From Extreme Behaviour to Optimal Closures — An Assemblage of Optimization Problems in 2D Turbulence

From Extreme Behaviour to Closures Models — An Assemblage of Optimization Problems in 2D Turbulence

By Pritpal Matharu, B.Sc., M.Sc.

A Thesis Submitted to the School of Graduate Studies in the Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy

McMaster University

Doctor of Philosophy (2022)

Hamilton, Ontario (Mathematics)

TITLE: From Extreme Behaviour to Closures Models — An Assemblage of Optimization Problems in 2D Turbulence

AUTHOR: Pritpal Matharu (McMaster University)

SUPERVISOR: Dr. Bartosz Protas

NUMBER OF PAGES: xvii, 147

Abstract

Turbulent flows occur in various fields and are a central, yet an extremely complex, topic in fluid dynamics. Understanding the behaviour of fluids is vital for multiple research areas including, but not limited to: biological models, weather prediction, and engineering design models for automobiles and aircraft. In this thesis, we study a number of fundamental problems that arise in 2D turbulent flows, using the 2D Navier-Stokes system. Introducing optimization techniques for systems described by partial differential equations (PDE), we frame these problems such that they can be solved using computational methods. We utilize adjoint calculus to build the computational framework to be implemented in an iterative gradient flow procedure, using the "optimize-then-discretize" approach. Pseudospectral methods are employed for solving PDEs in a numerically efficient manner. The use of optimization methods together with computational mathematics in this work provides an illuminating perspective on fluid mechanics.

We first apply these techniques to better understand enstrophy dissipation in 2D Navier-Stokes flows, in the limit of vanishing viscosity. By defining an optimization problem to determine optimal initial conditions, multiple branches of local maximizers were obtained each corresponding to a different mechanism producing maximum enstrophy dissipation. Viewing this quantity as a function of viscosity revealed quantitative agreement with an analytic bound, demonstrating the sharpness of this bound. We also introduce an extension of this problem, where enstrophy dissipation is maximized in the context of kinetic theory using the Boltzmann equation.

Secondly, these PDE-constrained optimization techniques were used to probe the fundamental limitations on the performance of the Leith eddy-viscosity closure model for 2D Large-Eddy Simulations of the Navier-Stokes system. Obtained by solving an optimization problem with a non-standard structure, the results demonstrate the optimal

eddy viscosities do not converge to a well-defined limit as regularization and discretization parameters are refined, hence the problem of determining an optimal eddy viscosity is ill-posed.

Further extending the problem of finding optimal eddy-viscosity closures, we consider imposing an additional nonlinear constraint on the control variable in the problem, in the form of requiring the time-averaged enstrophy be preserved. To address this problem in a novel way, we employ adjoint calculus to characterize a subspace tangent to the constraint manifold, which allows one to approximately enforce the constraint. Not only do we demonstrate that this produces better results when compared to the case without constraints, but this also provides a flexible computational framework for approximate enforcement of general nonlinear constraints. Lastly in this thesis, we introduce an optimization problem to study the Kolmogorov-Richardson energy cascade, where a pathway towards solutions is outlined.

Acknowledgements

First and foremost, I want to thank Dr. Bartosz Protas, my supervisor and mentor. Although it was uncertain if I was the correct fit for grad school, you guided me, taught me, provided me with incredible opportunities, and helped me grow immensely over the past 6 years. I cannot begin to express my appreciation for your constant support, motivation, encouragement to push my limits, and most importantly, your friendship. Truly, I couldn't have asked for a better supervisor, you have set the bar extremely high for mentors everywhere, and I hope someday to be half the mathematician and person that you are.

I want to thank Dr. Stephen Tullis for your guidance and help as my committee member over the years. I want to thank Dr. Nicholas Kevlahan for not only serving as a committee member and providing guidance during my committee meetings, but also for encouraging me explore new areas and for all of our wonderful discussions in Hamilton Hall. In addition, I want to thank Dr. Lia Bronsard. Lia, although you weren't on my committee, I always viewed you as one of my mentors and I cannot thank you enough for all your advice, guidance as a mathematician and teacher, and friendship throughout the years.

Being a member of the Department of Mathematics and Statistics at McMaster University has been an amazing experience. I always enjoyed coming into Hamilton Hall and getting to chat everyone, which made me deeply miss it during the years that we had to work remotely. Thank you to Dr. Stanley Alama, Dr. Dmitry Pelinovsky, and the late Dr. Walter Craig for all your helpful discussions. Special thanks to my friends, office mates, and colleagues, Samantha-Jo Caetano, Fletcher Gates, Nikolay Hristov, Adilbek Kairzhan, Graham Keiper, Michaela Kelly, Tyler Meadows, Mackenzie Neal, Ramsha Khan, Katarina Sacka, Yurij Salmaniw, Szymon Sobieszek, Martin Stelmach, Kala Studens, and Lee Van Brussel. Thank you for all your chats, mathematical discussions, and

for keeping me sane. In addition, thank you to the departmental administrative team for saving me from many headaches.

