: Reference, biased (case (a)), SO method (case (b)), and averaging SO method (case (c)) based analysis.

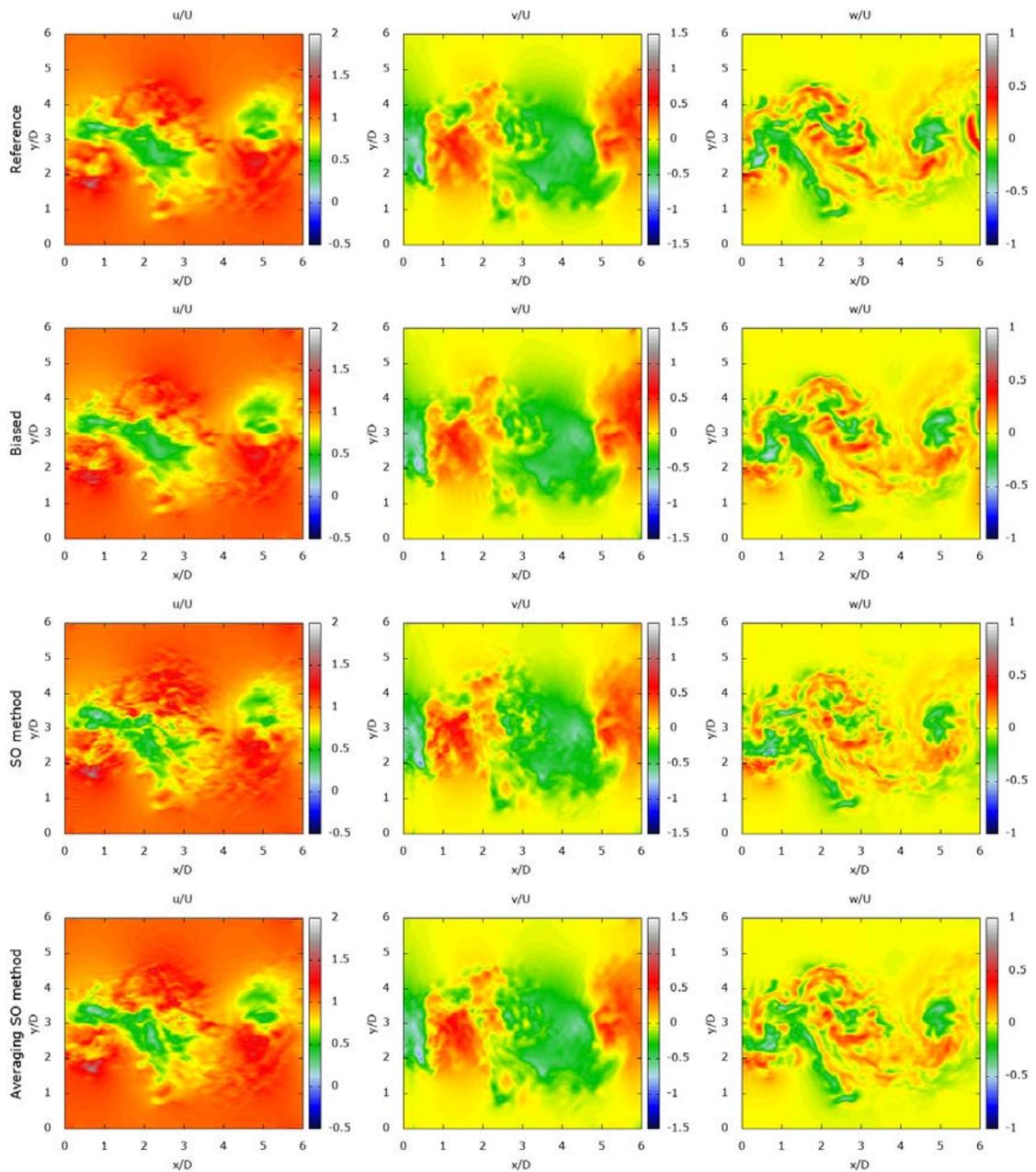


Figure 7.17: Instantaneous analysis and reference velocity snapshots at the final time of the assimilation window ( $t'_2$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: Reference, biased (case (a)), SO method (case (b)), and averaging SO method (case (c)) based analysis.

Table 7.4: Parameters for the different assimilation cases.

| Case | Observations | Background        | Background covariance | Coefficient optimisation |
|------|--------------|-------------------|-----------------------|--------------------------|
| (a)  | true 3D      | 3D biased         | Identity              | No                       |
| (b)  | true 3D      | SO                | Identity              | No                       |
| (c)  | true 3D      | Averaging SO      | Identity              | No                       |
| (d)  | Interpolated | SO                | Identity              | No                       |
| (e)  | Averaging SO | SO                | Identity              | No                       |
| (f)  | true 3D      | SO                | fully-defined         | No                       |
| (g)  | true 3D      | case (a) analysis | Identity              | Yes                      |
| (h)  | Averaging SO | SO                | fully-defined         | Yes                      |

### 7.3 Application - synthetic data emulating 2D experimental observations

This section focuses on assessing the algorithm capabilities with respect to partial observations as generally obtained from experimental methods. Due to the unavailability of validated cross-plane PIV observations for wake flow in the transitional Re range, an attempt has been made to modify synthetic data to emulate experimental data. This is achieved by extracting only cross-planar information from the synthetic simulation and building 3D observations from this limited data-set. The flow configuration for the synthetic data set is explained in the next section (§7.3.1). Procedure for the construction of 3D observations and background are provided in §7.3.2 and §7.3.3. This section concludes with the algorithm results and associated discussion in §7.3.4. The various cases studied in this chapter are all consolidated in table 7.4.

#### 7.3.1 Flow configuration

The same flow configuration to §7.2.1 is used for all simulations in this section. The same assimilation domain of size  $6D \times 6D \times 3.14D$  is extracted from the full-domain data with an identically discretised mesh. The significant alteration between the two time-line is the duration of the full-domain simulation which is run for 120000 time-steps after convergence with additional velocity snapshots recorded every 50 time-steps providing a comprehensive data-set of 2400 snapshots covering 72 vortex-sheddings. The corresponding time-line is shown in figure 7.18. This additional data collection procedure was also used in the previous section to reconstruct background trajectories using the SO methods. The StSp LES model is used for all the case-studies in this section as well, with the variance tensor calculated over the same spatial neighbourhood as before.

#### 7.3.2 Observations construction

In order to emulate an experimental cross-planar PIV data-set and its limited spatial extent, only observations on two cross planes are extracted from the reference simulation. The geometry, similar to that used in the SO based reconstructions, recalled in figure 7.19, consists of one inlet plane (in red) at  $x = 3D(x_0)$  behind the cylinder and an intersecting streamwise observation plane (in green) placed in the middle of the spanwise domain at  $z = 1.57D(z_0)$ . The data on these two planes are extracted over 360 non-dimensional time.

Two types of 3D observations are then constructed from the available data. Firstly,

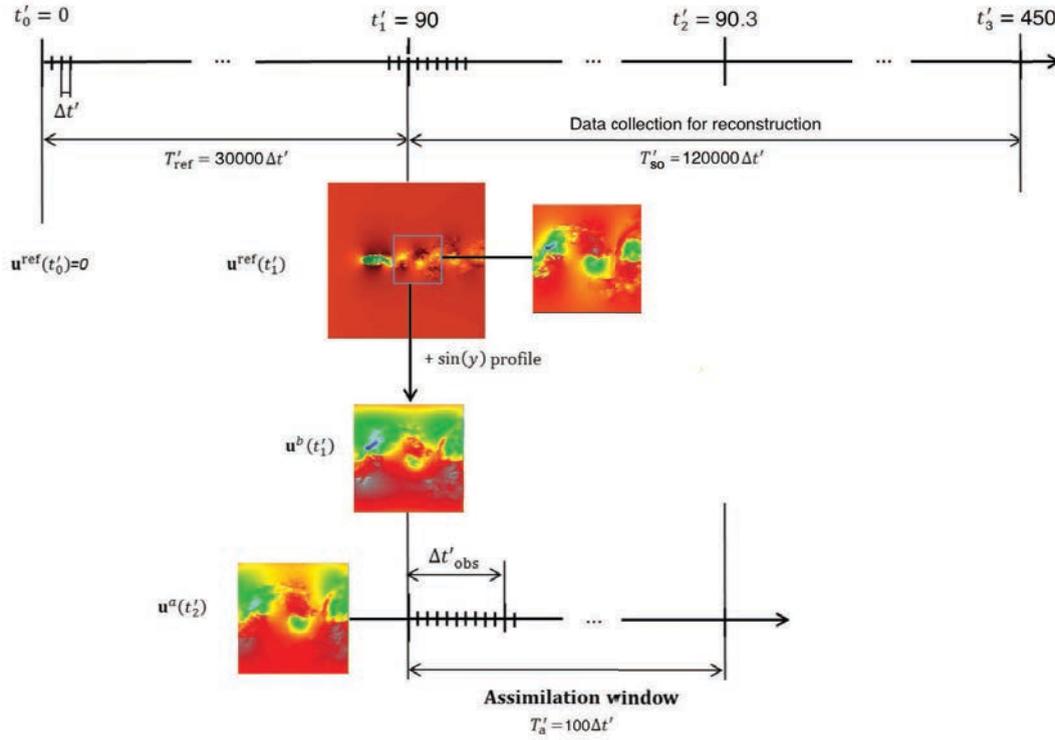


Figure 7.18: Time-line for reference simulation, data assimilation, and data collection procedures.

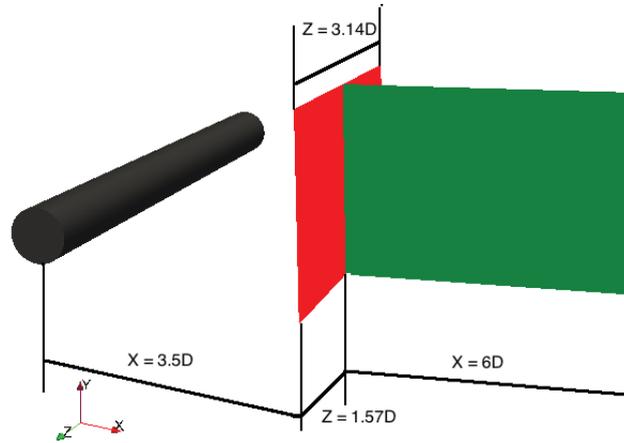


Figure 7.19: Geometry for cross-planar data collection of wake flow around a circular cylinder at  $Re\ 3900$  for reconstruction using the SO method. Red - inlet plane. Green - observation plane.

a simple interpolation method is employed to fill the empty space in the observations by weighting the data from the two planes (case (d)).

$$w_1 = \exp(-(x - x_0)^2 / (2.)) \quad (7.3.1)$$

$$w_2 = \exp(-(z - (\frac{L_z}{2}))^2 / (2.)) \quad (7.3.2)$$

$$\mathbf{u}(x, y, z) = (\mathbf{u}(x_0, y, z) * w_1 + \mathbf{u}(x, y, z_0) * w_2) / (w_1 + w_2) \quad (7.3.3)$$

Such a construction is shown in figure 7.20 on a streamwise plane off-set from the observation plane at the beginning of the assimilation window ( $t'_1$ ). The interpolation method was used by Robinson (2015) for assimilation with experimental observations. Thus, it provides a base case with which the second set of observations (case (e)) reconstructed using the SO method can be compared against. The averaging SO method is preferred over the SO method due to its ability to de-noise the reconstructed velocity field. Such a reconstruction is shown in figure 7.20 along with the true reference.

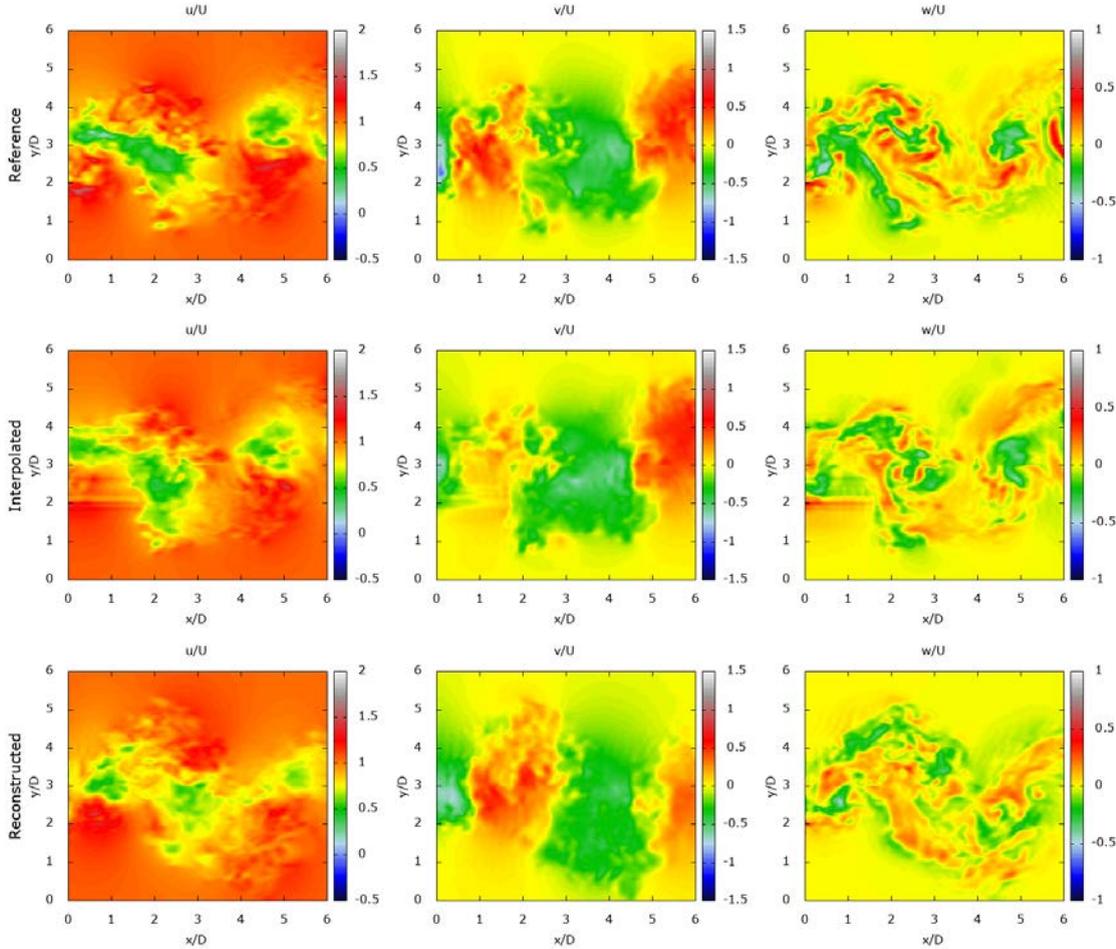


Figure 7.20: Observation velocity snapshots at the beginning of the assimilation window ( $t'_1$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: True reference, interpolated observations (case (d)), averaging SO reconstructed observations (case (e))

A more physical representation of the flow field can clearly be seen in case (e) as compared to case (d). The interpolation methodology introduces streaks near the inlet plane due to the weighting functions. The improvement brought by the reconstructed observations over the interpolated observations is better understood by comparing the velocity snapshots along a spanwise plane at  $x = 3D$  (see figure 7.21). The interpolation downstream from the inlet simply mirrors the observation plane leading to unphysical observations. This is majorly corrected in the reconstructed observations where turbulent structures are clearly visible. While the reconstruction is not a perfect fit with the reference data, major turbulent scales are captured. For an experimental case-study (PIV cross-

planar of wake flow for example), such a methodology would better represent the observed flow and the turbulent structures than a weighted interpolation algorithm.

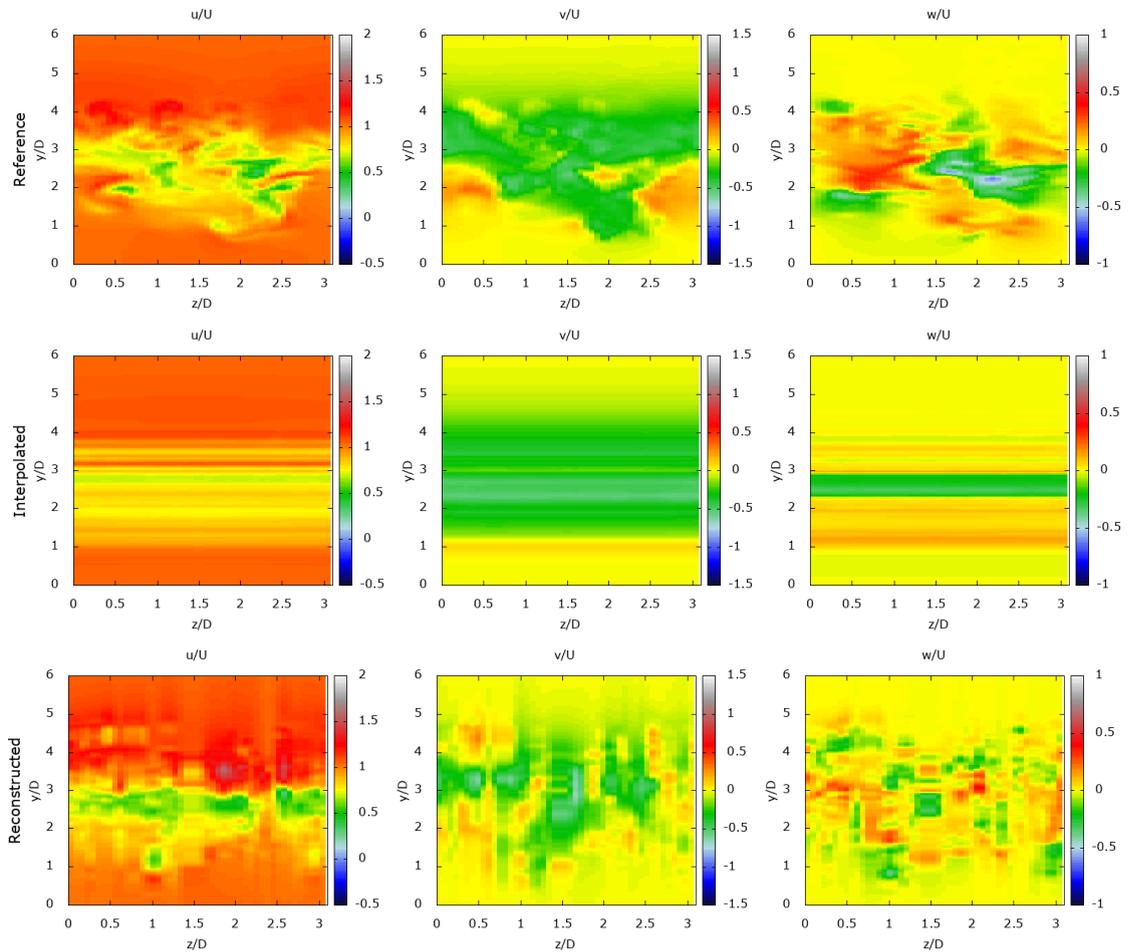


Figure 7.21: Reconstructed instantaneous observation velocity snapshots at the beginning of the assimilation window ( $t'_1$ ) in the spanwise plane at  $x = 3D$  using the averaged SO algorithm. From left to right: streamwise, lateral, and spanwise velocity field. From top to bottom: true reference, interpolated observations (case (d)), reconstructed observations (case (e)).