Thank you to the McMaster and Hamilton communities for giving me an outlet for stress relieve through various sports, and for introducing me to so many wonderful people. Special thanks to Arya Afsahi and Danielle Hayes for letting me vent many times during soccer and climbing sessions. I am also grateful for all the friends I have made internationally over the years. Special thanks to Dr. Tsuyoshi Yoneda for providing me with an insightful experience working with your group at Hitotsubashi University in Kunitachi, Tokyo, Japan. In addition, I am grateful to have lived with many fantastic people over the years. Special thanks to roommates and friends, Steven Cass, Tyler Horan, Sarah Robertson, Dillon Schnurr, and Szymon Sobieszek, for not only helping with my grad school struggles, but my personal ones too.

Finally, thank you to my family for your love, support, and encouragement. Thank you to my sisters, Gurpreet Matharu, Satinder Matharu, Kirandeep Matharu, and my parents, Balwinder Matharu and Gurdev Matharu, for being there for me, constant support, and encouragement, even if you don't fully understand what it is I do.

Contents

List of Figures

List of Tables

Declaration of Authorship

I, Pritpal MATHARU, declare that this thesis titled, "From Extreme Behaviour to Closures Models — An Assemblage of Optimization Problems in 2D Turbulence" and the work presented in it are my own. I confirm that:

- Chapter 1: Introduction
  - This chapter is my own work
- Chapter 2: On Maximum Enstrophy Dissipation in 2D Navier-Stokes Flows in the Vanishing Viscosity Limit
  - The content contained in this chapter has been accepted for publication in the Elsevier Journal Physica D: Nonlinear Phenomena
  - This chapter is the work of myself, Dr. Bartosz Protas, and Dr. Tsuyoshi Yoneda
- Chapter 3: Optimal Eddy Viscosity in Closure Models for 2D Turbulent Flows
  - The content contained in this chapter is copyright of the APS Journal Physical Review Fluids, used here under the publishing copyright agreement. Originally appeared Matharu and Protas (2022)
  - This chapter is the work of myself and Dr. Bartosz Protas
- Chapter 4: Enforcing Constraints Via Adjoint Analysis
  - This chapter is the work of myself and Dr. Bartosz Protas

- Chapter 5: Open Problems
  - This chapter is the work of myself, Dr. Bartosz Protas, Dr. Tsuyoshi Yoneda, Dr. Zhongyang Gu, and Hu Xin
- Chapter 6: Discussion and Conclusions
  - This chapter is my own work

All chapters were edited based on suggestions by Dr. Bartosz Protas.

Chapter 1

Introduction

1.1 Turbulence and the Navier-Stokes System

It is often stated that "turbulence is the last unsolved problem of classical physics" (Davidson 2015), while also a vital area of research with several challenges for scientists studying fundamental properties of turbulent flows, as well as engineers that utilize these flows. There are various open problems concerning turbulent flows, from understanding their fundamental behaviour such as the rate of dissipation of energy, to methods of accurately and efficiently computing them, which justifies referring to the problem of turbulence as "unsolved". Even defining "turbulent behaviour" is a difficult task, however it is often agreed upon that turbulent flows are chaotic, spatio-temporally complex, and/or have multiscale features (Davidson 2015). The most common model used to describe the motion of viscous incompressible fluids is the Navier-Stokes system which has been one of the most influential partial differential equations (PDE) in mathematical physics (Darrigol 2005). Derived from basic principles of conservation of mass and momentum of Newtonian fluids, the Navier-Stokes equations govern the motion of fluids from the macroscopic scale and are the main mathematical model used to describe flows and their complex properties. Here we consider a flow varying in time and space,

$(t, \mathbf{x}) \in (0, T] \times \Omega$, and although solid boundaries are a very important topic, for simplicity we will focus on periodic spatial domains Ω, as is quite common. The Navier-Stokes system for an incompressible Newtonian fluid with density $\rho = 1$ and kinematic viscosity ν, on a periodic domain $\Omega = \mathbb{T}^d$ ($d = 2$ or 3) can be written as (Davidson 2015)

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \boldsymbol{\nabla} p - \nu \boldsymbol{\Delta} \boldsymbol{u} = \boldsymbol{f} \qquad \text{in } (0, T] \times \Omega, \tag{1.1a}$$

$$\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0 \qquad \text{in } (0, T] \times \Omega, \tag{1.1b}$$

$$\boldsymbol{u}(t = 0) = \boldsymbol{u}_0 \qquad \text{in } \Omega, \tag{1.1c}$$

where $\boldsymbol{u}(t, \mathbf{x})$ is the velocity field, p is the pressure which ensures the fluid is incompressible, along with the incompressibility condition $\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$, $\boldsymbol{f}$ are external forces, and $\boldsymbol{u}_0$ denotes an initial condition. From weather behaviour to engineering design models for flows around objects, the Navier-Stokes system is central to various research areas. Known to produce solutions characterized by the utmost spatio-temporal complexity, the Navier-Stokes system is to describe turbulent behaviour. Consequently, there are open problems surrounding the Navier-Stokes system beyond questions related to its fundamental mathematical and physical properties.