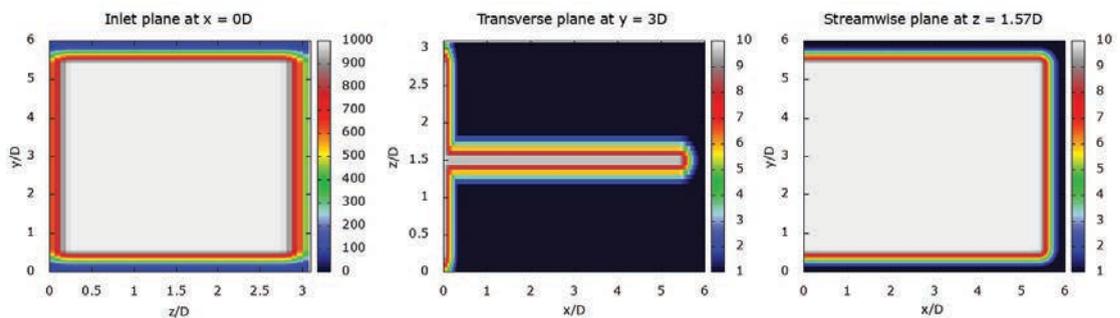


Figure 7.22: Inverted observation covariance matrix  $R^{-1}$  on planes perpendicular to the three principle directions for the reconstructed observations. From left to right: inlet plane at  $x = 0D$ , transverse plane at  $y = 3D$ , and streamwise plane at  $z = 1.57D$ .

The observations are not an exact representation of the true reference field and thus, this needs to be reflected in the inverse of the observation covariance matrix  $\mathbf{R}^{-1}$ . Only the two observed planes command high confidence while the rest have minimal confidence. Thus, maximum confidence is given to the points falling on these two planes with an exponential decrease in confidence away from the observed planes. The covariance matrix is thus constructed as,

$$\mathbf{R}^{-1}(x, y, z) = \max\left(\exp\left(-\frac{x^2}{0.05}\right), \exp\left(-\frac{(z - 0.5L_z)^2}{0.05}\right), R_{min}^{-1}\right) \quad (7.3.4)$$

$$\mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \exp\left(-\frac{(x - 0.9L_x)^2}{0.1}\right) \quad \forall x \geq 0.9L_x \quad (7.3.5)$$

$$\mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \exp\left(-\frac{(y - 0.1L_y)^2}{0.1}\right) \quad \forall y \leq 0.1L_y \quad (7.3.6)$$

$$\mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \exp\left(-\frac{(y - 0.9L_y)^2}{0.1}\right) \quad \forall y \geq 0.9L_y \quad (7.3.7)$$

$$\mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \exp\left(-\frac{(z - 0.1L_z)^2}{0.1}\right) \quad \forall z \leq 0.1L_z \quad (7.3.8)$$

$$\mathbf{R}^{-1}(x, y, z) = \mathbf{R}^{-1}(x, y, z) * \exp\left(-\frac{(z - 0.9L_z)^2}{0.1}\right) \quad \forall z \geq 0.9L_z \quad (7.3.9)$$

$$\mathbf{R}^{-1}(x, y, z) = \max(\mathbf{R}^{-1}(x, y, z), R_{min}^{-1}) * R_{max}^{-1} \quad (7.3.10)$$

On the mid streamwise plane and the inlet plane strong confidence is given taking into account boundary effect at the edges. Along the transverse plane, a steep exponential slope is enforced to reduce covariance as we move away from the two planes of true observation. The covariance on the three perpendicular planes are shown in figure 7.22.

### 7.3.3 Initial background construction

In the previous section, the optimal background estimate was obtained using the averaged SO reconstructed trajectory. However, considering that the observations are also constructed using the same algorithm, the classic SO method is preferred for constructing the background - this would assist in analysing the performance of the algorithm by providing different background and observations trajectories. The SO reconstruction is performed with the same data-set as the observation and is used as the background for all cases in this section. The background velocity map at the beginning of the assimilation window on an off-set plane at  $z = 1.31D$  is shown in figure 7.23. An alternative option for initial background construction is the ‘empty box’ technique of Robinson (2015). However, such a technique requires extensive pre-processing of the velocity fields and simulation of a DNS/LES with the true inlet imposed at every simulation time-step. This is impractical in reality - the time taken for the empty box method to fill the 3D domain would be in hours as compared to the few seconds taken by the SO algorithm. In addition, the ‘empty box’ techniques provides a solution which is rapidly (depending on the Re number) out of phase compared to the reality - Robinson (2015) observed such effect at Re 300 which would be enhanced in our higher Re 3900 case. The diagonal inverse background covariance matrix  $\mathbf{B}^{-1}$  is set to 1 for all cases in this section.

### 7.3.4 Results and Discussion

Two simulations were carried out with the SO reconstructed background and with interpolated observations (case (d)) and with averaged SO reconstructed observations (case (e)).

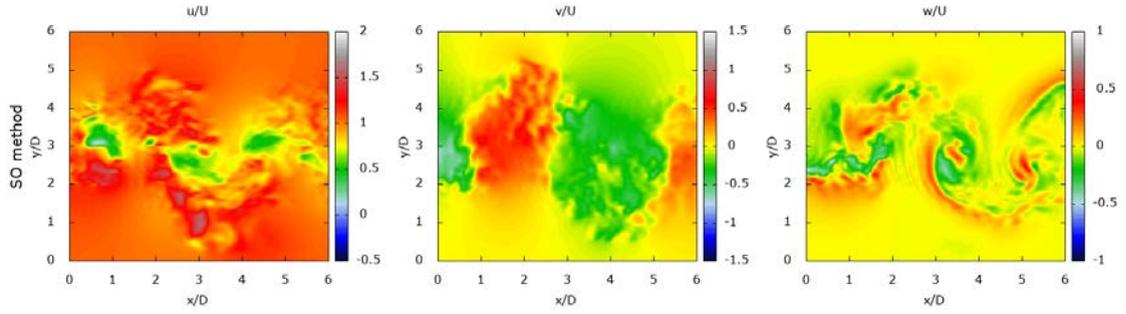


Figure 7.23: Reconstructed instantaneous background velocity snapshots at the beginning of the assimilation window ( $t_1'$ ) in the streamwise plane  $z = 1.31D$  using the SO algorithm. From left to right: Streamwise, lateral, and spanwise velocity field.

The background evolution is carried out with the estimated background initial condition. This is followed by the assimilation procedure where the initial condition, which is set identical to the background for the first iteration of the algorithm, is optimised. A single external loop iteration is carried out with the inner loop iterations determined by the LBFGS algorithm to ensure convergence. The corresponding results are now analysed.

### Optimisation characteristics

A smooth variation is observed for the cost function (see figure 7.24) for both cases which is a good indication of the algorithm capabilities. Due to the reduced confidence on the observation vis-à-vis the inverse observation covariance matrix, the cost function for both cases are drastically reduced from case (a)-(c) of the previous section. This has also reduced the difference in the cost function from its initial value to its optimised value - a  $\sim 50\%$  reduction is observed here as compared to the  $\sim 75\%$  with true 3D observations.

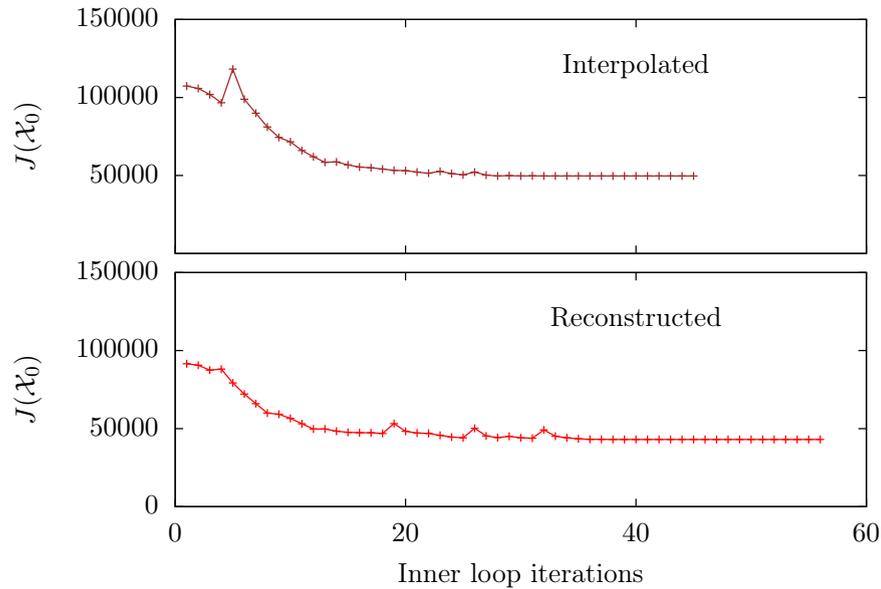


Figure 7.24: Evolution of the cost function  $J(\mathcal{X}_0)$  for the assimilation as a function of inner loop iterations with both observation data-sets.

The covariance change also plays a role on the gradient of the cost function (see figure 7.25) where any deviation from the background condition is still equally penalised but observation error is considerably reduced. A strong variance is observed in the gradient with stronger peaks observed for case (d) with interpolated observations. The lack of physical structures in case (d) could be the inducing factor for these significant oscillations as the dynamical model and the observations are not compatible.

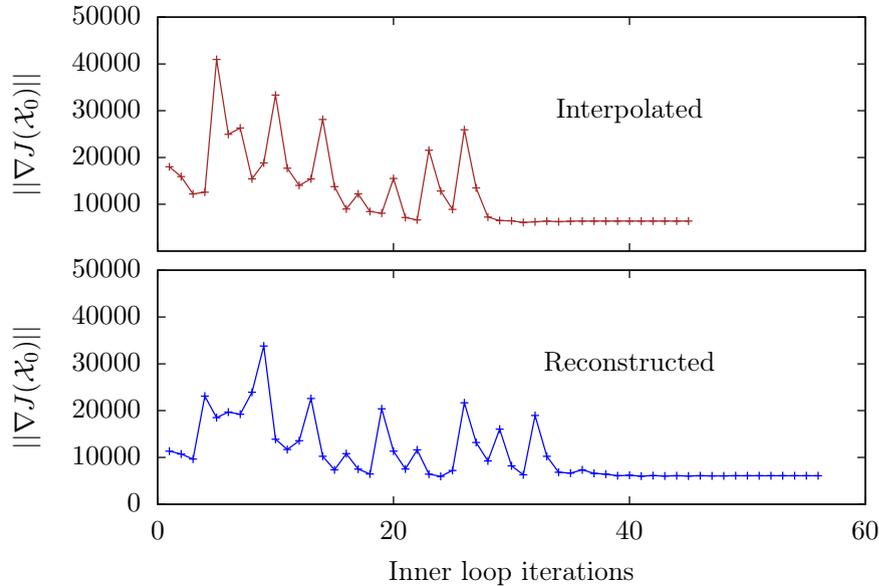


Figure 7.25: Evolution of the gradient of the cost function  $\|\nabla J(\chi_0)\|$  for the assimilation as a function of inner loop iterations with both observation data-sets.

### Optimised analysis trajectory

The velocity maps of the analysis trajectory are shown in figure 7.26 at the initial time and in figure 7.27 for the final time. The background velocity maps are displayed at the top of each figure followed by the observation and corresponding analysis velocity maps for each case. Both the optimised analyses show a tendency to stay close to the background condition - this is of prime importance in case (d) where the interpolated observations are unphysical. A more pronounced shift towards the observations can be seen in case (e) where turbulent flow structures appear to be a blend of the background and the observation. Given a case with sparse/limited observations, a compromise between the background and the observations is required as observed here.

The concern with respect to interpolated observations, apart from a lack of improvement in the initial condition, is the introduction of instabilities in the assimilation. The optimised initial condition obtained for case (d) was shown to be unstable when simulated with the non-linear NS code failing after a few time-steps. This defeats the purpose of performing DA, wherein past observations are used to identify an optimal condition with which future predictions can be made. This is not encountered with the reconstructed observations as seen in figure 7.27. The optimal analysis at the final time is once again a moderation between the background and the observation - the algorithm attempts to move towards an optima between background error and observation error.

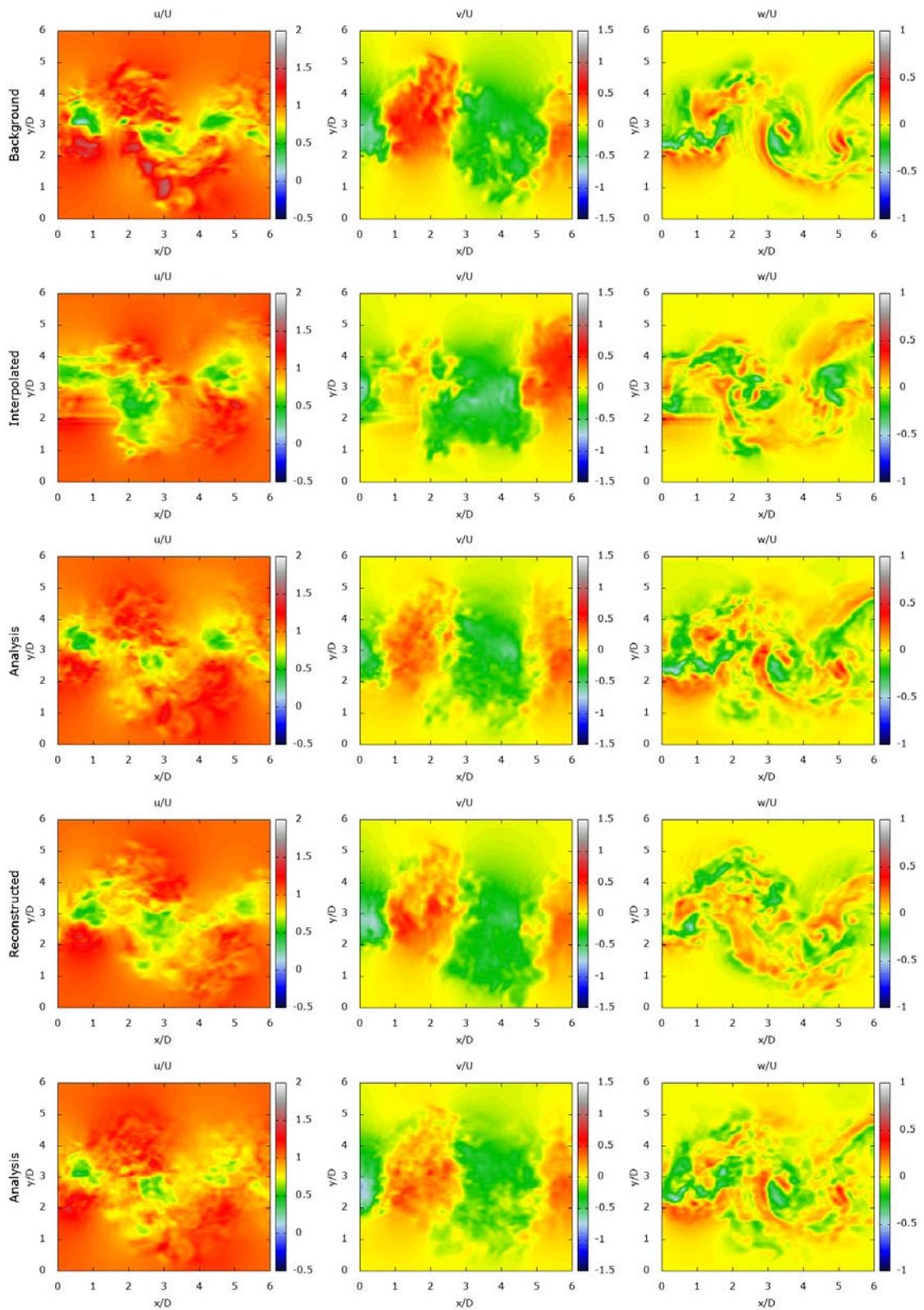


Figure 7.26: Instantaneous velocity snapshots at the initial time of the assimilation window ( $t'_1$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: Background, interpolated observation (case (d)), corresponding analysis, reconstructed observation (case (e)), and corresponding analysis.

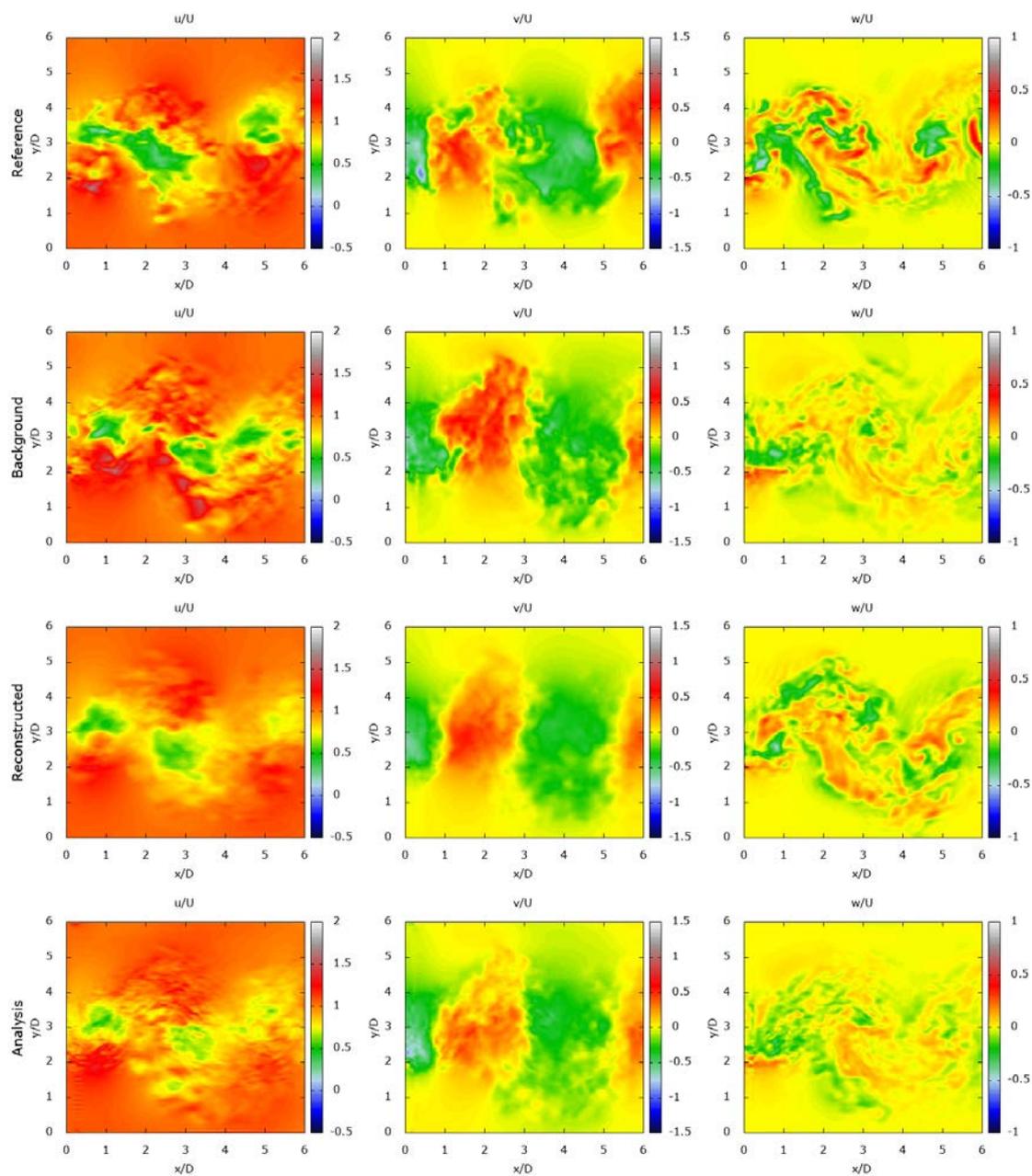


Figure 7.27: Instantaneous velocity snapshots at the final time of the assimilation window ( $t'_2$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: True reference, background, reconstructed observation (case (e)), and corresponding analysis.