The vorticity, which is defined as the curl of the velocity field, $\boldsymbol{w} = \boldsymbol{\nabla} \times \boldsymbol{u}$, can be viewed as "twice the angular velocity of a fluid" (Davidson 2015). We can write the governing system of $\boldsymbol{w}$ by taking the curl of (1.1) to obtain the so-called vorticity-velocity form of Navier-Stokes

$$\frac{\partial \boldsymbol{w}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{w} - \boldsymbol{w} \cdot \boldsymbol{\nabla} \boldsymbol{u} - \nu \boldsymbol{\Delta} \boldsymbol{w} = \boldsymbol{\nabla} \times \boldsymbol{f} \qquad \text{in } (0, T] \times \Omega, \tag{1.2a}$$

$$\boldsymbol{\Delta} \cdot \boldsymbol{u} = -\boldsymbol{\nabla} \times \boldsymbol{w} \qquad \text{in } (0, T] \times \Omega, \tag{1.2b}$$

$$\boldsymbol{w}(t = 0) = \boldsymbol{w}_0 \qquad \text{in } \Omega, \tag{1.2c}$$

where $\boldsymbol{w}_0 = \boldsymbol{\nabla} \times \boldsymbol{u}_0$. In (1.1a), we notice that the pressure is now eliminated however there is a new term, $\boldsymbol{w} \cdot \boldsymbol{\nabla} \boldsymbol{u}$, known as "vortex stretching". This term can produce vorticity, due to the stretching of vortices and vice-versa for the destruction of vorticity as a result of compression of vortices. This process is the key mechanism of the energy cascade, and this vortex stretching results in the transfer of energy down to small scales (Tennekes and Lumley 1972).

Although the Navier-Stokes system have been known for over a century and is vital for understanding fluid flows, there are still several open problems for understanding this seemingly straightforward set of equations. Most famously, the "blow-up problem" that concerns the question whether smooth (classical) solutions to the 3D Navier-Stokes system can spontaneously develop singularities in finite time remains one of the unsolved Millennium Problems recognized by The Clay Mathematics Institute (Fefferman 2006).

There had been systematic attempts to understand turbulence (Taylor 1935; Taylor 1938; Kármán and Howarth 1938), however a conceptual framework to understand this problem started with the Kolmogorov 1941 theory, now known as K41 (Kolmogorov 1941c; Kolmogorov 1941b; Kolmogorov 1941a; Frisch 1995). This work focused on statistically stationary turbulence invariant with respect to translations and rotations, hence homogeneous, isotropic turbulence. Using a statistical perspective, K41 theory established the four-fifths law for fully developed turbulence using third-order longitudinal structure functions. This is a key result providing some guidance to turbulence study, establishing an exact result for turbulent flows. A major consequence of the four-fifths law is the $-5/3$ law of the energy spectrum, which is a result of dimensional analysis, that largely agrees with experimental work. Further refinements were made to K41 theory (Kolmogorov 1962; Parisi and Frisch 1985), which built the standard theoretical framework for studying statistical properties of turbulent flows.

In turbulent flows, the viscosity ν is responsible for the loss of kinetic energy via

dissipation. Due to the nonlinearity present in the governing equations, in particular vortex stretching, in turbulent flow energy is transferred amongst scales and typically in three-dimensional (3D) turbulence, energy is cascaded down to the small scales where it is dissipated. In both numerical and physical experiments, as the viscosity ν is lowered with all other flow parameters fixed, the rate of energy dissipation in the system approaches a finite positive limit (Frisch 1995). Although this was used as an assumption in K41 theory and has been observed experimentally, it is still yet to be understood from a physical and mathematical point of view. This is a singular asymptotic limit, often referred to as the phenomenon of "anomalous dissipation" or the "zeroth law of turbulence". Anomalous dissipation is closely related to the "blow-up problem" and is of particular interest, since the velocity gradients are required to become unbounded to achieve this finite positive limit, which is directly related to questions concerning the singularity formation in the inviscid Euler equations (Constantin 2007). In this work, we make contributions to understanding the behaviour of turbulent flows in the limit of vanishing viscosity for two-dimensional (2D) flows.

Not only are there open problems regarding the foundations of turbulence, but also difficulties that arise when numerically computing solutions in an accurate and efficient manner. The closure problem is arguably one of the most important outstanding open problems in turbulence research. Given the extreme spatio-temporal complexity of turbulent flows, accurate numerical solutions of equations that govern turbulent flows, namely the Navier-Stokes system, requires resolutions exceeding the capability of commonly accessible computational resources, even at modest Reynolds numbers. As a result, this equation cannot be efficiently solved in most situations of practical interest and one must rely on its various simplifications. One such approach which has gained widespread popularity in engineering practice is Large Eddy Simulation (LES), a computationally efficient substitute for the "brute-force" Direct Numerical Simulation (DNS). LES depends on the solution of a suitably filtered version of the governing