The advantage of having a true 3D reference field can be exploited here by plotting the root mean square error (RMSE) between the true field and the background, reconstructed observation, and the analysis at each time step. Such an RMSE evolution for case (e) is plotted in figure 7.28 for all three velocity fields. Several conclusions can be inferred from this graph:

- The analysis trajectory provides a significant improvement over the background for all three velocity field over the entire assimilation domain;
- More strikingly, the analysis is closer to the true 3D observations as compared to the reconstructed observations for large sections of the assimilation timeframe;
- For the spanwise velocity field ( $w/U$ ), the analysis trajectory is substantially closer to the true reference than the reconstructed observations. As the reconstruction methodology slots in 2D (x-y) planes for each spanwise point, the continuity of the velocity field along the spanwise direction is destroyed resulting in unphysical spanwise velocity structures in the reconstructed observations and thus the analysis appears to respect the dynamical model more than the observations;
- The plots show that the reconstruction methodology applied at each observation time-step (marked by a cross in the plots) provides overall better estimate of the true field than reconstructing an initial condition and evolving said condition using the NS equations (i.e. the background (in blue)).

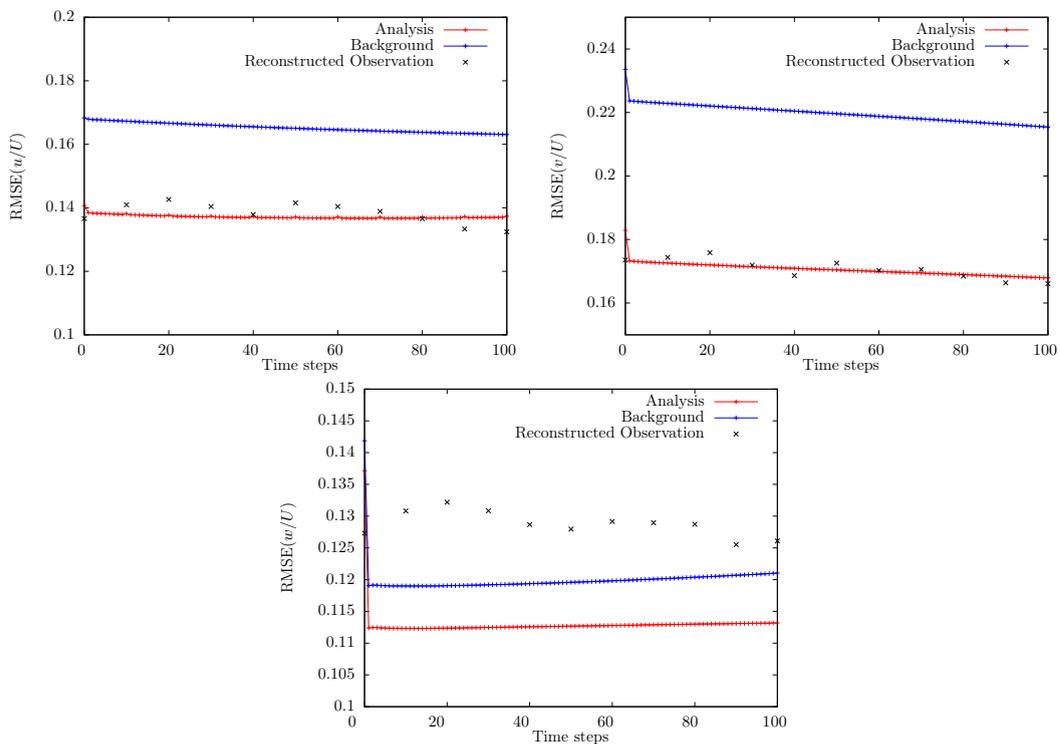


Figure 7.28: Evolution of the RMSE between the 3D true reference field and the analysis (red), the background (blue), and the reconstructed observation (black) for all three velocity fields.

The results stress on the importance of constructing a proper observation data-set. The interpolated observations introduced numerical instability in the algorithm resulting in a

numerical explosion of the analysis trajectory. The analysis evolution obtained with averaged SO reconstructed observations is stable and provides an acceptable solution. While multiple ways can be envisaged to create 3D observations from limited 2D data-sets, the SO or the averaged SO methods provide efficient and very fast algorithms for creating such observations as well as background initial conditions. It is interesting to note that using observations and background created using the same method (for example, the averaged SO method) still provides an improvement on the background as there are multiple observations encoding different turbulent structures as compared to the background which is constructed at the initial time and evolved in time.

## 7.4 Background covariance characterisation

The SO reconstruction methodology provides the capability to estimate a fully-defined background covariance matrix for use in the 4D-Var optimisation algorithm. The use of an accurate, well-defined background covariance matrix  $\mathbf{B}$  should provide significant improvements to the efficiency of the algorithm with faster convergence. However, the construction of the  $\mathbf{B}^{-1}$  matrix is usually cumbersome requiring extensive computations and/or memory due to its considerable size. In this section, we use an innovative method, assisted by the SO algorithm, for estimating this matrix using multiple reconstructions.

Consider the fully synthetic case with 3D observations using SO reconstructed background initial condition. The base simulation for producing snapshots is run for a longer period of time to produce 4000 snapshots spaced 50 time steps apart, with an addition 1600 snapshots recorded. Instead of using all the 4000 observation plane snapshots for one single reconstruction, we use a sliding window for the number of OP snapshots used to create multiple reconstructions i.e. we use the first 400 OP snapshots for reconstruction #1, 200-600 for #2, 400-800 for #3, and so on giving us 19 3D velocity fields at an instant of time  $t'_1$ . Removing the mean and decomposing the set of reconstructed fields using singular value decomposition (SVD) gives an estimate for the background covariance matrix through the velocity fluctuations ( $\mathbf{u}'$ ) as,

$$\mathbf{u}' = USV^T, \quad (7.4.1)$$

where  $U$  contains the left singular vectors as columns,  $V^T$  contains the right singular vectors as rows, and  $S$  is a diagonal matrix composed of the singular values in decreasing order of magnitude. The background covariance matrix and its inverse can then be defined as,

$$\mathbf{B} = \mathbf{u}'\mathbf{u}'^T = (USV^T)(USV^T)^T = US^2U^T. \quad (7.4.2)$$

$$\mathbf{B}^{-1} = (US^2U^T)^{-1} = US^{-2}U^T. \quad (7.4.3)$$

$$\mathbf{B}^{-\frac{1}{2}} = US^{-1}U^T. \quad (7.4.4)$$

With the computed  $\mathbf{B}^{-\frac{1}{2}}$ , the components of the cost function and its gradient with respect to the background can be modified as,

$$J(\mathcal{X}_0) = \frac{1}{2} \|\mathcal{X}_0^b - \mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 = \frac{1}{2} (\|\mathcal{X}_0^b - \mathcal{X}_0\|_{\mathbf{B}^{-\frac{1}{2}}}) (\mathbf{B}^{-\frac{1}{2}T} \|\mathcal{X}_0^b - \mathcal{X}_0\|^T) \quad (7.4.5)$$

$$\frac{\partial J}{\partial(\delta\mathcal{X}_0)} = (\delta\mathcal{X}_0 \mathbf{B}^{-\frac{1}{2}}) ((\mathbf{B}^{-\frac{1}{2}})^T). \quad (7.4.6)$$

Such an evaluation takes into account the effect a change in velocity at one point has on other points in a neighbourhood through the inverse of the  $\mathbf{B}$  matrix. This should result

in faster convergence due to a distributed optimisation procedure where convergence at each mesh-points assists in a better estimate of the field at other points with respect to the background condition. This is studied by performing an assimilation with this estimated  $\mathbf{B}^{-\frac{1}{2}}$  matrix (case (f)). Identical parameters to case (c) are used with the averaging SO reconstructed background and full 3D observations.

The cost functional and its gradient evolution with the updated  $\mathbf{B}$  matrix are shown in figure 7.29 and figure 7.30 and compared with case (c) where the  $\mathbf{B}$  matrix was set to identity. A significantly faster convergence is obtained with case (f) (26 inner loop iterations) requiring only half the iterations of case (c) (58 inner loop iterations) to converge to an optimal analysis. The final cost is marginally higher in this case due to the increased weighting of the background error caused by the fully defined background covariance matrix. The gradient of the cost function for the initial set of iterations are identical for both cases but as the optimal trajectory deviates further from the background, the covariance matrix plays a role and thus, the two gradient curves deviate from each other. These curves suggest that a good estimate of the background covariance matrix provides a significant computational cost reduction by improving the efficiency of the optimisation algorithm. However, the accuracy of the algorithm needs to be analysed further.

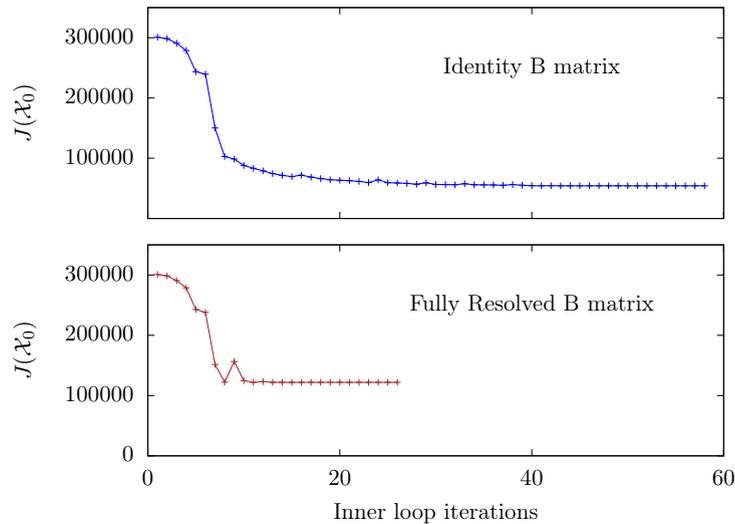


Figure 7.29: Evolution of the cost function  $J(\mathcal{X}_0)$  for the assimilation as a function of inner loop iterations with identity (case (c)) and fully-resolved (case (f))  $\mathbf{B}$  matrix.

The optimal analysis provided by the fully-resolved  $\mathbf{B}$  matrix case is shown in figure 7.31 at time  $t'_1$ . Comparing with case (c) (see figure 7.17), we can see that the two analysis trajectories are comparable and both match well with the reference. The performance is better analysed through the error function for the velocity plotted in figure 7.32. The assimilation algorithm improves on the background condition with either covariance matrix, however, the analysis from case (c) is closer to the true reference than with case (f). This discrepancy, despite better information provided through the background covariance matrix, can be attributed to the lack of detailed information on the observations e.g. through the observation covariance matrix, or on the inlet condition. The significant reduction in computation cost obtained with case (f) is slightly off-set by the loss of accuracy in the analysis. However, this could be improved by estimating the observation covariance matrix using a similar algorithm and by weighting appropriately the various contributions to the cost function.

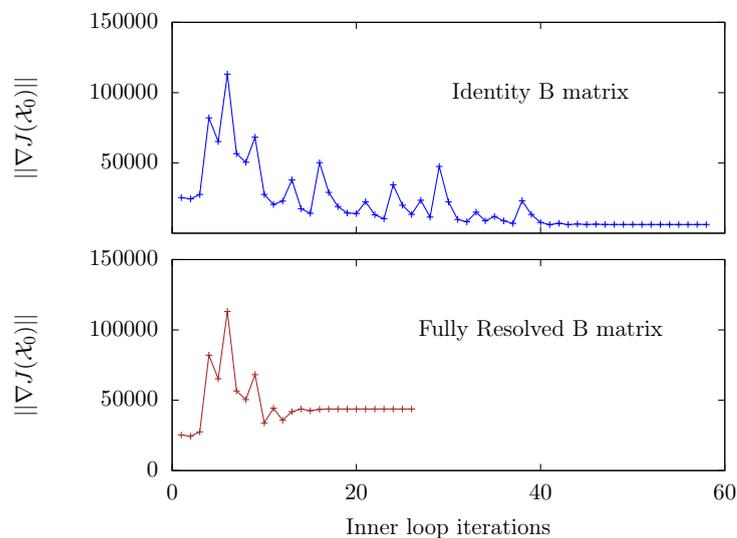


Figure 7.30: Evolution of the gradient of the cost function  $\|\nabla J(\chi_0)\|$  for the assimilation as a function of inner loop iterations with identity (case (c)) and fully-resolved (case (f))  $B$  matrix.

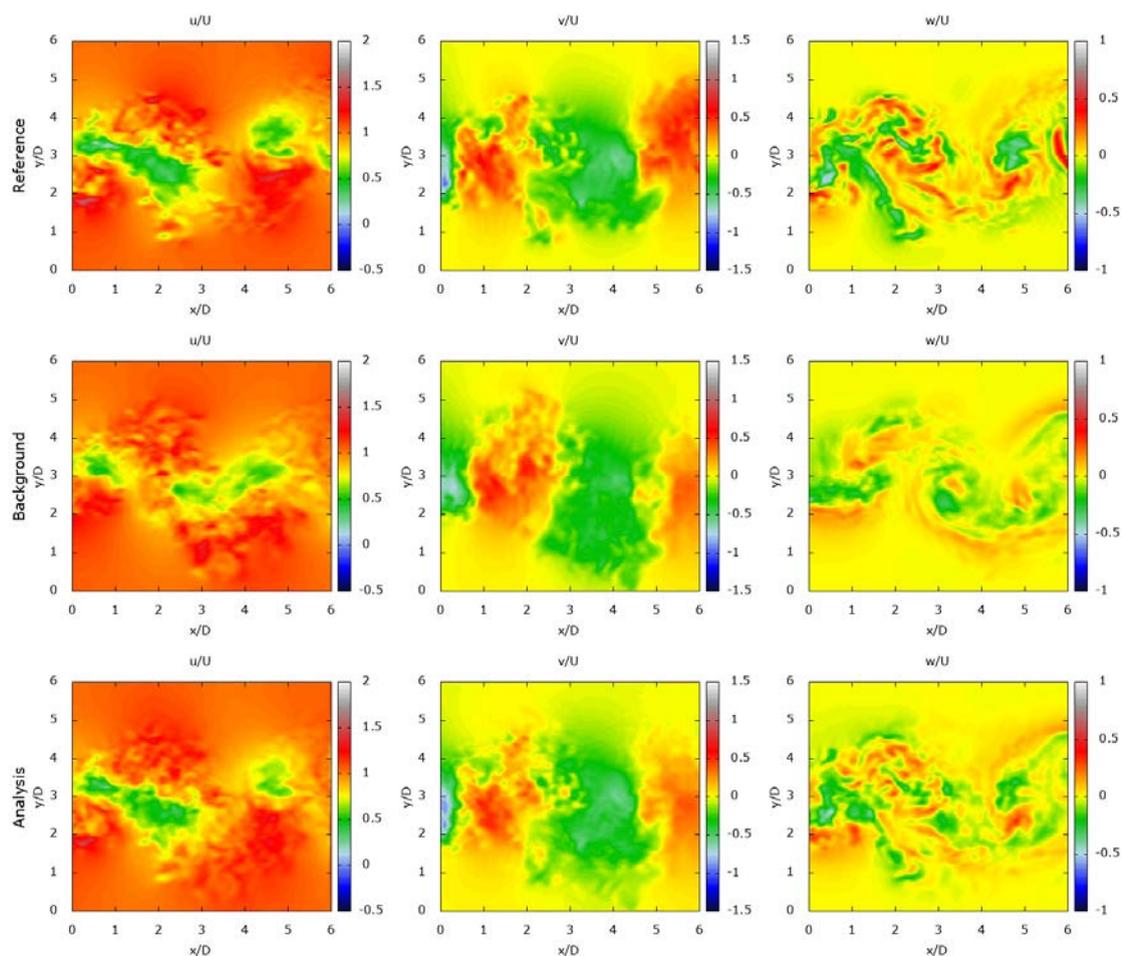


Figure 7.31: Instantaneous velocity snapshots at the initial time of the assimilation window ( $t'_1$ ) in the streamwise plane  $z = 1.31D$ . From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: Reference, background, analysis (case (f)).

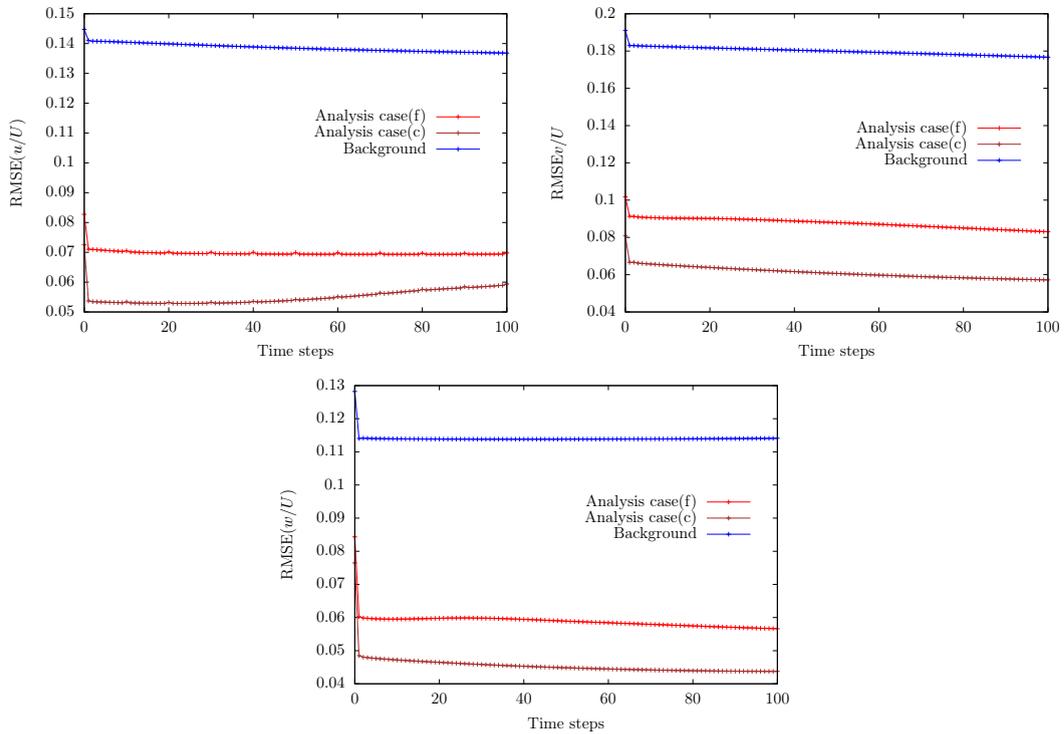


Figure 7.32: Evolution of the RMSE between the 3D true reference field and the analysis for case (f) (red), analysis for case (c) (brown) and the background (blue) for all three velocity fields.

## 7.5 Model coefficient estimation

This section focuses on the possibility of estimating model coefficient using the 4D- $\text{Var}_{les}$  algorithm. Estimating of coefficient of a turbulence model using data assimilation in 3D is a relatively new concept and thus, is still in its infancy. This chapter provides a proof of concept towards the capability of the 4D- $\text{Var}_{les}$  to estimate the model coefficient for a turbulent flow. Thus, an ideal case study (case (g)) is analysed with full-scale 3D observations and a background set to the optimised initial condition from case (a) with the background covariance matrix set to Identity. An initial estimate for the coefficient is provided using eq. (2.2.47) - for a uniform mesh, this results in a constant coefficient estimate in time and space. The optimised coefficient is enforced to be constant in time but allowed to vary in space. A strict penalty is enforced on the coefficient to ensure it does not vary far from the background value provided. This is important for multiple reasons:

- The coefficient defines the amount of dissipation brought in by the model and the stability of the flow is predicated on this;
- The contribution to the cost function of the coefficient is only through the deviation from the initial estimate provided and this is multiple orders of magnitude smaller than the observation error. Thus, the algorithm tends to modify freely the coefficient in order to reduce the cost function in lieu of the initial/inlet condition.