system which only accounts for the large-scale motions. Accounting for the effect of the small-scale motions on the resolved scales is the essence of the notorious "closure problem". Mathematically, the closure problem consists of there being more unknowns than equations in the LES equations, in the form of a quantity known as the Reynolds stresses, which describes the interactions between motions at large resolved scales and unresolved small-scale motions. To address this, a representation for the unresolved expression is introduced as a suitable model for turbulent motions occurring at small unresolved scales, which is typically called the closure model. The eddy viscosity model, arguably the most popular type of closure model, assumes a functional dependence of the eddy viscosity on the resolved solution, typically proportional to the magnitude of the resolved velocity gradient in the form of the resolved rate of strain tensor, and has been utilized in several engineering applications. In particular, the Smagorinsky model (Smagorinsky 1963) and many of its modifications are widely used as closure models to compute LES. These Smagorinsky models have common drawbacks, including being too dissipative near walls (Davidson 2015) and assuming the closure term to be strictly dissipative (Rodi et al. 2013). It is important to realize that, despite their popularity, the development of Smagorinsky models and their various improvements has been traditionally dominated by empiricism. This work provides a new perspective on determining an optimal closure model and some fundamental performance limitations for LES in 2D flows.

1.2 2D Turbulence

Two-dimensional turbulence has been the main focus of several studies and understanding its complex behaviour has been the key for various research applications, e.g., geophysical flows (Boffetta and Musacchio 2010; Boffetta and Ecke 2012). In addition to the importance of 2D turbulence in several fields, 2D turbulence is also sometimes studied

due to its similarities with turbulence in 3D, but is simpler from both the mathematical and computational perspective. Notwithstanding, 2D turbulence possesses some properties that are intrinsically different than 3D turbulence.

Let us consider a 2D velocity field $\boldsymbol{u}(t,\mathbf{x}) = (u_1, u_2)^T$, i.e. $d = 2$ for (1.1), governed by the incompressible Navier-Stokes equation with uniform unit density, $\rho = 1$, and periodic boundary conditions, where Ω is a 2D torus. Then, the vorticity field $\boldsymbol{w} = [0, 0, w]^T$ has only one nonzero component perpendicular to the plane of motion, and can be represented by the scalar field $w = -\boldsymbol{\nabla}^{\perp} \cdot \boldsymbol{u}$ with $\boldsymbol{\nabla}^{\perp} = (\partial_{x_2}, -\partial_{x_1})^T$. Now when taking the curl of (1.1a), we write the 2D Navier-Stokes equation for the scalar vorticity field, and obtain the vorticity equation $\partial_t w + \boldsymbol{u} \cdot \boldsymbol{\nabla} w - \nu \boldsymbol{\Delta} w = f_\omega$ in Ω where $f_\omega = \boldsymbol{\nabla} \times \boldsymbol{f}$, noting the absence of the vortex stretching term when compared to (1.2a). Defining the stream function ψ by $\boldsymbol{u} = \boldsymbol{\nabla}^{\perp}\psi$, we obtain the relation between vorticity and streamfunction in the form of the following Poisson equation $\boldsymbol{\Delta}\psi = -w$, as a consequence of the Helmholtz decomposition. Using these two equations, along with a suitable initial condition w_0, we define the PDE governing the time evolution of $w = w(t,\mathbf{x})$ in the vorticity-streamfunction formulation

$$\frac{\partial w}{\partial t} + \boldsymbol{\nabla}^{\perp}\psi \cdot \boldsymbol{\nabla} w - \nu \boldsymbol{\Delta} w = f_\omega \qquad \text{in } (0,T] \times \Omega, \tag{1.3a}$$

$$\boldsymbol{\Delta}\psi = -w \qquad \text{in } (0,T] \times \Omega, \tag{1.3b}$$

$$w(t=0) = w_0 \qquad \text{in } \Omega, \tag{1.3c}$$

where we require $\int_\Omega w_0(\mathbf{x}) d\mathbf{x} = 0$, to ensure that the Laplacian in (1.3b) can be inverted on the periodic domain Ω. The PDE system (1.3) is known to produce regular solutions and is globally well-posed in the classical sense (Kreiss and Lorenz 2004). One key physical effect described by the 3D Navier-Stokes system but absent in 2D is vortex stretching, $\boldsymbol{w} \cdot \boldsymbol{\nabla} \boldsymbol{u}$ in (1.2), which is crucial for energy transfer, as mentioned before. The Batchelor-Kraichnan theory (Kraichnan 1967; Batchelor 1969) also provides the heuristic

principles behind the different mechanisms of energy transfer in 2D turbulence. Unlike 3D turbulence, 2D turbulence is characterized by the phenomenon of inverse energy cascade, in addition to a forward enstrophy cascade. As a result, energy accumulates at the large scales which fill the domain (Davidson 2015). This occurs after vortex filaments are formed and vortex sheets appear in the fully developed turbulent state, where eddies grow in size. With the fascinating behaviour of an inverse energy and forward enstrophy cascade, the dynamics of 2D turbulence has been investigated in both the strictly decaying and forced setting (Farazmand et al. 2011; Lindborg and Vallgren 2010; Vallgren and Lindborg 2011; Bracco and McWilliams 2010; Boffetta and Musacchio 2010). The Kraichnan-Batchelor-Leith (Kraichnan 1967; Batchelor 1969; Leith 1968) theory provides a heuristic, albeit well-founded, theoretical framework for studying 2D statistically stationary turbulence.

1.3 PDE optimization

PDE optimization has versatile applications, and can be applied to various problems to gain insight beyond what standard analytic methods provide. Let us consider a general PDE-constrained optimization problem of the form

$$\begin{cases} \min\limits_{(u,\Gamma)} \Lambda(u,\Gamma), \\ \\ \text{subject to:} \quad \mathcal{G}(u,\Gamma) = 0, \end{cases} \tag{1.4}$$

where u and Γ are the state and control variables, respectively, and $\mathcal{G}$ is the constraint typically a PDE with appropriate initial and boundary conditions. We can introduce Lagrange multipliers λ to handle the constraint using a standard approach, which leads to a min-max saddle problem

$$\min_{(u,\Gamma)} \max_{\lambda} \left[\Lambda(u,\Gamma) - \lambda\, \mathcal{G}(u,\Gamma)\right].$$

This problem is often difficult to solve numerically when $\mathcal{G}$ represents a complicated PDE. If the constraint $\mathcal{G}(u, \Gamma) = 0$ can be solved for $u = u(\Gamma)$, with the help of, e.g., the implicit function theorem, then we can define a "reduced" objective functional $\Phi := \Lambda(u(\Gamma), \Gamma)$. The constrained problem (1.4) is then equivalent to the unconstrained optimization problem

$$\min_{\Gamma} \Phi(\Gamma).$$

Now, a local minimizer Γ can be found using a (discrete) gradient flow.

Adjoint calculus has had a long history in PDE-constrained optimization (Lions 1968), including problems in fluid mechanics (Gunzburger 2003). Defining suitable cost functionals, dependent on a control parameter and constrained by a PDE, we can determine minimizers or maximizers of this objective function via adjoint-based gradients. When dealing with non-convex optimization problems, the optimal solution for the control parameter cannot guarantee to be a global minimizer of the objective functional and is only a local minimizer or maximizer of the objective functional, which becomes particularly pertinent when dealing with solving these problems computationally. This introduces two approaches that can be implemented when using adjoint calculus to solve these problems computationally: *optimize-then-discretize* or *discretize-then-optimize*. These approaches lead to different approximations because the optimization and discretization procedures do not commute in general (Gunzburger 2003), however there are some special cases where they do indeed commute.

In this we work, we focus on the *optimize-then-discretize* approach, where optimality conditions, constraints, and gradient expressions are all formulated at the continuous PDE level, and then discretized to be numerically solved. This allows us to study the mathematical properties and analytical structure of the optimal solutions, while utilizing computationally flexible methods. These methods have been used in various

applications, including but not limited to: transfer of heat in steel production (Yang et al. 2021), aeroelastic optimization of wings (Maute et al. 2003), and optimal design of open and closed loop controls for nonlinear mechanical systems (Pappalardo and Guida 2018). In terms of applications, the novelty of the work presented here is to use optimization methods to shed light on the inner workings of the Navier-Stokes system.

1.4 Summary of the Main Results

The main focus of this work concerns three problems, two of which are closely related, which provide insight about some fundamental theoretical and computational problems in 2D turbulence. The main approach we utilize is suitably selected optimization problems, and solving them using an *optimize-then-discretize* approach. Using this approach allows us to formulate the optimization techniques in the continuous setting and then numerically solve the optimization problems via an iterative gradient descent/ascent method. Several novel methods were developed in order to solve these problems as they possess non-standard structure, and we introduce new techniques now available for future work. The contributions in each chapter is described below.

In Chapter 2, we consider enstrophy dissipation in 2D Navier-Stokes flows and focus on how this quantity behaves in the limit of vanishing viscosity. After recalling a number of a priori estimates providing lower and upper bounds on this quantity, we state an optimization problem aimed at probing the sharpness of these estimates as functions of viscosity. More precisely, solutions of this problem are the initial conditions with fixed palinstrophy and possessing the property that the resulting 2D Navier-Stokes flows locally maximize the enstrophy dissipation over a given time window. This problem is solved numerically with an adjoint-based gradient ascent method and solutions obtained for a broad range of viscosities and lengths of the time window reveal the presence of multiple branches of local maximizers, each associated with a distinct mechanism for the

amplification of palinstrophy. The dependence of the maximum enstrophy dissipation on viscosity is shown to be in quantitative agreement with the estimate due to (Ciampa et al. 2021), demonstrating the sharpness of this bound.