The cost function including the additional control on the coefficient is given as,

$$J(\delta\mathcal{X}_0, \delta C) = \frac{1}{2}\|\delta\mathcal{X}_0\|_{\mathbf{B}^{-1}}^2 + \frac{1}{2}\|\delta C\|_{\mathbf{B}_c^{-1}}^2 + \frac{1}{2}\int_{t_0}^{t_f} \|\mathbb{H}(\mathcal{X}_t) - \mathcal{Y}(t)\|_{\mathbf{R}^{-1}}^2 dt, \quad (7.5.1)$$

where  $\delta C = C_0 - C_b$  is the difference between the coefficient estimate and the background value provided, and  $\mathbf{B}_c^{-1}$  is the corresponding covariance which is set to Identity. The gradient, using the adjoint approach, is defined as,

$$\frac{\partial J}{\partial(\delta\mathcal{X}_0)} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}^{-1}\delta\mathcal{X}_0, \quad (7.5.2)$$

$$\frac{\partial J}{\partial(\delta C)} = -\boldsymbol{\lambda}_{t_0} + \mathbf{B}_c^{-1}\delta C + (\partial_{\mathbf{u}}\mathbb{M})^*\boldsymbol{\lambda}. \quad (7.5.3)$$

where  $(\partial_{\mathbf{u}}\mathbb{M})^*$  refers to the adjoint dynamical model with respect to the coefficient formulated using TAPENADE. It is interesting to note that due to the linear nature of the coefficient in the dynamical model, the tangent dynamical model with respect to the coefficient is the model itself without the coefficient.

The optimisation characteristics for the assimilation (case (g)) are shown in figures 7.33a and 7.33b. The significance of the coefficient is well expressed through the maxima in the cost function obtained after the first iteration of the algorithm. Even an incremental variation of the coefficient has a considerable effect on the flow. A large value of the gradient for this iteration demonstrates the strong dependance of the flow and its evolution on the LES model through the model coefficient. However, despite the initial large peaks, the algorithm converges quickly to a cost which is slightly smaller than the initial cost - the error introduced by the coefficient is offset by an improvement in match with the observation.

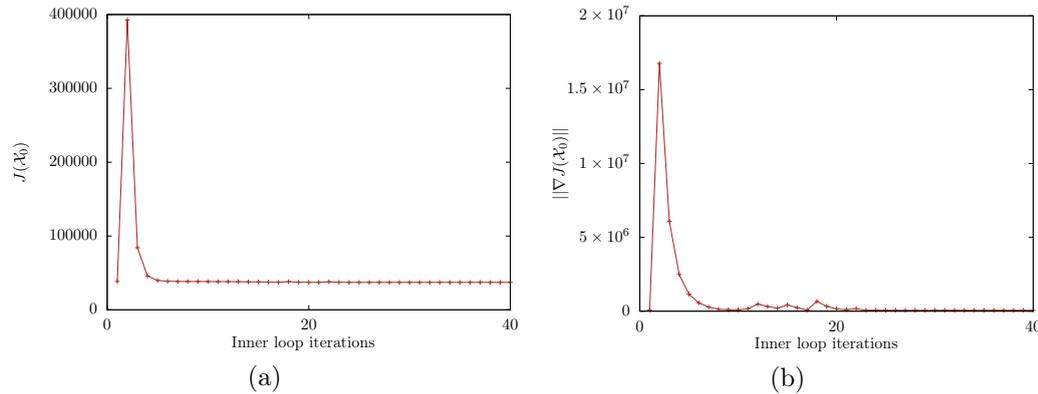


Figure 7.33: Evolution of the cost function  $J(\mathcal{X}_0)$  and its gradient  $\|\nabla J(\mathcal{X}_0)\|$  for the assimilation as a function of inner loop iterations for coefficient estimation case (g).

The reference, background, and analysis velocity maps at the final time of the assimilation  $t'_2$  are shown in figure 7.34. Given that the background provided to the algorithm was already an optimal condition, minimal variation is observed in the analysis and background both of which match well with the reference. The root mean square between the reference and the background/analysis for the three velocity components is plotted in figure 7.35. By optimising the coefficient, we obtain a marginal improvement on the analysis mainly in the streamwise component - the analysis evolves more accurately when the coefficient has been optimised as shown by the red curve for the streamwise velocity in figure 7.35

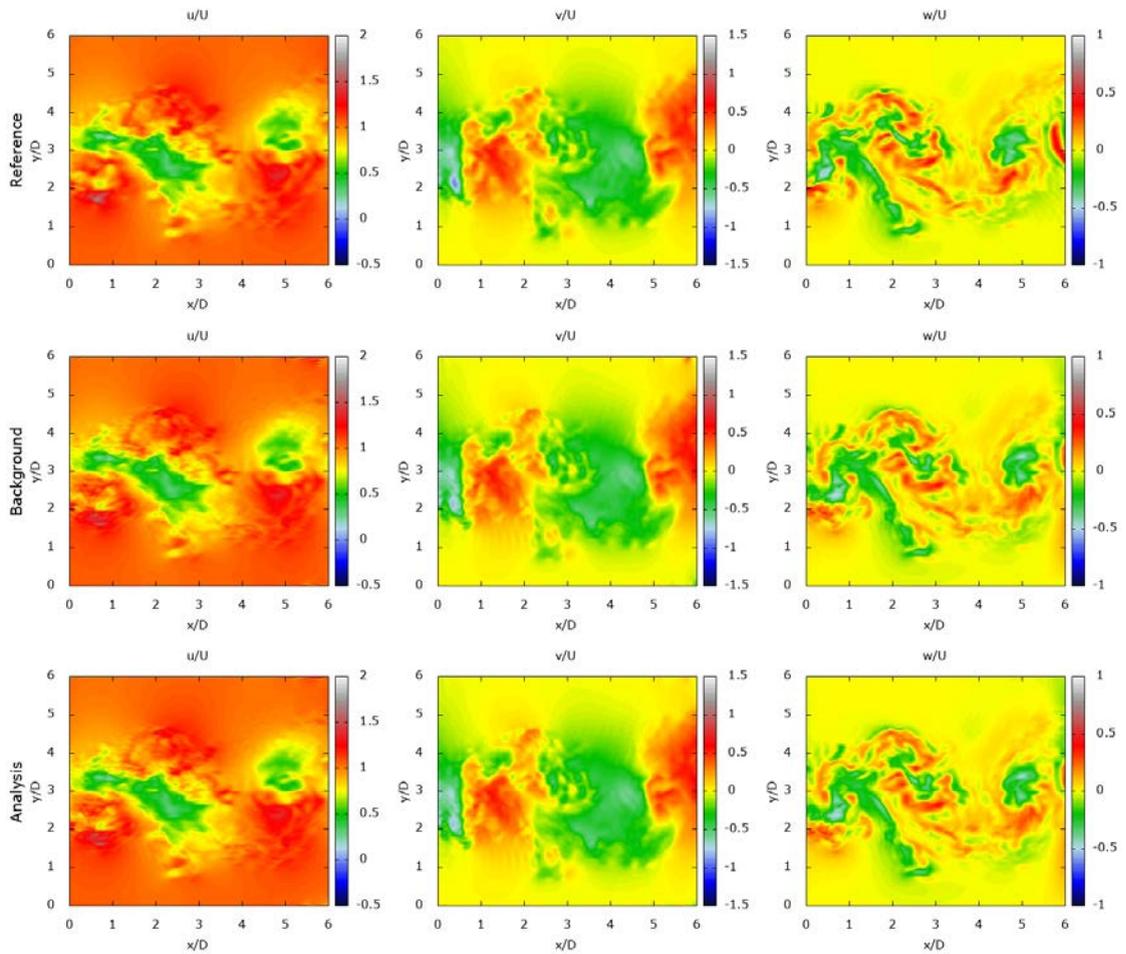


Figure 7.34: Instantaneous velocity snapshots at the final time of the assimilation window ( $t'_2$ ) in the streamwise plane  $z = 1.31D$  for case (g). From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: reference, background, and analysis for case (g).

The estimated coefficient is shown in figure 7.36 along with the initial estimate provided to the algorithm. The optimised coefficient is seen to vary from the initial estimate over certain regions which can be correlated to vortices shed from the cylinder i.e. regions with higher turbulent fluctuations. This suggests that the optimised procedure adheres to flow physics in optimising the coefficient. A region of high turbulence implies stronger small scale fluctuations and thus increased contribution of the turbulence model. Thus, regions with higher model coefficient implies a stronger impact of the small scales on the large scale resolved velocity. It also suggests that the parameter should be a function of space and of time especially as the assimilation window is enlarged to account for multiple vortex sheddings - a possible focus for future work. The effect of the modified coefficient is difficult to observe on the velocity maps for this idealised case-study. By providing an initial condition that is already optimised, the algorithm is limited in its ability to explore the range of coefficient values to provide an improvement. This is in fact the counter argument to our assertion that if a free reign on the coefficient is provided to the algorithm, it could exclusively modify the coefficient and not the other control parameters. Thus, a moderation needs to be achieved between the two extremes and this is studied in the next and final

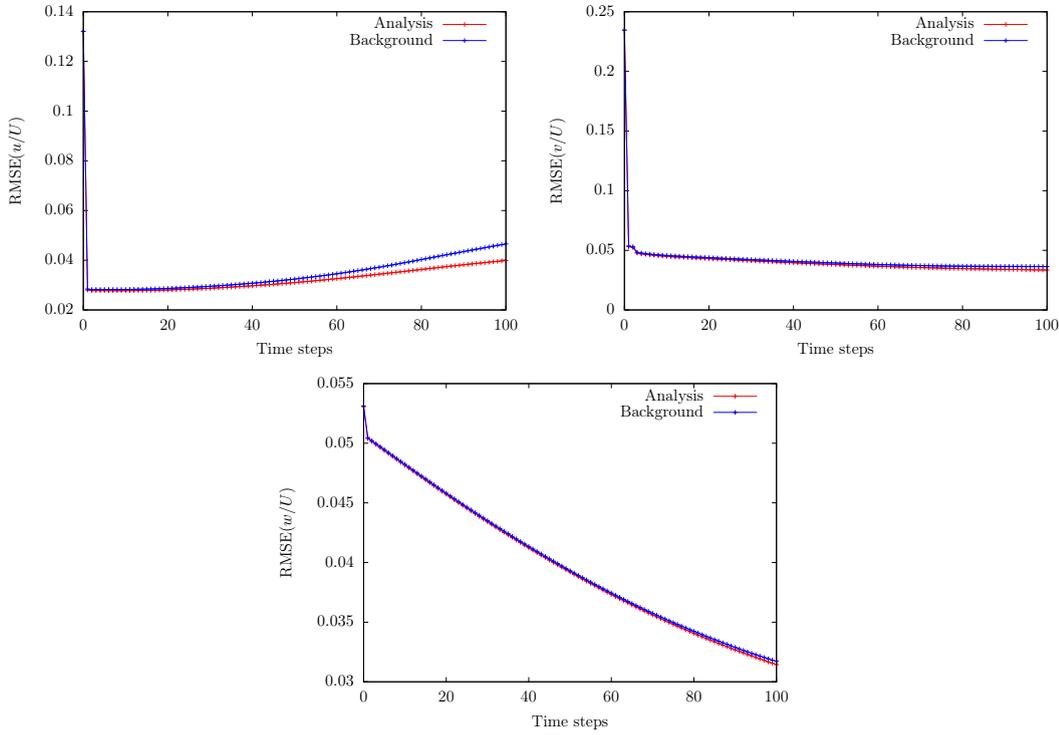


Figure 7.35: Evolution of the RMSE between the 3D true reference field and the analysis (red), and the background (blue) for all three velocity fields for case (g).

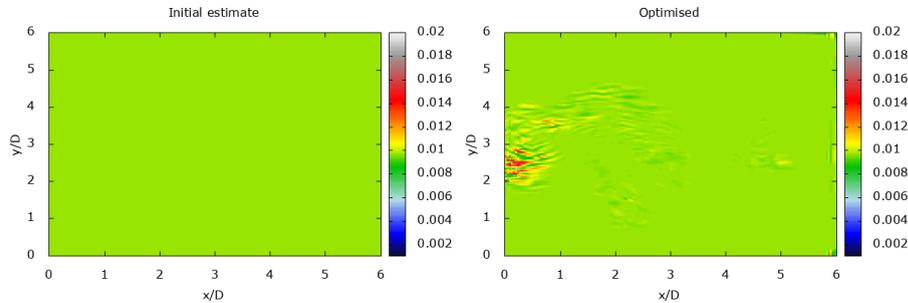


Figure 7.36: LES model coefficient contour map in the streamwise plane  $z = 1.31D$  for case (g). From left to right: background, and analysis.

section for the complex case of coefficient optimisation with 2D reconstructed observations, reconstructed background, and fully-defined background covariance matrix.

## 7.6 Combined case study

This chapter has successfully analysed individually the effects of different background conditions, reconstructed volumetric observations, background covariance contribution, and the ability to estimate the model coefficient. This final section now combines all the analysis in to one final assimilation with averaging SO reconstructed volumetric observations, SO reconstructed background condition, a full estimate of the background covariance matrix, and coefficient optimisation. A two step optimisation procedure is carried out: first, the initial and inlet condition are optimised until a relative error tolerance of 1% in the

cost function between consecutive iterations. This is followed by the optimisation of the initial and inlet condition in combination with the coefficient - this is one approach towards achieving the moderation discussed in the previous section. Such a two-step optimisation is necessary as a full optimisation from the start would tend to focus on the coefficient which provides more freedom to the algorithm than the initial/inlet condition. The background covariance matrix as estimated in §7.4 is used in the algorithm.

All the components used for the assimilation in this case-study are obtained and processed using algorithms developed in this thesis. The reconstruction algorithms are used to develop meaningful initial background and volumetric observations. Multiple reconstructions are used to estimate the background covariance matrix. The LES based assimilation algorithm is used to estimate the coefficient of the incorporated turbulence model. Unlike §7.5, the background provided is not an optimal analysis and the coefficient estimation is performed with only an estimate of the background condition.

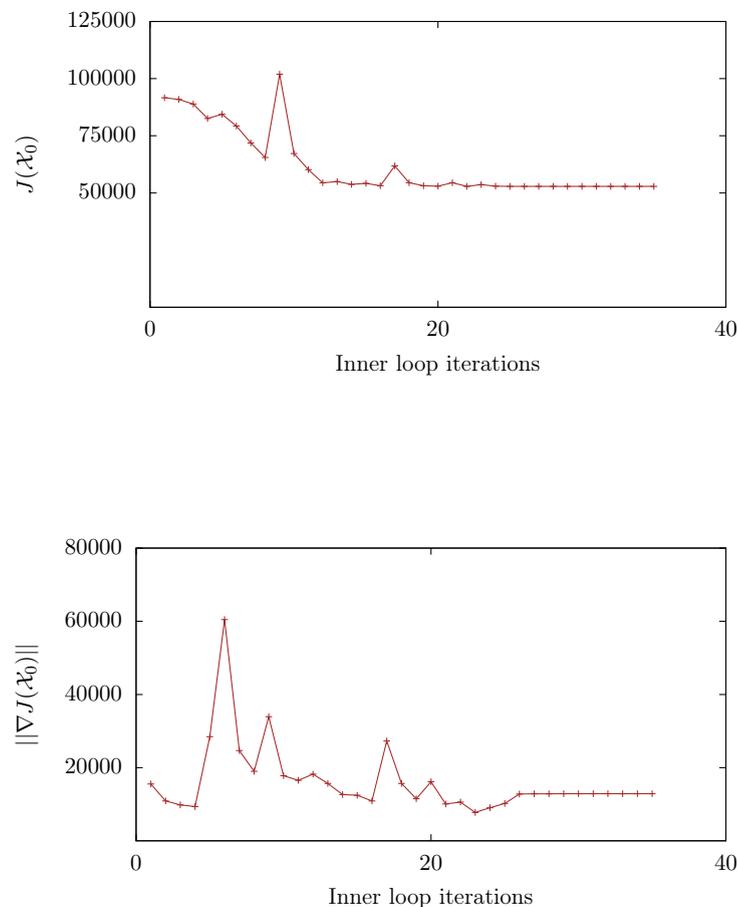


Figure 7.37: Evolution of the cost function  $J(\mathcal{X}_0)$  for the assimilation as a function of inner loop iterations for the combined case (h).

The cost functional and its gradient for case h are plotted in figure 7.37. A 50% reduction in cost is observed similar to case (e), with only two-thirds the number of iteration as required for case (e). This is due to the well-defined background covariance matrix which was shown to reduce significantly the computational cost for performing variational assimilation. The analysis velocity at time  $t_1$  is a compromise between the observation

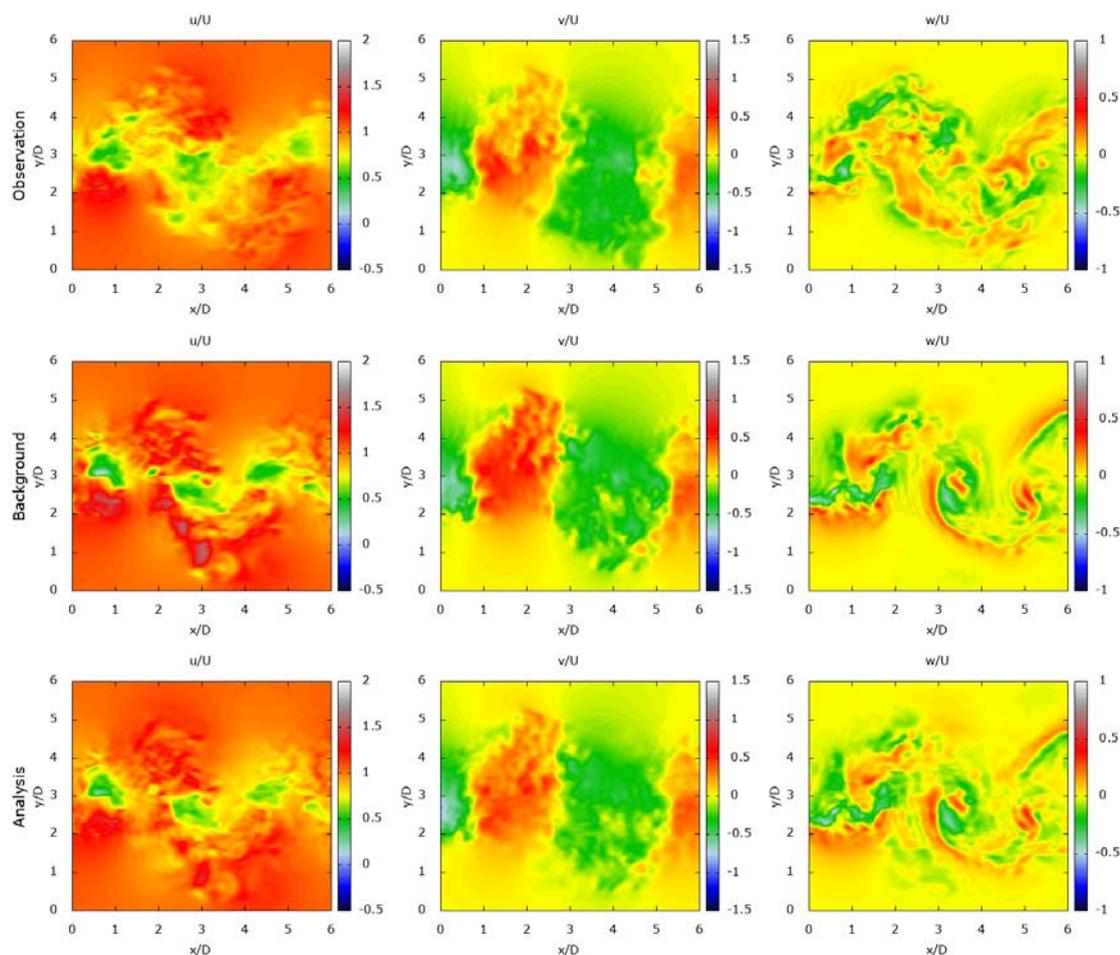


Figure 7.38: Instantaneous velocity snapshots at the initial time of the assimilation window ( $t'_1$ ) in the streamwise plane  $z = 1.31D$  for the combined case (h). From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: reference, background, and analysis.

and the background (see figure 7.38). The reduced confidence on the observation (see figure 7.22), results in an initial optimal trajectory that is quite close to the background. This compromise between the two is sustained through the evolution of the initial condition to the final assimilation time  $t_2$  (see figure 7.39). The deviation from the background is more pronounced at time  $t_2$  as compared to time  $t_1$ .