In Chapter 3, we consider the question of fundamental limitations on the performance of eddy-viscosity closure models for turbulent flows, focusing on the Leith model for 2D Large-Eddy Simulation. Optimal eddy viscosities depending on the magnitude of the vorticity gradient are determined subject to minimum assumptions by solving PDE-constrained optimization problems defined such that the corresponding optimal Large-Eddy Simulation best matches the filtered Direct Numerical Simulation. The framework for determining optimal closure models was developed in Matharu and Protas (2020), where we used a simple one-dimensional model problem. Here, we first consider pointwise match in the physical space and the main finding is that with a fixed cutoff wavenumber k_c, the performance of the Large-Eddy Simulation systematically improves as the regularization in the solution of the optimization problem is reduced and this is achieved with the optimal eddy viscosities exhibiting increasingly irregular behaviour with rapid oscillations. Since the optimal eddy viscosities do not converge to a well-defined limit as the regularization vanishes, we conclude that in this case the problem of finding an optimal eddy viscosity does not in fact have a solution and is thus ill-posed. We argue that this observation is consistent with the physical intuition concerning closure problems. The second problem we consider involves matching time-averaged vorticity spectra over small wavenumbers. It is shown to be better behaved and to produce physically reasonable optimal eddy viscosities. We conclude that while better behaved and hence practically more useful eddy viscosities can be obtained with stronger regularization or by matching quantities defined in a statistical sense, the corresponding Large-Eddy Simulations will not achieve their theoretical performance limits.

The optimization problems considered in Matharu and Protas (2020) and in Chapter 3

rely on pointwise observations to match a target solution and the flow field predicted by the Navier-Stokes system. Unless all observations were matched exactly, this does not ensure that certain key quantities, e.g., enstrophy, is preserved. Due to the complexity of these nonstandard optimization problems, enforcing a complicated constraint in a computationally efficient and accurate manner is not straightforward. In Chapter 4, we develop a novel method for approximately enforcing nonlinear constraints defined on complex manifolds with low codimension. Using adjoint calculus and a judiciously defined adjoint system, we characterize a subspace tangent to the constraint manifold. This is done by describing the tangent subspace with a suitable "normal vector", which then allows one to define a projected gradient flow, to approximately satisfy the nonlinear constraint. This method is used to enforce a time-averaged enstrophy constraint, while solving an optimization problem to determine an optimal eddy viscosity, as was described in Chapter 3.

The results described in Chapter 4 shed some light on the optimal form of eddy viscosities needed in order for the flow to maintain energy-type quantities, such as enstrophy. Interestingly, the tools developed in this research introduce a new point of view for utilizing adjoint calculus in optimization problems. Adjoint calculus has a long history in being used to compute gradient directions (Gunzburger 2003), however previously it had not been utilized for enforcing complicated nonlinear constraints. Instead, techniques such as Newton's method were used to enforce such constraints, as done in Farazmand and Sapsis (2017). As a computationally efficient alternative, the present work demonstrates adjoint calculus can be generalized to approximately enforce the control variable onto the constraint manifold. Employing this method now makes a class of previously intractable optimization problems possible to solve, and also opens avenues to use adjoint analysis in a less traditional manner to solve novel problems.

In Chapter 5, certain extensions of some of the problems considered in the thesis

are discussed together with an outline of possible solutions approaches. An extension of the problem in Chapter 2 is considered from the perspective of kinetic theory using the Boltzmann equation. The idea behind this problem is to study enstrophy dissipation in 2D Navier-Stokes using the Boltzmann equation, while utilizing the framework of Chapter 2. We also consider self-similar flow structures arising in the Kolmogorov-Richardson energy cascade, by studying the energy transfer in 1D Burgers flow with an optimally determined initial condition.

Chapter 2

On Maximum Enstrophy Dissipation in 2D Navier-Stokes Flows in the Limit of Vanishing Viscosity

2.1 Introduction

The physical phenomenon of "anomalous dissipation", also referred to as the "zeroth law of turbulence", is one of the oldest problems in turbulence (Frisch 1995). This empirical law states that the energy dissipation in either forced or decaying three-dimensional (3D) turbulent flows approaches a nonzero limit as the fluid viscosity $\nu > 0$ vanishes, all other flow parameters remaining fixed. There is a lot of evidence coming from both experiments and numerical simulations supporting this anomalous behavior of the energy dissipation (Sreenivasan 1998; Yeung et al. 2015), but we are still far from being able to understand this problem from the mathematical point of view. The main consequence of the dissipation anomaly is an unbounded increase of velocity gradients which would in

turn imply finite-time singularities in solutions of the inviscid Euler equations (Dascaliuc and Grujić 2012). Similar dissipation anomalies are also known to occur in the behavior of passive scalars (Sreenivasan 2019; Mazzucato 2022).

The dissipation anomaly arises in solutions of the one-dimensional (1D) Burgers equation (Eyink and Drivas 2015). As regards 2D flows, the relevant question is about the behavior of the enstrophy dissipation in the limit of vanishing viscosity. The assumption that enstrophy dissipation tends to a finite (nonzero) limit as $\nu \to 0$ underlaid Batchelor's theory of 2D turbulence (Batchelor 1969). However, in Tran and Dritschel (2006) it was argued that this quantity in fact vanishes in the inviscid limit such that Navier-Stokes flows in 2D are not subject to dissipation anomaly. This result was confirmed by rigorous analysis of the inviscid limit of 2D Navier-Stokes flows (Filho et al. 2006).

While there is no dissipation anomaly in 2D flows, it is interesting to know the worst-case (slowest) rate at which the enstrophy dissipation vanishes in the limit $\nu \to 0$. A number of theoretical results, in the form of both lower and upper bounds on the dependence of the enstrophy dissipation on ν, have been established and are reviewed below. The goal of the present study is to address this question computationally by finding flows with the largest possible enstrophy dissipation as the viscosity vanishes. Such "extreme" flows will be found by solving suitably defined optimization problems with constraints in the form of partial differential equations (PDEs). This will provide insights about the sharpness of various rigorous bounds on the enstrophy dissipation in the inviscid limit. While methods of PDE optimization have had a long history in various applied areas (Gunzburger 2003), they have recently been employed to study certain fundamental problems concerning extreme behavior in fluid mechanics (Protas 2022). In particular, problems somewhat related to the subject of the present study were investigated using such techniques in Ayala and Protas (2014a), Ayala and Protas (2014b), and Ayala et al. (2018).

We consider the incompressible Navier-Stokes system on a 2D periodic domain $\Omega := \mathbb{T}^2 = [0,1]^2$ (":=" means "equal to by definition") which can be written in the vorticity form as

$$\frac{\partial w_\nu}{\partial t} + \boldsymbol{\nabla}^\perp \psi_\nu \cdot \boldsymbol{\nabla} w_\nu = \nu \boldsymbol{\Delta} w_\nu \qquad \text{in } (0,T] \times \Omega, \tag{2.1a}$$

$$-\boldsymbol{\Delta} \psi_\nu = w_\nu \qquad \text{in } (0,T] \times \Omega, \tag{2.1b}$$

$$w_\nu(t=0) = \phi \qquad \text{in } \Omega, \tag{2.1c}$$

where w_ν and ψ_ν are the vorticity component perpendicular to the plane of motion and the corresponding streamfunction, both assumed to satisfy the periodic boundary conditions in the space variable $\mathbf{x}$, whereas $T > 0$ is the length of the time window considered. The symbol ϕ denotes the initial condition which without loss of generality is assumed to have zero mean, i.e.,

$$\int_\Omega \phi(\mathbf{x})\, d\mathbf{x} = 0. \tag{2.2}$$

Problem (2.1) is known to be globally well-posed in the classical sense (Kreiss and Lorenz 2004). Its solutions are characterized by the enstrophy and palinstrophy defined, respectively, as[1]

$$\mathcal{E}(w_\nu(t,\cdot)) := \frac{1}{2} \int_\Omega |w_\nu(t,\mathbf{x})|^2 \, d\mathbf{x}, \tag{2.3}$$

$$\mathcal{P}(w_\nu(t,\cdot)) := \frac{1}{2} \int_\Omega |\boldsymbol{\nabla} w_\nu(t,\mathbf{x})|^2 \, d\mathbf{x}, \tag{2.4}$$

which satisfy the relation

$$\frac{d\mathcal{E}(t)}{dt} = -2\nu \mathcal{P}(t). \tag{2.5}$$

[1]For consistency with the convention used in our earlier studies, cf. Protas (2022), both these quantities are defined with a factor of 1/2. Without the risk of confusion we will sometimes use the simplified notation $\mathcal{E}(t) = \mathcal{E}(w_\nu(t,\cdot))$ and $\mathcal{P}(t) = \mathcal{P}(w_\nu(t,\cdot))$.

We then define our main quantity of interest as

$$\chi_\nu(\phi) := \frac{2\nu}{T}\int_0^T \mathcal{P}(t)\,dt = \frac{\nu}{T}\int_0^T\int_\Omega |\boldsymbol{\nabla} w_\nu(t,\mathbf{x};\phi)|^2\,d\mathbf{x}dt = \frac{\mathcal{E}(0)-\mathcal{E}(T)}{T}, \tag{2.6}$$

which represents the enstrophy dissipation per unit of time and will be viewed here as a function of the initial data ϕ.

The enstrophy dissipation (2.6) has been the subject of numerous estimates. For technical reasons we will hereafter assume that $0 < \nu < 1$. We refer to the following result as a "conjecture" since it relies on some assumptions, albeit well justified, about the form of the spectrum of the solutions of (2.1).

Conjecture 1 (Tran & Dritschel (Tran and Dritschel 2006))**.** *The enstrophy dissipation in solutions of system (2.1) is bounded above by*

$$\chi_\nu \leq C\,[-\ln(\nu)]^{-\frac{1}{2}}, \tag{2.7}$$

for some constant $C > 0$ depending on the initial condition ϕ and the length T of the time window.

Hereafter $C = C(T)$ will denote a generic positive constant depending on the length T of the considered time window with numerical values differing from one instant to another.