A more quantitative comparison is performed by plotting the RMSE between the three fields, namely observation, background, and analysis, and the true reference field for each velocity component in figure 7.40. The analysis trajectory at all time-steps is considerably closer to the true reference field than the background condition. The observations for the streamwise ( $u$ ) and lateral ( $v$ ) velocity have the least error which is an added acclaim for the SO reconstruction method. However, for the spanwise ( $w$ ) velocity, due to discontinuities, the observation error is quite high. The assimilation algorithm, by enforcing the divergence free condition, provides an analysis trajectory of lower error than both the observation and the background for the spanwise velocity field while for the other two components it achieves a compromise between the background and the provided observations.

Finally, the initial and optimised coefficient field is plotted in figure 7.41. The optimised

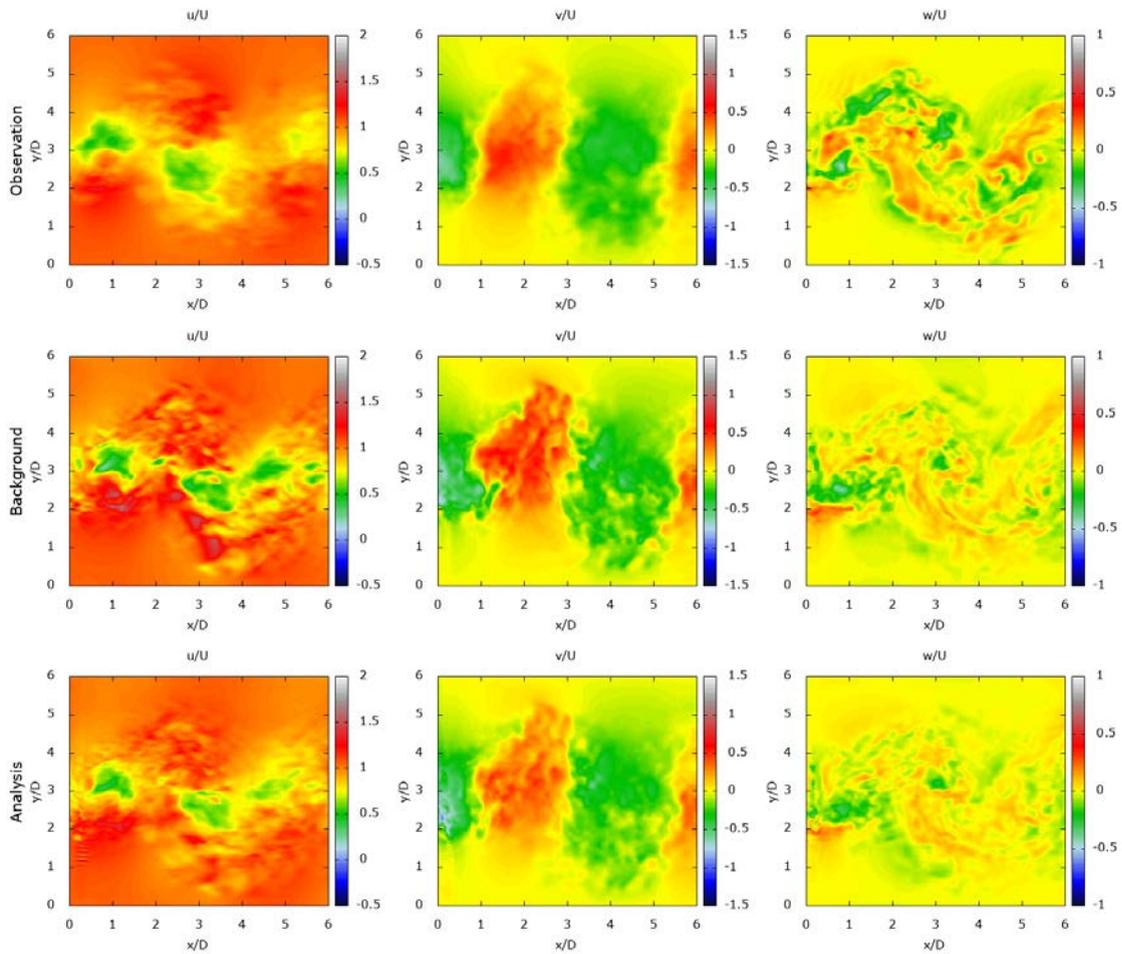


Figure 7.39: Instantaneous velocity snapshots at the final time of the assimilation window ( $t'_2$ ) in the streamwise plane  $z = 1.31D$  for the combined case (h). From left to right: Streamwise, lateral, and spanwise velocity field. From top to bottom: reference, background, and analysis.

coefficient conforms to the turbulent regimes of the flow. A correlation can indeed be observed between the regions with high value of the coefficient and the vortex shedding observed in the streamwise velocity field. Compared with case (g), a more widespread modulation is obtained for the coefficient - this is due to the higher degree of freedom allowed to the algorithm in case (h) as compared to case (g). The physical influence of this coefficient can also be observed on the analysis velocity maps at time  $t_2$  for case (h) (figure 7.39) by comparing it with case (e) (figure 7.27). The difference between the two cases is the use of a fully defined background covariance matrix and control via coefficient optimisation. The former plays no role in the time-evolution of the flow field and one can see that at the initial time  $t_1$  the analysis for cases (e) and (h) are quite similar (see figure 7.26 and figure 7.38). However, at the final time, the velocity maps for case (h) are comparatively smoother than case (e) - this difference should be due to the higher value of the coefficient which would entail higher dissipation and thus smoother velocity fields. A longer time evolution of the flow would provide more prominent variations between the two cases due to the different coefficient maps.

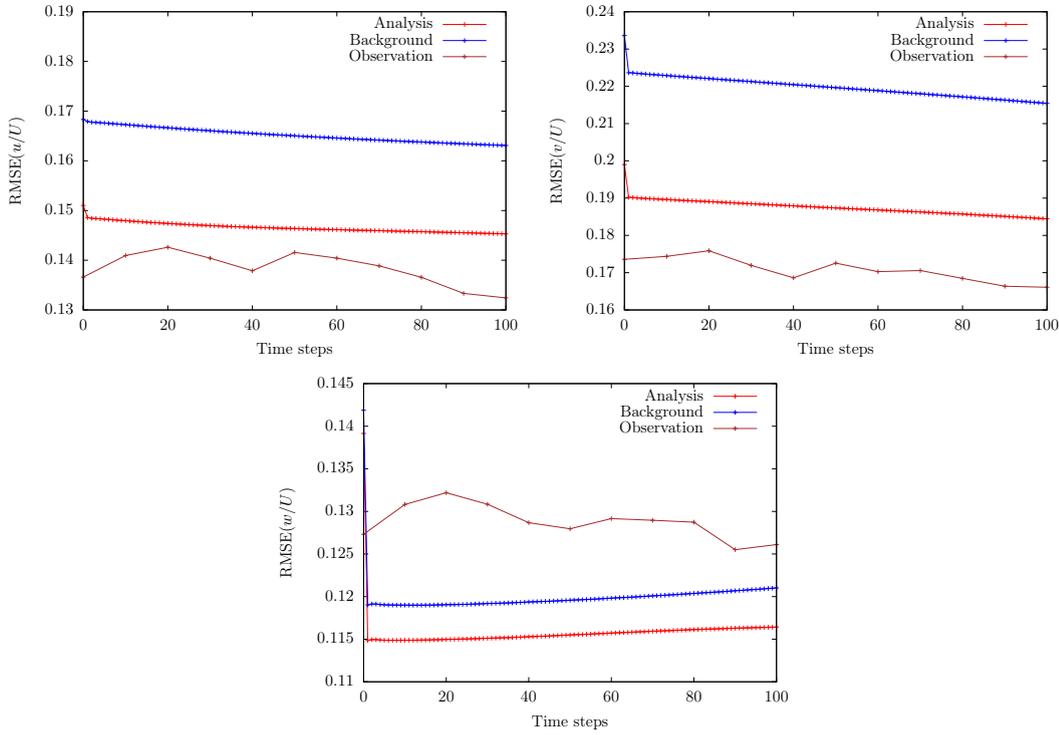


Figure 7.40: Evolution of the RMSE between the 3D true reference field and the analysis (red), the background (blue), and the reconstructed observation (brown) for all three velocity fields for case (h).

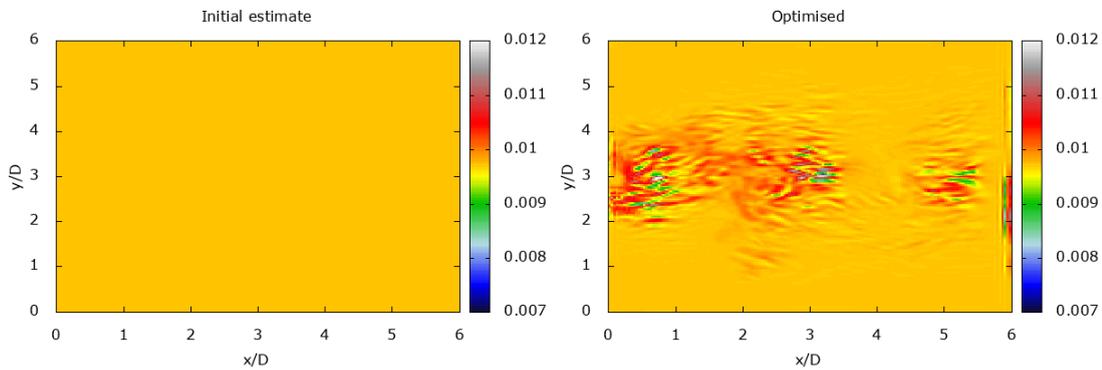


Figure 7.41: LES model coefficient contour map in the streamwise plane  $z = 1.31D$  for case (g). From left to right: background, and analysis.

## 7.7 Conclusions

This chapter successfully combines the various tools and formulations developed in the preceding chapters to formulate, validate, and apply the  $4D\text{-Var}_{les}$  code. The code development is enumerated in §7.1 alongside the concepts of auto-differentiation and associated numerical tools/software. A first assessment of the algorithm is done in §7.2 using true 3D observations. Multiple background conditions are analysed with the algorithm: the background constructed from the averaged SO reconstruction method (developed in chapter 6) is observed to be the optimal choice with faster convergence and an optimal analysis

trajectory. The algorithm is shown to work with ‘incorrect’ estimates for the background obtained by biasing the observations. The algorithm is also applied to observations emulating an experimental 2D-3C cross-planar PIV data-set. The extracted data is used to construct 3D-3C volumetric observations using two reconstruction algorithms: the weighted interpolation method (Robinson, 2015), and the averaged SO reconstruction method. The latter is shown to provide good results with the assimilation while the former destabilised the flow - the interpolation method is applicable to flows at low Reynolds number but for more turbulent flows, such a method introduces instabilities.

A new method is proposed to estimate the background covariance matrix which, when incorporated, is shown to reduce significantly the computational time required for optimising the run with similar accuracy. An attempt is also made to estimate the LES model coefficient using the 4D-Var<sub>les</sub> code with 3D true observation and a background defined from the optimised initial condition of case (a). A physically relevant coefficient map is obtained with higher estimate in turbulent regions - the enforcement of a strict penalty (to ensure stability) tends to limit the deviation of the estimation from the initial estimate provided. This application is a demonstration towards the capabilities of the algorithm to provide a physically meaningful estimate for the parameter of interest, namely, the model coefficient. Finally, the algorithms capabilities are analysed with the complex case of partial observations, reconstructed background condition, a fully defined background covariance matrix, along with coefficient estimation. The results are optimistic towards the expansion of the variational assimilation methodology, combined with the various techniques, to higher Reynolds number flows. Further assessment with respect to experimental observations, observation covariance matrices, different flow configurations and long-time assimilation windows are necessary to generalise algorithm accuracy and stability.



# Conclusions and perspectives

## Conclusions

Data assimilation (DA) has become an integral part of fluid mechanics over the last two decades due to its ability to bridge the limitations of EFD and CFD. However, DA is significantly more complex to implement and requires stringent pre-processing using a wide variety of tools in order to make it feasible. A major restriction of DA is the computational cost requirements. The 4D variational assimilation procedure (4DVar) requires at least one order of magnitude higher computational power than a DNS for the same flow parameters. This thesis has attempted to redress the various limitations of the variational DA procedure.

A reduction in computational cost can be achieved by moving towards turbulence modelling following along the path laid by Smagorinsky in 1963 for fluid simulations. Existing turbulence modelling approaches and models for flow simulations are enumerated in chapter 2. Large eddy simulation (LES) is preferred over other techniques as it is better suited for representing turbulent structures in a time-resolved simulation. Chapter 3 analyses the performance of various LES models for turbulent flows. Classical LES models developed using deterministic principles such as classic and dynamic Smagorinsky are compared with the stochastically derived models under location uncertainty (MULU). The local spatial variance model under the framework of MULU are shown to improve on the classical models for two turbulent flows: channel flow at  $Re_\tau$  of 590, and wake flow around a circular cylinder at  $Re$  of 3900. Hence, this model is chosen for implementation in the 4DVar<sub>les</sub> algorithm.

A non-exhaustive literature on DA is briefly provided in chapter 5. The mathematical principles behind variational assimilation and its incremental alternative along with optimisation of control parameters are also enumerated. Chapter 6 formulates and tests a novel volumetric reconstruction technique using 2D cross-planar data-sets. Such reconstruction techniques has numerous applications for DA, namely to:

- Reconstruct physical background initial conditions.
- Reconstruct volumetric observations from sparse 2D observations.
- Provide a representation of the background covariance matrix for the minimisation algorithm.

The techniques and methodologies observed and formulated in the previous chapters are combined with the 4DVar procedure in chapter 7 to give the 4DVar<sub>les</sub> algorithm. A brief introduction to the DNS based 4DVar algorithm is provided along with the necessary tools required for code formulation. Validation procedures for the tangent and adjoint formulations of the dynamics are enumerated. The development of the 4DVar<sub>les</sub> code is also presented in detail. A validation study of the code is performed with 3D observations using multiple background conditions - a biased background, a SO reconstructed background, and

an averaged SO reconstructed background. The utility of the SO reconstruction methods is seen via faster convergence and a better estimate of the analysis trajectory with the reconstructed background conditions. It was interesting to note that despite an ‘incorrect’ biased background, the algorithm was capable of recovering the observation trajectory. A novel method is introduced to define a fully-resolved background covariance matrix from the SO reconstruction methodology. This well-defined matrix is shown to provide computational advantages in terms of cost and time.

The algorithm capabilities are further analysed with a more realistic case of 2D observation. The reconstruction methodology from chapter 6 is used to construct 3D volumetric observation from the 2D data-sets. An alternative observation set is also constructed using a mirroring technique. The results demonstrate the need for a properly constructed observation data-set. With the use of mirror imaged observations, the algorithm was shown to predict an initial condition which was unphysical and thus led to numerical explosion. An acceptable solution is obtained by using the SO reconstructed observations.

Another application of the algorithm focuses on coefficient estimation for the LES model. By introducing the coefficient as a control parameter, an estimate of the coefficient can be obtained from the optimisation procedure. A preliminary study, using an optimal analysis trajectory from previous case as the background condition, provides a physically meaningful estimate for the coefficient with higher values in turbulent regions and lower values when the flow is laminar. A more complex study involving partial observations, and reconstructed background was also shown to estimate a physically meaningful coefficient along with an acceptable analysis trajectory.

A detour is taken from DA in the final chapter of this thesis work where the fully stochastic version of the MULU is implemented and studied in the context of bridging turbulence inhomogeneity, streaks and wave-current interaction. An attempt is made to theorise the origin of streaks in wall-bounded flows by borrowing analogies from oceanic modelling concepts, namely Langmuir circulation and vortex forces. The underlying physics behind streak formation in turbulent channel flow is explained in terms of vortex forces. The MULU are mathematically shown to incorporate these vortex forces through the variance tensor modelling turbulence inhomogeneity. The signature of these vortex forces are observed in a high resolution DNS of channel flow. In addition, a large-scale simulation at low-resolution of the same channel is shown to recover the streaks when the turbulence is modelled using the MULU but not with classical LES modelling.

## Perspectives

Further study into the following subjects could have high import and interest:

### **Assimilation with experimental observations**

A proof of concept is provided in this thesis for variational data assimilation using one turbulence model and synthetic data-sets. Its expansion to experimental data-sets and an analysis of the algorithm’s ability to work with noisy, sparse experimental data-sets is the logical next-step to this thesis work. With increasingly efficient experimental techniques, assimilation of 3D volumetric data obtained directly from experiments (tomo-PIV, for example) is also a promising avenue of research.

### **Outflow and boundary condition control**

Additional control can be implemented on the outflow as well as the boundaries in order to reduce errors in the assimilation procedure. Numerical oscillation are induced by the

algorithm due to incorrect outflow condition but these can be controlled and damped by optimising the outflow at each time-step in line with the inflow optimisation. Inclusion of boundary condition control, for Dirichlet, Neumann, or mixed conditions, would facilitate assimilation of complex flows with a wide variety of boundary conditions not limited to the free-slip/periodicity condition enforced for case-studies in this thesis.

#### **Additional model integration**

This thesis proved the possibility of numerically merging variational assimilation principles with a sub-grid scale turbulence model. However, as is well known in the turbulence field, one model is rarely able to capture and reproduce accurately all types of flows. A variety of models need to be implemented to facilitate a generic algorithm capable to simulating a large variety of turbulent flows.

Expansion to steady state modelling through RANS methodologies can have some interesting applications. One application, currently under study, is to assimilate pressure measurements around high-rise structures to obtain detailed flow patterns and pressure distribution useful for building design optimisation studies. The shift to steady state modelling would provide highly efficient and fast assimilation algorithms as compared to time-resolved LES or DNS based algorithms.

#### **Numerical framework for multi-source application**

In a numerical sense, to restrict the application of a procedure to one computer configuration and one programming language is strictly prohibitive. It will be a future necessity to shift towards a general framework that can be implemented and evaluated on multiple configurations (namely CPU or GPU) and with multiple programming languages. This work would be along the lines of the OP2/OPS project undertaken at the University of Oxford e-Research centre. OPS and OP2 are high-level frameworks for structured and un-structured grids, respectively, that are capable of generating parallel executables that can be run on any of single-threaded CPU/ CUDA parallelised GPUs/ OpenMP based multi-threaded CPU systems. The framework also supports development in both C++ and FORTRAN 90.



# Bibliography

- R. J. Adrian. Scattering particle characteristics and their effect on pulsed laser measurements of fluid flow: speckle velocimetry vs particle image velocimetry. *Appl. Opt.*, 23(11):1690–1691, 1984.
- S. Ahuja and C. W. Rowley. Feedback control of unstable steady states of flow past a flat plate using reduced-order estimators. *J. Fluid Mech.*, 645:447–478, 2010.
- J. D. Anderson. Brief History of the Early Development of Theoretical and Experimental Fluid Dynamics. In Richard Blockley and Wei Shyy, editors, *Encyclopedia of Aerospace Engineering*. John Wiley & Sons, Ltd, Chichester, UK, 2010.
- G. Arroyo, J. Carlier, P. Georgeault, A. Guibert, D. Heitz, and L. Wallian. FROILOC<sup>®</sup> a novel food safety technology through localised cold and clean airflow. In *8th International Conference on the Food Factory of the Future*, 2016.
- M. P. Arroyo and C. A. Greated. Stereoscopic particle image velocimetry. *Meas. Sci. Technol.*, 2(12):1181, 1991.
- G. Artana, A. Cammilleri, J. Carlier, and E. Mémin. Strong and weak constraint variational assimilations for reduced order fluid flow modeling. *J. Comput. Phys.*, 231(8):3264–3288, 2012.
- O. O. Badmus, S. Chowdhury, and C. N. Nett. Nonlinear control of surge in axial compression systems. *Automatica*, 32(1):59–70, 1996.
- S. Bagheri, L. Brandt, and D. S. Henningson. Input–output analysis, model reduction and control of the flat-plate boundary layer. *J. Fluid Mech.*, 620:263–298, 2009.
- B. S. Baldwin and T. J. Barth. A One-Equation Turbulence Transport Model for High Reynolds Number Wall-Bounded Flows,. Technical Report 102847, NASA TM, 1990.
- B. S. Baldwin and H. Lomax. Thin Layer Approximation and Algebraic for Separated Turbulent Flows. Alabama, 1978.
- B. Bays-Muchmore and A. Ahmed. On streamwise vortices in turbulent wakes of cylinders. *Phys. Fluids A*, 5(2):387–392, 1993.
- P. Beaudan and P. Moin. Numerical experiments on the flow past a circular cylinder at sub-critical Reynolds number. *Stanford Univ. CA Thermosciences Div.*, 1994.
- A. Bensoussan and R. Temam. Equations stochastiques du type Navier-Stokes. *J. Funct. Anal.*, 13(2):195–222, 1973.
- P. Bergthórsson and B. R. Döös. Numerical Weather Map Analysis. *Tellus*, 7(3):329–340, 1955.

- K.Y. Billah and R.H. Scanlan. Resonance, Tacoma Narrows bridge failure, and undergraduate physics textbooks. *Am. J. Phys.*, 59:118–124, 1991.
- J. P. Boris. On large eddy simulation using subgrid turbulence models Comment 1. In *Whither turbulence? Turbulence at the crossroads.*, pages 344–353, Berlin, Heidelberg, 1990. Springer.
- J. P. Boris, F. F. Grinstein, E. S. Oran, and R. L. Kolbe. New insights into large eddy simulation. *Fluid Dyn. Res.*, 10(4-6):199–228, 1992.
- J. Boussinesq. *Essai sur la théorie des eaux courantes*. Impr. nationale, 1877.
- C. Braud, D. Heitz, P. Braud, G. Arroyo, and J. Delville. Analysis of the wake–mixing-layer interaction using multiple plane PIV and 3d classical POD. *Experiments in fluids*, 37(1):95–104, 2004.
- M. Breuer, N. Jovičić, and K. Mazaev. Comparison of DES, RANS and LES for the separated flow around a flat plate at high incidence. *Int. J. Numer. Methods Fluids*, 41(4):357–388, 2003.
- Ch Brücker. Digital-particle-image-velocimetry (DPIV) in a scanning light-sheet: 3d starting flow around a short cylinder. *Experiments in Fluids*, 19(4):255–263, 1995.
- Ch Brücker, D. Hess, and J. Kitzhofer. Single-view volumetric PIV via high-resolution scanning, isotropic voxel restructuring and 3d least-squares matching (3d-LSM). *Measurement Science and Technology*, 24(2):024001, 2012.
- A. Cahuzac, J. Boudet, P. Borgnat, and E. Lévêque. Smoothing algorithms for mean-flow extraction in large-eddy simulation of complex turbulent flows. *Phys. Fluids*, 22(12):125104, 2010.
- C. Canuto and D. Funaro. The Schwarz Algorithm for Spectral Methods. *SIAM J. Numer. Anal.*, 25(1):24–40, 1988.
- T. A. Casey, J. Sakakibara, and S. T. Thoroddsen. Scanning tomographic particle image velocimetry applied to a turbulent jet. *Physics of fluids*, 25(2):025102, 2013.
- P. Chandramouli, E. Memin, and D. Heitz. 4d Turbulent Wake Reconstruction using Large Eddy Simulation based Variational Data Assimilation. In *2nd Workshop on Data Assimilation & CFD Processing for Particle Image and Tracking Velocimetry*, Delft, Netherlands, 2017a.
- P. Chandramouli, E. Memin, D. Heitz, and S. Laizet. A Comparative Study of LES Models Under Location Uncertainty. In *Proceedings of the Congrès Français de Mécanique.*, 2017b.
- P. Chandramouli, E. Memin, D. Heitz, and S. Laizet. Coarse large-eddy simulations in a transitional wake flow with flow models under location uncertainty. *Comput. Fluids*, 168:170–189, 2018.
- C. Chemel and C. Staquet. A formulation of convective entrainment in terms of mixing efficiency. *J Fluid Mech.*, 580:169–178, 2007.

- C. Chemel, C. Staquet, and J.-P. Chollet. Estimating the diffusive heat flux across a stable interface forced by convective motions. *Nonlinear Processes Geophys.*, 17(2):187–200, 2010.
- L. Chen, S. W. Coleman, J. C. Vassilicos, and Z. Hu. Acceleration in turbulent channel flow. *J. Turbul.*, 11:N41, 2010.
- B. N. Cheng. A duality between forward and adjoint MPI communication routines. *Comput. Methods Sci. Technol.*, pages 23–24, 2006.
- S. I. Chernyshenko and M. F. Baig. The mechanism of streak formation in near-wall turbulence. *J. Fluid Mech.*, 544(-1):99–131, November 2005. ISSN 0022-1120, 1469-7645.
- M. Chevalier, J. Høpfner, T. R. Bewley, and D. S. Henningson. State estimation in wall-bounded flow systems. Part 2. Turbulent flows. *J. Fluid Mech.*, 552:167–187, 2006.
- T. M. Chin and A. J. Mariano. A particle filter for inverse Lagrangian prediction problems. *J. Atmos. Oceanic Tech.*, 27(2):371–384, 2010.
- P. Yo. Chou. On velocity correlations and the solutions of the equations of turbulent fluctuation. *Quart. Applied Mathematics*, 3(1):38–54, 1945.
- T. J. Chung. *Computational fluid dynamics*. Cambridge University Press, Cambridge ; New York, 2nd ed edition, 2010.
- C. K. Chyu and D. Rockwell. Near-wake structure of an oscillating cylinder: effect of controlled shear-layer vortices. *J. Fluid Mech.*, 322:21–49, 1996.
- C. H. Colburn, J. B. Cessna, and T. R. Bewley. State estimation in wall-bounded flow systems. Part 3. The ensemble Kalman filter. *J. Fluid Mech.*, 682:289–303, 2011.
- B. Combés, D. Heitz, A. Guibert, and E. Mémin. A particle filter to reconstruct a free-surface flow from a depth camera. *Fluid Dyn. Res.*, 47(5):051404, 2015.
- M. Cornick, B. Hunt, E. Ott, H. Kurtuldu, and M. F. Schatz. State and parameter estimation of spatiotemporally chaotic systems illustrated by an application to Rayleigh–Bénard convection. *Chaos*, 19(1):013108, 2009.
- T. Corpetti, D. Heitz, G. Arroyo, E. Mémin, and A. Santa-Cruz. Fluid experimental flow estimation based on an optical-flow scheme. *Exp. Fluids*, 40(1):80–97, 2006.
- C. J. Cotter, G. A. Gottwald, and D. D. Holm. Stochastic partial differential fluid equations as a diffusive limit of deterministic Lagrangian multi-time dynamics. *arXiv preprint arXiv:1706.00287*, 2017.
- P. Courtier and O. Talagrand. Variational assimilation of meteorological observations with the adjoint vorticity equation. II: Numerical results. *Q. J. R. Meteorolog. Soc.*, 113(478): 1329–1347, 1987.
- P. Courtier, J.-N. Thépaut, and A. Hollingsworth. A strategy for operational implementation of 4d-Var, using an incremental approach. *Q. J. R. Meteorolog. Soc.*, 120(519): 1367–1387, 1994.

- A. D. Craik and S. Leibovich. A rational model for Langmuir circulations. *J. Fluid Mech.*, 73(3):401–426, 1976.
- G. P. Cressman. An operational objective analysis system. *Mon. Weather Rev.*, 87(10):367–374, 1959.
- D. Crisan, F. Flandoli, and D. D. Holm. Solution properties of a 3d stochastic Euler fluid equation. *arXiv preprint arXiv:1704.06989*, 2017.
- A. Cuzol and E. Memin. A Stochastic Filtering Technique for Fluid Flow Velocity Fields Tracking. *IEEE Trans. Pattern Anal. Mach. Intell.*, 31(7):1278–1293, 2009.
- T. Cziesla, G. Biswas, and N.K. Mitra. Large eddy simulation in a turbulent channel flow using exit boundary conditions. *Int. J. Numer. Methods Fluids*, 30(6):763–773, 1999.
- J. D’adamo, N. Papadakis, E. Mémin, and G. Artana. Variational assimilation of POD low-order dynamical systems. *J. Turbul.*, 8:N9, 2007.
- R. Dolganov, B. Dubrulle, and J.-P. Laval. LES-Langevin Approach for Turbulent Channel Flow. In *Progress in Wall Turbulence: Understanding and Modeling*, pages 239–248. Springer, 2011.
- S. Dong and G.E. Karniadakis. DNS of flow past a stationary and oscillating cylinder at. *J. Fluids Struct.*, 20(4):519–531, 2005.
- B. Dubrulle, J. P. Laval, and S. Nazarenko. Langevin Models of Turbulence. In *Progress in Turbulence*, pages 77–86. Springer, 2005.
- F. Ducros, P. Comte, and M. Lesieur. Large-eddy simulation of transition to turbulence in a boundary layer developing spatially over a flat plate. *J. Fluid Mech.*, 326(-1):1, 1996.
- M. Eknes and G. Evensen. An ensemble Kalman filter with a 1-D marine ecosystem model. *J. Mar. Syst.*, 36(1-2):75–100, 2002.
- G. E. Elsinga, F. Scarano, B. Wieneke, and B. W. van Oudheusden. Tomographic particle image velocimetry. *Exp. Fluids*, 41(6):933–947, 2006.
- G. Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res. Oceans*, 99(C5):10143–10162, 1994.
- G. Evensen. *Data Assimilation*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- E. P. Fabry. 3d holographic PIV with a forward-scattering laser sheet and stereoscopic analysis. *Exp. Fluids*, 24(1):39–46, 1998.
- E. A. Fadlun, R. Verzicco, P. Orlandi, and J. Mohd-Yusof. Combined Immersed-Boundary Finite-Difference Methods for Three-Dimensional Complex Flow Simulations. *J. Comput. Phys.*, 161(1):35–60, 2000.
- C. Faure and P. Dutto. *Extension of Odyssee to the MPI library-Reverse mode*. PhD Thesis, INRIA, 1999.
- Y. Favennec, Y. Rouizi, and D. Petit. On the use of reduced models obtained through identification for feedback optimal control problems in a heat convection–diffusion problem. *Comput. Methods Appl. Mech. Eng.*, 199(17-20):1193–1201, 2010.

- B. Ferron and L. Hascoët. Capacites actuelles de la differentiation automatique: l'adjoint d'opa 9.0 par tapenade. In *In Colloque National sur l'Assimilation de Données*, Toulouse, France, 1996.
- E. J. Fertig, J. Harlim, and B. R. Hunt. A comparative study of 4d-VAR and a 4d Ensemble Kalman Filter: perfect model simulations with Lorenz-96. *Tellus A*, 59(1):96–100, 2007.
- D. P. G. Foures, N. Dovetta, D. Sipp, and P. J. Schmid. A data-assimilation method for Reynolds-averaged Navier–Stokes-driven mean flow reconstruction. *J. Fluid Mech.*, 759:404–431, 2014.
- J. Franke and W. Frank. Large eddy simulation of the flow past a circular cylinder at  $Re_D = 3900$ . *J. Wind Eng. Ind. Aerodyn.*, 90(10):1191–1206, 2002.
- N. Fujisawa, S. Tanahashi, and K. Srinivas. Evaluation of pressure field and fluid forces on a circular cylinder with and without rotational oscillation using velocity data from PIV measurement. *Meas. Sci. Technol.*, 16(4):989–996, 2005.
- C. Fureby and F. F. Grinstein. Monotonically Integrated Large Eddy Simulation of Free Shear Flows. *AIAA journal*, 37(5):544–556, 1999.
- B. Ganapathisubramani, K. Lakshminarasimhan, and N. T. Clemens. Investigation of three-dimensional structure of fine scales in a turbulent jet by using cinematographic stereoscopic particle image velocimetry. *Journal of fluid mechanics*, 598:141–175, 2008.
- L. S. Gandin. The Objective Analysis of Meteorological Fields (in Russian), *Gidrometeorologicheskoe Izdatel'stvo*, Leningrad, USSR, 1963. In *English translation, Israel Program for Scientific Translations*, Jerusalem, 1965.
- R. Gautier, S. Laizet, and E. Lamballais. A DNS study of jet control with microjets using an immersed boundary method. *Int. J. Comput. Fluid Dyn.*, 28(6-10):393–410, 2014.
- M. Germano. A proposal for a redefinition of the turbulent stresses in the filtered Navier–Stokes equations. *Phys. Fluids*, 29(7):2323, 1986.
- M. Germano, U. Piomelli, P. Moin, and W. H. Cabot. A dynamic subgrid-scale eddy viscosity model. *Phys. Fluids A*, 3(7):1760, 1991.
- S. Gesemann, F. Huhn, D. Schanz, and Andreas Schröder. From noisy particle tracks to velocity, acceleration and pressure fields using B-splines and penalties. In *18th international symposium on applications of laser and imaging techniques to fluid mechanics, Lisbon, Portugal*, pages 4–7, 2016.
- S. Ghosal, T. S. Lund, P. Moin, and K. Akselvoll. A dynamic localization model for large-eddy simulation of turbulent flows. *J. Fluid Mech.*, 286:229–255, 1995.
- D. Goldstein, R. Handler, and L. Sirovich. Modeling a no-slip flow boundary with an external force field. *J. Comput. Phys.*, 105(2):354–366, 1993.
- N. J. Gordon, D. J. Salmond, and A. F. M. Smith. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. *IEE Proceedings F - Radar and Signal Processing*, 140(2):107, 1993.
- C. Gray, C. A. Greated, D. R. McCluskey, and W. J. Easson. An analysis of the scanning beam PIV illumination system. *Meas. Sci. Technol.*, 2(8):717, 1991.

- R. Griesse and A. Walther. Evaluating gradients in optimal control: continuous adjoints versus automatic differentiation. *J. Optim. Theory. Appl.*, 122(1):63–86, 2004.
- A. Gronsksis, D. Heitz, and E. Mémin. Inflow and initial conditions for direct numerical simulation based on adjoint data assimilation. *J. Comput. Phys.*, 242:480–497, 2013.
- Y. Gu and D. S. Oliver. An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation. *SPE J.*, 12(04):438–446, 2007.
- J. Gullbrand and F. K. Chow. The effect of numerical errors and turbulence models in large-eddy simulations of channel flow, with and without explicit filtering. *J. Fluid Mech.*, 495:323–341, 2003.
- R. Hain, C. J. Kähler, and D. Michaelis. Tomographic and time resolved PIV measurements on a finite cylinder mounted on a flat plate. *Exp. Fluids*, 45(4):715–724, 2008.
- F. Hamba. A Hybrid RANS/LES Simulation of Turbulent Channel Flow. *Theor. Comput. Fluid Dyn.*, 16(5):387–403, 2003.
- J. Hamdi, H. Assoum, K. Abed-Meraïm, and A. Sakout. Volume reconstruction of an impinging jet obtained from stereoscopic-PIV data using POD. *European Journal of Mechanics-B/Fluids*, 67:433–445, 2018.
- T. M. Hamill and C. Snyder. A Hybrid Ensemble Kalman Filter–3d Variational Analysis Scheme. *Mon. Weather Rev.*, 128:15, 2000.
- R. R. Harcourt. An improved second-moment closure model of Langmuir turbulence. *J. Phys. Oceanogr.*, 45(1):84–103, 2015.
- J.-L. Harion, M. Favre-Marinet, and B. Camano. An improved method for measuring velocity and concentration by thermo-anemometry in turbulent helium-air mixtures. *Exp. Fluids*, 22(2):174–182, 1996.
- S. K. Harouna and E. Mémin. Stochastic representation of the Reynolds transport theorem: revisiting large-scale modeling. *Comput. Fluids*, 156:456–469, 2017.
- K. L. Harrison and D. G. Bogard. Comparison of RANS turbulence models for prediction of film cooling performance. In *ASME Turbo Expo 2008: Power for Land, Sea, and Air*, pages 1187–1196. American Society of Mechanical Engineers, 2008.
- L. Hascoët and V. Pascual. TAPENADE 2.1 user’s guide. Technical report, 2004.
- L. Hascoët and V. Pascual. The Tapenade Automatic Differentiation tool: principles, model, and specification. *ACM Trans. Math. Software*, 39(3):20, 2013.
- D. C. Haworth and S. B. Pope. A generalized Langevin model for turbulent flows. *Phys. Fluids*, 29(2):387–405, 1986.
- T. Hayase. Numerical simulation of real-world flows. *Fluid Dyn. Res.*, 47(5):051201, 2015.
- T. Hayase and S. Hayashi. State Estimator of Flow as an Integrated Computational Method With the Feedback of Online Experimental Measurement. *J. Fluids Eng*, 119(4):814–822, 1997.

- P. Héas, E. Mémin, D. Heitz, and P. D. Mininni. Power laws and inverse motion modelling: application to turbulence measurements from satellite images. *Tellus A*, 64(1):10962, 2012.
- D. Heitz. Coupling experimental and computational fluid dynamics: Synopsis of approaches, issues and perspectives. In *2nd Workshop on Data Assimilation & CFD Processing for PIV and Lagrangian Particle Tracking*, Delft, Netherlands, 2017.
- S. Hickel, N. A. Adams, and J. A. Domaradzki. An adaptive local deconvolution method for implicit LES. *J. Comput. Phys.*, 213(1):413–436, 2006.
- J. Höpffner, M. Chevalier, T. R. Bewley, and D. S. Henningson. State estimation in wall-bounded flow systems. Part 1. Perturbed laminar flows. *J. Fluid Mech.*, 534:263–294, 2005.
- M. Högberg, T. R. Bewley, and D. S. Henningson. Linear feedback control and estimation of transition in plane channel flow. *J. Fluid Mech.*, 481:149–175, 2003.
- D. D. Holm. The ideal craik-leibovich equations. *Physica D*, 98(2-4):415–441, 1996.
- D. D. Holm. Variational principles for stochastic fluid dynamics. *Proc. R. Soc. London, Ser. A*, 471(2176), 2015.
- J. A. C. Humphrey and F. Pourahmadi. Prediction of curved channel flow with an extended k-epsilon model of turbulence. *AIAA journal*, 21(10):1365–1373, 1983.
- B. R. Hunt, E. Kalnay, E. J. Kostelich, E. Ott, D. J. Patil, T. Sauer, I. Szunyogh, J. A. Yorke, and A. V. Zimin. Four-dimensional ensemble Kalman filtering. *Tellus A*, 56(4):273–277, 2004.
- B. R. Hunt, E. J. Kostelich, and I. Szunyogh. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, 230(1-2):112–126, 2007.
- Z. Jiang, Z. Xia, Y. Shi, and S. Chen. Large-eddy simulation of plane channel flow with Vreman’s model. *J. Turbul.*, 17(8):807–822, 2016.
- J. Jiménez. Coherent structures in wall-bounded turbulence. *J. Fluid Mech.*, 842, May 2018. ISSN 0022-1120, 1469-7645. doi: 10.1017/jfm.2018.144.
- W. P. Jones and B. E. Launder. The prediction of laminarization with a two-equation model of turbulence. *Int. J. Heat Mass Transfer*, 15(2):301–314, 1972.
- C. J. Kähler and J. Kompenhans. Fundamentals of multiple plane stereo particle image velocimetry. *Experiments in Fluids*, 29(1):S070–S077, 2000.
- R. E. Kalman. A new approach to linear filtering and prediction problems. *J. Basic Eng.*, 82(1):35–45, 1960.
- H. Kato, A. Yoshizawa, G. Ueno, and S. Obayashi. A data assimilation methodology for reconstructing turbulent flows around aircraft. *J. Comput. Phys.*, 283:559–581, 2015.
- A. Kempf, F. Flemming, and J. Janicka. Investigation of lengthscales, scalar dissipation, and flame orientation in a piloted diffusion flame by LES. *Proc. Combust. Inst.*, 30(1):557–565, January 2005.

- C. A. Kennedy, M. H. Carpenter, and R. M. Lewis. Low-storage, explicit Runge–Kutta schemes for the compressible Navier–Stokes equations. *Appl. Numer. Math.*, 35(3):177–219, 2000.
- J. C. Kent, N. Trigui, W. C. Choi, Y. G. Guezennec, and R. S. Brodkey. Photogrammetric calibration for improved three-dimensional particle-tracking velocimetry (3d PTV). In *Optical Diagnostics in Fluid and Thermal Flow*, volume 2005, pages 400–413. International Society for Optics and Photonics, 1993.
- J. Kim, P. Moin, and R. Moser. Turbulence statistics in fully developed channel flow at low Reynolds number. *J. Fluid Mech.*, 177(-1):133, 1987.
- S. K. Kim and S. K. Chung. An investigation on airflow in disordered nasal cavity and its corrected models by tomographic PIV. *Meas. Sci. Technol.*, 15(6):1090–1096, 2004.
- E. Kit, O. Krivonosova, D. Zhilenko, and D. Friedman. Reconstruction of large coherent structures from SPIV measurements in a forced turbulent mixing layer. *Exp. Fluids*, 39(4):761–770, 2005.
- A. N. Kolmogorov. Equations of motion of an incompressible turbulent fluid. *Izv Akad Nauk SSSR Ser Phys*, 6(6):56–58, 1942.
- R. H. Kraichnan. The structure of isotropic turbulence at very high Reynolds numbers. *J. Fluid Mech.*, 5(04):497–543, 1959.
- A. G. Kravchenko and P. Moin. Numerical studies of flow over a circular cylinder at  $Re_D=3900$ . *Phys. Fluids*, 12(2):403, 2000.
- H. Kunita. *Stochastic flows and stochastic differential equations.*, volume 24. Cambridge University Press, 1997.
- S. Laizet and E. Lamballais. High-order compact schemes for incompressible flows: A simple and efficient method with quasi-spectral accuracy. *J. Comput. Phys.*, 228(16):5989–6015, 2009.
- S. Laizet and N. Li. Incompact3d: A powerful tool to tackle turbulence problems with up to  $O(10^5)$  computational cores. *Int. J. Numer. Methods Fluids*, 67(11):1735–1757, 2011.
- E. Lamballais, J. Silvestrini, and S. Laizet. Direct numerical simulation of a separation bubble on a rounded finite-width leading edge. *Int. J. Heat Fluid Fl.*, 29(3):612–625, 2008.
- E. Lamballais, J. Silvestrini, and S. Laizet. Direct numerical simulation of flow separation behind a rounded leading edge: Study of curvature effects. *Int. J. Heat Fluid Fl.*, 31(3):295–306, 2010.
- E. Lamballais, V. Fortuné, and S. Laizet. Straightforward high-order numerical dissipation via the viscous term for direct and large eddy simulation. *J. Comput. Phys.*, 230(9):3270–3275, 2011.
- I Langmuir. Surface motion of water induced by wind. *Science*, 87(2250):119–124, 1938.
- E. Lauga and T. R. Bewley. Performance of a linear robust control strategy on a nonlinear model of spatially developing flows. *J. Fluid Mech.*, 512:343–374, 2004.

- B. E. Launder and B. I. Sharma. Application of the energy-dissipation model of turbulence to the calculation of flow near a spinning disc. *Lett. Heat Mass Trans.*, 1:131–138, 1974.
- B. E. Launder, G. J. Reece, and W. Rodi. Progress in the development of a Reynolds-stress turbulence closure. *J. Fluid Mech.*, 68(03):537, 1975.
- J.-P. Laval and B. Dubrulle. A LES-Langevin model for turbulence. *Eur. Phys. J. B - Condensed Matter and Complex Systems*, 49(4):471–481, 2006.
- J.-P. Laval, B. Dubrulle, and J. C. McWilliams. Langevin models of turbulence: Renormalization group, distant interaction algorithms or rapid distortion theory? *Phys. Fluids*, 15(5):1327–1339, 2003.
- K. Law, A. Stuart, and K. Zygalakis. *Data Assimilation*, volume 62 of *Texts in Applied Mathematics*. Springer International Publishing, Cham, 2015.
- W. Layton and R. Lewandowski. A simple and stable scale-similarity model for large eddy simulation: energy balance and existence of weak solutions. *Appl Math Lett*, 16(8):1205–1209, 2003.
- W. Layton and R. Lewandowski. On a well-posed turbulence model. *Discrete Continuous Dyn Syst Ser B*, 6(1):111, 2006a.
- W. Layton and R. Lewandowski. Residual stress of approximate deconvolution models of turbulence. *J. Turbul.*, 7:N46, 2006b.
- F.-X. Le Dimet and O. Talagrand. Variational algorithms for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus A*, 38(2):97–110, 1986.
- J. Lee, H. Choi, and N. Park. Dynamic global model for large eddy simulation of transient flow. *Phys. Fluids*, 22(7):075106, 2010.
- M. Lee and R. D. Moser. Direct numerical simulation of turbulent channel flow up to  $Re\tau \sim 5200$ . *J. Fluid Mech.*, 774:395–415, 2015.
- P. J. V. Leeuwen, Y. Cheng, and S. Reich. *Nonlinear Data Assimilation*. Frontiers in Applied Dynamical Systems: Reviews and Tutorials. Springer International Publishing, 2015.
- S. Leibovich. On wave-current interaction theories of Langmuir circulations. *J. Fluid Mech.*, 99(4):715–724, 1980.
- S. Leibovich. The form and Dynamics of Langmuir Circulations. *Annu. Rev. Fluid Mech.*, 15(1):391–427, January 1983. ISSN 0066-4189, 1545-4479.
- S. K. Lele. Compact finite difference schemes with spectral-like resolution. *J. Comput. Phys.*, 103(1):16–42, 1992.
- M. Lemke and J. Sesterhenn. Adjoint-based pressure determination from PIV data in compressible flows—validation and assessment based on synthetic data. *Eur. J. Mech. B. Fluids*, 58:29–38, 2016.
- A. Leonard. Energy cascade in large-eddy simulations of turbulent fluid flows. *Adv. Geophys.*, 18:237–248, 1975.

- R. Leroux and L. Cordier. Dynamic mode decomposition for non-uniformly sampled data. *Exp. Fluids*, 57(5):94, 2016.
- M. Lesieur. *Turbulence in fluids*. Number 84 in Fluid Mechanics and its Applications. Springer, Dordrecht, 4. rev. and enlarged ed edition, 2008. OCLC: 255895179.
- M. Lesieur, C. Staquet, P. Le Roy, and P. Comte. The mixing layer and its coherence examined from the point of view of two-dimensional turbulence. *J Fluid Mech.*, 192: 511–534, 1988.
- M. Lesieur, P. Comte, Y. Dubief, E. Lamballais, O. Métais, and S. Ossia. From two-point closures of isotropic turbulence to LES of shear flows. *Flow Turbul. Combust.*, 63(1-4): 247–267, 2000.
- R. Lewandowski. Vorticities in a LES model for 3d periodic turbulent flows. *J. Math. Fluid Mech.*, 8(3):398–422, 2006.
- N. Li and S. Laizet. 2decomp&FFT – A Highly Scalable 2d Decomposition Library and FFT Interface, 2010.
- Z. Li and I. M. Navon. Optimality of variational data assimilation and its relationship with the Kalman filter and smoother. *Q. J. R. Meteorolog. Soc.*, 125(572):661–683, 2001.
- D. K. Lilly. The representation of small scale turbulence in numerical simulation experiments. In *Proceedings of the IBM Scientific Computing Symposium on Environmental Sciences.*, Yorktown Heights, NY, 1967.
- D. K. Lilly. A proposed modification of the Germano subgrid-scale closure method. *Phys. Fluids A*, 4(3):633–635, 1992.
- J. Lions. Optimal control of systems governed by partial differential equations problèmes aux limites. *Berlin: Springer*, page 407, 1971.
- A. C. Lorenc. The potential of the ensemble Kalman filter for NWP—a comparison with 4d-Var. *Q. J. R. Meteorolog. Soc.*, 129(595):3183–3203, 2003.
- L. Lourenco and A. Krothapalli. The role of photographic parameters in laser Speckle of particle image displacement velocimetry. *Exp. Fluids*, 5(1):29–32, 1986.
- L. M. Lourenco and C. Shih. Characteristics of the plane turbulent near wake of a circular cylinder. *A Particle Image Velocimetry Study.*, 1993.
- H. Lübecke, St. Schmidt, T. Rung, and F. Thiele. Comparison of LES and RANS in bluff-body flows. *J. Wind Eng. Ind. Aerodyn.*, 89(14):1471–1485, 2001.
- D. G. Luenberger. Observing the state of a linear system. *IEEE Transactions on Military Electronics*, 8(2):74–80, 1964.
- T. S. Lund. The use of explicit filters in large eddy simulation. *Comput. Math. Appl.*, 46(4):603–616, 2003.
- T. S. Lund and H. J. Kaltenbach. Experiments with explicit filtering for LES using a finite-difference method. *Annual Research Briefs (CTR)*, pages 91–105, 1995.
- P. Lynch. The origins of computer weather prediction and climate modeling. *J. Comput. Phys.*, 227(7):3431–3444, 2008.

- X. Ma, G.S. Karamanos, and G. E. Karniadakis. Dynamics and low-dimensionality of a turbulent near wake. *J. Fluid Mech.*, 410:29–65, 2000.
- J. M. MacInnes and F. V. Bracco. Stochastic particle dispersion modeling and the tracer-particle limit. *Phys. Fluids A*, 4(12):2809, 1992.
- L. G. Margolin and W. J. Rider. The design and construction of implicit LES models. *Int. J. Numer. Methods Fluids*, 47(10-11):1173–1179, 2005.
- L. G. Margolin, W. J. Rider, and F. F. Grinstein. Modeling turbulent flow with implicit LES. *J. Turbul.*, 7, 2006.
- R. Martinuzzi and A. Pollard. Comparative study of turbulence models in predicting turbulent pipe flow. I - Algebraic stress and k-epsilon models. *AIAA journal*, 27(1): 29–36, 1989.
- P. J. Mason and D. J. Thomson. Stochastic backscatter in large-eddy simulations of boundary layers. *J. Fluid Mech.*, 242(-1):51, 1992.
- J. C. McWilliams, P. P. Sullivan, and C.-H. Moeng. Langmuir turbulence in the ocean. *J. Fluid Mech.*, 334:1–30, 1997.
- M. Meldi and A. Poux. A reduced order model based on Kalman filtering for sequential data assimilation of turbulent flows. *J. Comput. Phys.*, 347:207–234, 2017.
- E. Mémin. Fluid flow dynamics under location uncertainty. *Geophys. Astro. Fluid*, 108(2): 119–146, 2014.
- C. Meneveau, T. S. Lund, and W. H. Cabot. A Lagrangian dynamic subgrid-scale model of turbulence. *J. Fluid Mech.*, 319(-1):353, 1996.
- E. R. Menter. Zonal Two Equation  $k-\omega$ , Turbulence Models for Aerodynamic Flows. Orlando, Florida, 1993.
- F. R. Menter. Two-equation eddy-viscosity turbulence models for engineering applications. *AIAA journal*, 32(8):1598–1605, 1994.
- P. Mercier and M. Deville. A multidimensional compact higher-order scheme for 3-D Poisson’s equation. *J. Comput. Phys.*, 39(2):443–455, 1981.
- O. Métais and M. Lesieur. Spectral large-eddy simulation of isotropic and stably stratified turbulence. *J. Fluid Mech.*, 239(-1):157, 1992.
- J. Meyers and P. Sagaut. On the model coefficients for the standard and the variational multi-scale Smagorinsky model. *J. Fluid Mech.*, 569:287, 2006.
- J.-P. Minier. On Lagrangian stochastic methods for turbulent polydisperse two-phase reactive flows. *Prog. Energy Combust. Sci.*, 50:1–62, 2015.
- T. Misaka, T. Ogasawara, S. Obayashi, I. Yamada, and Y. Okuno. Assimilation Experiment of Lidar Measurements for Wake Turbulence. *J. Fluid Sci. Technol.*, 3(4):512–518, 2008.
- E. A. Misawa and J. K. Hedrick. Nonlinear observers - a state-of-the-art survey. *J. Dyn. Syst. Meas. Contr.*, 111(3):344–352, 1989.

- R. Mittal and P. Moin. Suitability of Upwind-Biased Finite Difference Schemes for Large-Eddy Simulation of Turbulent Flows. *AIAA journal*, 35(8):1415–1417, 1997.
- P. Moin and J. Kim. Numerical investigation of turbulent channel flow. *J. Fluid Mech.*, 118(-1):341, 1982.
- V. Mons, J.-C. Chassaing, T. Gomez, and P. Sagaut. Is isotropic turbulence decay governed by asymptotic behavior of large scales? An eddy-damped quasi-normal Markovian-based data assimilation study. *Phys. Fluids*, 26(11):115105, 2014.
- V. Mons, J.-C. Chassaing, T. Gomez, and P. Sagaut. Reconstruction of unsteady viscous flows using data assimilation schemes. *J. Comput. Phys.*, 316:255–280, 2016.
- R. D. Moser, J. Kim, and N. N. Mansour. Direct numerical simulation of turbulent channel flow up to  $Re_\tau = 590$ . *Phys. Fluids*, 11(4):943–945, 1999.
- S. Nadarajah and A. Jameson. A comparison of the continuous and discrete adjoint approach to automatic aerodynamic optimization. In *38th Aerospace Sciences Meeting and Exhibit*, page 667, 2000.
- M. Ndoye, J. Delville, D. Heitz, and G. Arroyo. Parameterizable constant temperature anemometer: a new method for the analysis of velocity-temperature coupling in turbulent heat transfer. *Meas. Sci. Technol.*, 21(7):075401, 2010.
- F. Nicoud and F. Ducros. Subgrid-scale stress modelling based on the square of the velocity gradient tensor. *Flow Turbul. Combust.*, 62(3):183–200, 1999.
- J. Nocedal. Updating quasi-Newton matrices with limited storage. *Math. Comput.*, 35(151):773–782, 1980.
- C. Norberg. LDV measurements in the near wake of a circular cylinder. In *Proceedings of the Conference on Bluff Body Wakes and Vortex-Induced Vibration*, pages 1–12, Washington, D.C., 1998. Cornell University, Ithaca, NY, 1998.
- N. M. Nouri, S. Sekhvat, and A. Mofidi. Drag reduction in a turbulent channel flow with hydrophobic wall. *J. Hydrodyn. Ser. B*, 24(3):458–466, 2012.
- T. Okuno, Y. Sugii, and S. Nishio. Image measurement of flow field using physics-based dynamic model. *Meas. Sci. Technol.*, 11(6):667, 2000.
- L. Ong and J. Wallace. The velocity field of the turbulent very near wake of a circular cylinder. *Exp. Fluids*, 20(6):441–453, 1996.
- E. S. Oran and J. P. Boris. Computing Turbulent Shear Flows — A Convenient Conspiracy. *Comput. Phys.*, 7(5):523, 1993.
- H. Ouvrard, B. Koobus, A. Dervieux, and M. V. Salvetti. Classical and variational multi-scale LES of the flow around a circular cylinder on unstructured grids. *Comput. Fluids*, 39(7):1083–1094, 2010.
- N. Papadakis and E. Mémin. Variational Assimilation of Fluid Motion from Image Sequence. *SIAM J. Imag. Sci.*, 1(4):343–363, 2008.
- N. Papadakis, E. Mémin, A. Cuzol, and N. Gengembre. Data assimilation with the weighted ensemble Kalman filter. *Tellus A*, 62(5):673–697, 2010.

- H. M. Park and W. J. Lee. Feedback control of natural convection. *Comput. Methods Appl. Mech. Eng.*, 191(8-10):1013–1028, 2001.
- N. Park, S. Lee, J. Lee, and H. Choi. A dynamic subgrid-scale eddy viscosity model with a global model coefficient. *Phys. Fluids*, 18(12):125109, 2006.
- P. Parnaudeau, J. Carlier, D. Heitz, and E. Lamballais. Experimental and numerical studies of the flow over a circular cylinder at Reynolds number 3900. *Phys. Fluids*, 20(8):085101, 2008.
- R. Pasquetti. Spectral vanishing viscosity method for LES: sensitivity to the SVV control parameters. *J. Turbul.*, 6:N12, 2005.
- O.M. Phillips. *The Dynamics of the Upper Ocean*. Cambridge University Press, 2nd edition edition, 1977.
- L. C. Pinto, E. B. C. Schettini, and J. H. Silvestrini. Numerical analysis of the immersed boundary method applied to the flow around a forced oscillating cylinder. *J. Phys. Conf. Ser.*, 296:012011, 2011.
- A. Pollard and R. Martinuzzi. Comparative study of turbulence models in predicting turbulent pipe flow. II - Reynolds stress and k-epsilon models. *AIAA journal*, 27(12):1714–1721, 1989.
- S. Pope. *Turbulent Flows*. Cambridge University Press, Cambridge, 2000.
- S. B. Pope. Lagrangian PDF methods for turbulent flows. *Annu. Rev. Fluid Mech.*, 26(1):23–63, 1994.
- L. Prandtl. Bericht über untersuchungen zur ausgebildeten turbulenz. *Zs. angew. Math. Mech.*, 5:136–139., 1925.
- L. Prandtl. Über ein neues Formelsystem für die ausgebildete Turbulenz. In *Nachr. Alad. Wiss. Göttingen*, pages 6–19, Göttingen, 1945.
- T. H. Pulliam and D. W. Zingg. *Fundamental Algorithms in Computational Fluid Dynamics*. Scientific Computation. Springer International Publishing, Cham, 2014.
- S. Rajagopalan and R. A. Antonia. Flow around a circular cylinder-structure of the near wake shear layer. *Exp. Fluids*, 38(4):393–402, 2005.
- T. C. Rebollo and R. Lewandowski. *Mathematical and Numerical Foundations of Turbulence Models and Applications*. Modeling and Simulation in Science, Engineering and Technology. Springer New York, New York, NY, 2014.
- S. Reich and C. Cotter. *Probabilistic forecasting and bayesian data assimilation | Computational science*. Cambridge University Press, 2015.
- V. Resseguier. *Mixing and fluid dynamics under location uncertainty*. PhD Thesis, Université de Rennes 1, 2017.
- V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, Part II Quasi-geostrophy and efficient ensemble spreading. *Geophys. Astro. Fluid*, 111(3):177–208, 2017a.

- V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, Part III SQG and frontal dynamics under strong turbulence conditions. *Geophys. Astro. Fluid*, 111(3):209–227, 2017b.
- V. Resseguier, E. Mémin, and B. Chapron. Geophysical flows under location uncertainty, Part I Random transport and general models. *Geophys. Astro. Fluid*, 111(3):149–176, 2017c.
- V. Resseguier, E. Mémin, D. Heitz, and B. Chapron. Stochastic modelling and diffusion modes for proper orthogonal decomposition models and small-scale flow analysis. *J. Fluid Mech.*, 826:888–917, 2017d.
- R. D. Richtmyer. Proposed numerical method for calculation of shocks. LANL Report 671, LA, 1948.
- C. Robinson. *Image assimilation techniques for Large Eddy Scale models : Application to 3D reconstruction*. Scientific, Université de Rennes 1, Rennes, 2015.
- J. Rotta. Statistische Theorie nichthomogener Turbulenz. *Zeitschrift für Physik*, 131:51–77, 1951.
- P. Sagaut. *Large eddy simulation for incompressible flows: an introduction*. Scientific computation. Springer, Berlin ; New York, 3rd ed edition, 2006.
- Y. Sasaki. An Objective Analysis Based on the Variational Method. *J. Meteorolog. Soc. Jpn. Ser. II*, 36(3):77–88, 1958.
- Y. Sasaki. Some basic formalisms in numerical variational analysis. *Mon. Weather Rev.*, 98(12):875–883, 1970.
- B. L. Sawford. Generalized random forcing in random-walk turbulent dispersion models. *Phys. Fluids*, 29(11):3582–3585, 1986.
- F. Scarano. Tomographic PIV: principles and practice. *Measurement Science and Technology*, 24(1):012001, 2012.
- F. Scarano and P. Moore. An advection-based model to increase the temporal resolution of PIV time series. *Exp. Fluids*, 52(4):919–933, 2012.
- M. Schanen, U. Naumann, L. Hascoët, and J. Utke. Interpretative adjoints for numerical simulation codes using MPI. *Procedia Comput. Sci.*, 1(1):1825–1833, 2010.
- D. Schanz, S. Gesemann, and A. Schröder. Shake-The-Box: Lagrangian particle tracking at high particle image densities. *Exp. Fluids*, 57(5), 2016.
- Peter J. Schmid. Dynamic mode decomposition of numerical and experimental data. *J. fluid mech.*, 656:5–28, 2010.
- F. G. Schmitt. About Boussinesq’s turbulent viscosity hypothesis: historical remarks and a direct evaluation of its validity. *Comptes Rendus Mécanique*, 335(9-10):617–627, 2007.
- J. Schneiders, G. C. A. Caridi, A. Sciacchitano, and F. Scarano. Instantaneous Pressure Measurements from Large-Scale Tomo-PTV with HFSB Tracers past a Surface-Mounted Finite Cylinder. American Institute of Aeronautics and Astronautics, 2016.

- J. F. G. Schneiders and F. Scarano. Dense velocity reconstruction from tomographic PTV with material derivatives. *Exp. Fluids*, 57(9), 2016.
- J. F. G. Schneiders, R. P. Dwight, and F. Scarano. Time-supersampling of 3d-PIV measurements with vortex-in-cell simulation. *Exp. Fluids*, 55(3):1692, 2014.
- U. Schumann. Stochastic backscatter of turbulence energy and scalar variance by random subgrid-scale fluxes. *Proc. R. Soc. Lond. A*, 451(1941):293–318, 1995.
- A. S. Sharma and B. J. McKeon. On coherent structure in wall turbulence. *J. Fluid Mech.*, 728:196–238, 2013.
- T. H. Shih, N. S. Liu, and K. H. Chen. A non-linear k-epsilon model for turbulent shear flows. In *34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, page 3983, 1998.
- J. H. Silvestrini and E. Lamballais. Direct Numerical Simulation of Wakes with Virtual Cylinders. *Int. J. Comput. Fluid Dyn.*, 16(4):305–314, 2002.
- J. H. Silvestrini and E. Lamballais. Direct numerical simulation of oblique vortex shedding from a cylinder in shear flow. *Int. J. Heat Fluid Fl.*, 25(3):461–470, 2004.
- S. P. Singh and S. Mittal. Energy spectra of flow past a circular cylinder. *Int. J. Comput. Fluid Dyn.*, 18(8):671–679, 2004.
- A. Sinha. *Development of reduced-order models and strategies for feedback control of high-speed axisymmetric jets*. PhD Thesis, The Ohio State University, 2011.
- L. Sirovich. Turbulence and the dynamics of coherent structures. I. Coherent structures. *Q. Appl. Math.*, 45(3):561–571, 1987.
- J. Smagorinsky. General circulation experiments with the primitive equations. *Mon. Weather Rev.*, 91(3):99–164, 1963.
- C. R. Smith and J. D. Walker. Turbulent wall-layer vortices. In *Fluid vortices*, pages 235–289. Springer, 1995.
- K. Sodjavi and J. Carlier. Experimental study of thermal mixing layer using variable temperature hot-wire anemometry. *Exp. Fluids*, 54(10), 2013.
- H. W. Sorenson. *Kalman filtering: theory and application*. IEEE, 1985.
- J. Soria and C. Atkinson. Towards 3c-3d digital holographic fluid velocity vector field measurement—tomographic digital holographic PIV (Tomo-HPIV). *Meas. Sci. Technol.*, 19(7):074002, 2008.
- P. Spalart and S. Allmaras. A one-equation turbulence model for aerodynamic flows. American Institute of Aeronautics and Astronautics, 1992.
- P. R. Spalart. Comments on the feasibility of LES for wings, and on a hybrid RANS/LES approach. In *Proceedings of First AFOSR International Conference on DNS/LES*, 1997.
- P. R. Spalart. Detached-Eddy Simulation. *Annu. Rev. Fluid Mech.*, 41(1):181–202, 2009.

- C. G. Speziale, S. Sarkar, and T. B. Gatski. Modelling the pressure–strain correlation of turbulence: an invariant dynamical systems approach. *J. Fluid Mech.*, 227:245–272, 1991.
- P. K. Stansby. The effects of end plates on the base pressure coefficient of a circular cylinder. *Aeronaut. J.*, 78(757):36–37, 1974.
- A. M. Steinberg, J. F. Driscoll, and S. L. Ceccio. Three-dimensional temporally resolved measurements of turbulence–flame interactions using orthogonal-plane cinematographic PIV. *Experiments in fluids*, 47(3):527–547, 2009.
- M. Strelets. Detached eddy simulation of massively separated flows. In *39th Aerospace Sciences Meeting and Exhibit*, page 879, 2001.
- J. T. Stuart. The production of intense shear layers by vortex stretching and convection. Technical report, Advisory group for aerospace research and development Paris (France), 1965.
- N. Sugiura, T. Awaji, S. Masuda, T. Mochizuki, T. Toyoda, T. Miyama, H. Igarashi, and Y. Ishikawa. Development of a four-dimensional variational coupled data assimilation system for enhanced analysis and prediction of seasonal to interannual climate variations. *J. Geophys. Res. Oceans*, 113(C10), 2008.
- T. Suzuki, H. Ji, and F. Yamamoto. Unsteady PTV velocity field past an airfoil solved with DNS: Part 1. Algorithm of hybrid simulation and hybrid velocity field at  $Re = 103$ . *Exp. Fluids*, 47(6):957–976, 2009a.
- T. Suzuki, A. Sanse, T. Mizushima, and F. Yamamoto. Unsteady PTV velocity field past an airfoil solved with DNS: Part 2. Validation and application at Reynolds numbers up to  $Re \leq 104$ . *Exp. Fluids*, 47:977–994, 2009b.
- K. Takehara, R. J. Adrian, G. T. Etoh, and K. T. Christensen. A Kalman tracker for super-resolution PIV. *Exp. Fluids*, 29(1):S034–S041, 2000.
- M. H. Tber, L. Hascoet, A. Vidard, and B. Dauvergne. Building the tangent and adjoint codes of the ocean general circulation model OPA with the automatic differentiation tool tapenade. *arXiv preprint arXiv:0711.4444*, 2007.
- J. N. Thépaut, R. N. Hoffman, and P. Courtier. Interactions of dynamics and observations in a four-dimensional variational assimilation. *Mon. Weather Rev.*, 121(12):3393–3414, 1993.
- G. Tissot, L. Cordier, N. Benard, and B. R. Noack. Dynamic mode decomposition of PIV measurements for cylinder wake flow in turbulent regime. In *TSPF Digital Library Online*. Begel House Inc., 2013.
- G. Tissot, L. Cordier, N. Benard, and B. R. Noack. Model reduction using dynamic mode decomposition. *C. R. Mécanique*, 342(6-7):410–416, 2014.
- A. A. Townsend. Momentum and Energy Diffusion in the Turbulent Wake of a Cylinder. *Proc. R. Soc. London, Ser. A*, 197(1048):124–140, 1949a.
- A. A. Townsend. The fully developed wake of a circular cylinder. *Aust. J. Chem.*, 2(4):451–468, 1949b.

- A. Travin, M. Shur, M. Strelets, and P. Spalart. Detached-eddy simulations past a circular cylinder. *Flow Turbul. Combust.*, 63(1-4):293–313, 2000.
- M. Uchiyama and K. Hakomori. Measurement of instantaneous flow rate through estimation of velocity profiles. *IEEE Trans. Autom. Control*, 28(3):380–388, 1983.
- J. Utke, L. Hascoet, P. Heimbach, C. Hill, P. Hovland, and U. Naumann. Toward adjoinable MPI. In *Parallel & Distributed Processing, 2009. IPDPS 2009. IEEE International Symposium on*, pages 1–8. IEEE, 2009.
- J. Utke, L. Hascoët, V. Pascual, P. Heimbach, C. Hill, U. Naumann, and M. Schanen. Gradient of MPI-parallel codes, 2012.
- J. Van Bladel. On Helmholtz’s Theorem in Finite Regions. *IRE Transaction on Antennas and Propagation*, 1958.
- H. van der Ven. A family of large eddy simulation (LES) filters with nonuniform filter widths. *Phys. Fluids*, 7(5):1171–1172, 1995.
- E. R. Van Driest. The problem of aerodynamic heating. *Aeronaut. Eng. Rev.*, 15(10):26–41, 1956.
- M.R. Visbal, P.E. Morgan, and D.P. Rizzetta. An Implicit LES approach based on high-order compact differencing and filtering schemes (Invited). Orlando, Florida, 2003.
- S. Völker, R. D. Moser, and P. Venugopal. Optimal large eddy simulation of turbulent channel flow based on direct numerical simulation statistical data. *Phys. Fluids*, 14(10):3675–3691, 2002.
- A. W. Vreman. An eddy-viscosity subgrid-scale model for turbulent shear flow: Algebraic theory and applications. *Phys. Fluids*, 16(10):3670–3681, 2004.
- M. Weickert, G. Teike, O. Schmidt, and M. Sommerfeld. Investigation of the LES WALE turbulence model within the lattice Boltzmann framework. *Comput. Math. Appl.*, 59(7):2200–2214, 2010.
- J. Weinkauff, D. Michaelis, A. Dreizler, and B. Böhm. Tomographic PIV measurements in a turbulent lifted jet flame. *Exp. Fluids*, 54(12), 2013.
- J. Westerweel. Efficient detection of spurious vectors in particle image velocimetry data. *Exp. Fluids*, 16(3-4):236–247, 1994.
- J. Westerweel, G. E. Elsinga, and R. J. Adrian. Particle Image Velocimetry for Complex and Turbulent Flows. *Annu. Rev. Fluid Mech.*, 45(1):409–436, 2013.
- B. Wieneke. Iterative reconstruction of volumetric particle distribution. *Meas. Sci. Technol.*, 24(2):024008, 2012.
- D. C. Wilcox. Reassessment of the scale-determining equation for advanced turbulence models. *AIAA journal*, 26(11):1299–1310, 1988.
- D. C. Wilcox. Comparison of two-equation turbulence models for boundary layers with pressure gradient. *AIAA journal*, 31(8):1414–1421, 1993.
- R. B. Wilhelmson and J. H. Ericksen. Direct solutions for Poisson’s equation in three dimensions. *J. Comput. Phys.*, 25(4):319–331, 1977.

- C. E. Willert and M. Gharib. Digital particle image velocimetry. *Exp. Fluids*, 10(4):181–193, 1991.
- C. H. K. Williamson. Vortex dynamics in the cylinder wake. *Annu. Rev. Fluid Mech.*, 28:477–539, 1996.
- J. H. Williamson. Low storage Runge-Kutta schemes. *J. Comput. Phys.*, 35:48–56, 1980.
- G. S. Winckelmans, A. A. Wray, O. V. Vasilyev, and H. Jeanmart. Explicit-filtering large-eddy simulation using the tensor-diffusivity model supplemented by a dynamic Smagorinsky term. *Phys. Fluids*, 13(5):1385–1403, 2001.
- A. Wray and J. Hunt. Algorithms for classification of turbulent structures. *Topological Fluid Mechanics: Proceedings of the IUTAM Symposium*, pages 95–104, 1990.
- K. Yagoh, K. Ogawara, and S. Iida. Application of Kalman filter to three-dimensional flow measuring system based on digital image processing. *JSME Int. J. Ser. B Fluids Therm. Eng.*, 36(2):285–293, 1993.
- V. Yakhot and S. A. Orszag. Renormalization group analysis of turbulence. I. Basic theory. *J. Sci. Comput.*, 1(1):3–51, 1986.
- V. Yakhot, S. A. Orszag, S. Thangam, T. B. Gatski, and C. G. Speziale. Development of turbulence models for shear flows by a double expansion technique. *Phys. Fluids A*, 4(7):1510–1520, 1992.
- Y. Yang. *Study of Variational Ensemble Methods for Image Assimilation*. PhD Thesis, Université de Rennes 1, Rennes, 2014.
- Y. Yang, C. Robinson, D. Heitz, and E. Mémin. Enhanced ensemble-based 4dvar scheme for data assimilation. *Comput. Fluids*, 115:201–210, 2015.
- Z. Yang and T. H. Shih. New time scale based k-epsilon model for near-wall turbulence. *AIAA journal*, 31(7):1191–1198, 1993.
- C. H. Yau, A. K. Bajaj, and O. D. I. Nwokah. Active control of chaotic vibration in a constrained flexible pipe conveying fluid. *J. Fluids Struct.*, 9(1):99–122, 1995.
- D. You and P. Moin. Large-eddy simulation of flow separation over an airfoil with synthetic jet control. *Annual Research Briefs (CTR)*, pages 337–346, 2006.
- W. Zhang, R. Hain, and C. J. Kähler. Scanning PIV investigation of the laminar separation bubble on a SD7003 airfoil. *Experiments in Fluids*, 45(4):725–743, 2008.
- L. Zhao, C. Marchioli, and H. I. Andersson. Slip velocity of rigid fibers in turbulent channel flow. *Phys. Fluids*, 26(6):063302, 2014.
- D. Županski and F. Mesinger. Four-dimensional variational assimilation of precipitation data. *Mon. Weather Rev.*, 123(4):1112–1127, 1995.

**Titre :** Reconstruction d'écoulements turbulents complexes par assimilation de données images dans des modèles grandes échelles

**Mots clés :** assimilation de données, modèles grandes échelles,

**Résumé :** L'assimilation de données en tant qu'outil pour la mécanique des fluides a connu une croissance exponentielle au cours des dernières décennies. La possibilité de combiner des mesures précises mais partielles avec un modèle dynamique complet est précieuse et a de nombreuses applications dans des domaines allant de l'aérodynamique, à la géophysique et à l'aéraulique. Cependant, son utilité reste limitée en raison des contraintes imposées par l'assimilation de données notamment en termes de puissance de calcul, de besoins en mémoire et en informations préalables. Cette thèse tente de remédier aux différentes limites de la procédure d'assimilation pour faciliter plus largement son utilisation en mécanique des fluides.

Un obstacle majeur à l'assimilation des données est un coût de calcul prohibitif pour les écoulements complexes. Une modélisation de la turbulence à grande échelle est intégrée à la procédure d'assimilation afin de réduire considérablement la coût de calcul et le temps requis.

La nécessité d'une information volumétrique préalable pour l'assimilation est abordée à l'aide d'une nouvelle méthodologie de reconstruction développée et évaluée dans cette thèse. L'algorithme d'optimisation reconstruit les champs 3D à partir d'observations dans deux plans orthogonaux en exploitant l'homogénéité directionnelle. La méthode et ses variantes fonctionnent bien avec des ensembles de données synthétiques et expérimentaux fournissant des reconstructions précises.

La méthodologie de reconstruction permet également d'estimer la matrice de covariance d'ébauche indispensable à un algorithme d'assimilation efficace.

Tous les ingrédients sont combinés pour effectuer avec succès l'assimilation de données variationnelles d'un écoulement turbulent dans le sillage d'un cylindre à un nombre de Reynolds transitoire. L'algorithme d'assimilation est validé pour des observations volumétriques synthétiques et est évalué sur des observations expérimentales dans deux plans orthogonaux.

**Title :** Turbulent complex flows reconstruction via data assimilation in large eddy models

**Keywords :** data assimilation, large eddy simulation, 3d flow reconstruction, stochastic modelling

**Abstract :** Data assimilation as a tool for fluid mechanics has grown exponentially over the last few decades. The ability to combine accurate but partial measurements with a complete dynamical model is invaluable and has numerous applications to fields ranging from aerodynamics, geophysics, and internal ventilation. However, its utility remains limited due to the restrictive requirements for performing data assimilation in the form of computing power, memory, and prior information. This thesis attempts at redressing various limitations of the assimilation procedure in order to facilitate its wider use in fluid mechanics.

A major roadblock for data assimilation is the computational cost which is restrictive for all but the simplest of flows. Following along the lines of Joseph Smagorinsky, turbulence modelling through large-eddy simulation is incorporated in to the assimilation procedure to significantly reduce computing power and time required.

The requirement for prior volumetric information for assimilation is tackled using a novel reconstruction methodology developed and assessed in this thesis. The snapshot optimisation algorithm reconstructs 3D fields from 2D cross-planar observations by exploiting directional homogeneity. The method and its variants work well with synthetic and experimental data-sets providing accurate reconstructions.

The reconstruction methodology also provides the means to estimate the background covariance matrix which is essential for an efficient assimilation algorithm.

All the ingredients are combined to perform variational data assimilation of a turbulent wake flow around a cylinder successfully at a transitional Reynolds number. The assimilation algorithm is validated with synthetic volumetric observation and assessed on 2D cross-planar observations emulating experimental data.