Bounds on enstrophy dissipation are closely related to another problem which has recently received considerable attention, namely, the question of the convergence as $\nu \to 0$ of Navier-Stokes flows to solutions of the inviscid Euler equations obtained by setting $\nu = 0$ in (2.1a) and corresponding to the same initial condition ϕ. More specifically, noting (2.5), the fact that solutions of the inviscid Euler system conserve the enstrophy

and using the reverse triangle inequality, we have

$$\begin{aligned}
\chi_\nu(\phi) &= \frac{\nu}{T}\int_0^T \|\boldsymbol{\nabla} w_\nu(t,\mathbf{x};\phi)\|^2_{L^2(\Omega)}\,dt = \frac{2\nu}{T}\int_0^T \mathcal{P}(t)\,dt \\
&= \frac{1}{T}\left[\mathcal{E}(0)-\mathcal{E}(T)\right] = \frac{1}{T}\left[\|\phi\|^2_{L^2(\Omega)} - \|w_\nu(T,\mathbf{x};\phi)\|^2_{L^2(\Omega)}\right] \\
&= \frac{1}{T}\left[\|w(T,\mathbf{x};\phi)\|^2_{L^2(\Omega)} - \|w_\nu(T,\mathbf{x};\phi)\|^2_{L^2(\Omega)}\right] \\
&\leq \frac{1}{T}\left[\|w(T,\mathbf{x};\phi)\|_{L^2(\Omega)} + \|w_\nu(T,\mathbf{x};\phi)\|_{L^2(\Omega)}\right]\,\|w(T,\mathbf{x};\phi)-w_\nu(T,\mathbf{x};\phi)\|_{L^2(\Omega)} \\
&\leq \frac{2}{T}\,\|\phi\|_{L^2(\Omega)}\,\|w(T,\mathbf{x};\phi)-w_\nu(T,\mathbf{x};\phi)\|_{L^2(\Omega)}\,, \qquad (2.8)
\end{aligned}$$

where $w(t,\mathbf{x}) := w_0(t,\mathbf{x})$ denotes the vorticity in the inviscid Euler flow. The above relation shows that the enstrophy dissipation over the time window $[0,T]$ can be bounded from above in terms of the difference of the vorticity fields in the viscous and inviscid flows obtained with the same initial data ϕ at time $t=T$. Quantifying this difference in terms of viscosity as $\nu\to 0$ has been the subject of some recent studies. In Constantin et al. (2022) the authors showed the strong convergence of w_ν to w as $\nu\to 0$ when $\phi\in L^\infty(\Omega)$, implying the vanishing of the right-hand side (RHS) in (2.8). Moreover, the following estimate was established in the case when $\phi\in L^\infty(\Omega)\cap B^s_{2,\infty}(\Omega)$ for some $s>0$, where L^p and $B^s_{p,q}$ are the usual Lebesgue and Besov spaces,

$$\sup_{t\in[0,T]}\|w(t,\cdot)-w_\nu(t,\cdot)\|_{L^p(\Omega)} \leq C(\nu T)^{\frac{s\,e^{-2CTM_\infty}}{p(1+s\,e^{-CTM_\infty})}}, \qquad (2.9)$$

where $M_\infty := \|\phi\|_{L^\infty(\Omega)}$. This problem was revisited in Ciampa et al. (2021) where it was proved that

$$\sup_{t\in[0,T]}\|w(t,\cdot)-w_\nu(t,\cdot)\|_{L^p(\Omega)} \leq C\,M_\infty^{1-\frac{1}{p}}\max\left\{\gamma_{\phi,p,M_\infty}(C\,\nu^{\frac{e^{-CT}}{2}}),\left(C\,\nu^{\frac{e^{-CT}}{2}}\right)^{\frac{e^{-CT}}{2p}}\right\}, \qquad (2.10)$$

where now $C=C(T,M_\infty)$ and $\gamma_{\phi,p,M_\infty}:\mathbb{R}^+\to\mathbb{R}^+$ is a continuous function such that $\gamma_{\phi,p,M_\infty}(0)=0$. Additional results were also obtained recently in Lopes et al. (2021)

and Seis (2021). In particular, the following bound was produced in Seis (2021), which improves the rate of the weak convergence of w_ν to w as $\nu \to 0$,

$$\sup_{t\in[0,T]} \|w(t,\cdot) - w_\nu(t,\cdot)\|_{\dot{H}^{-1}(\Omega)} \leq C \left[\frac{\nu}{|\ln(\nu)|}\right]^{\frac{e^{-CT}}{2}} . \tag{2.11}$$

We reiterate that, in the light of relation (2.8), inequalities (2.9)–(2.10) imply viscosity-dependent upper bounds on the enstrophy dissipation (2.6). This is not the case for estimate (2.11) as it involves a weaker norm than in (2.8). We will nonetheless refer to this estimate when we discuss our results in Section 2.4 with the hope that our findings may inspire further work on refining this estimate. On the other hand, as is evident from the following theorem, a lower bound on the maximum enstrophy dissipation is also available.

Theorem 1 (Jeong & Yoneda (Jeong and Yoneda 2021))**.** *Let w_ν be the unique solution to (2.1). Then, there exists initial data ϕ such that the enstrophy dissipation is bounded below by*

$$\chi_\nu \geq C\nu \, [-\ln(\nu)]^{\frac{1}{2}} . \tag{2.12}$$

Upper bounds on the energy and enstrophy dissipation in 2D Navier-Stokes flows in the presence of external forcing were obtained in Alexakis and Doering (2006).

In the present study we construct families of 2D Navier-Stokes flows which at fixed values of the viscosity ν locally maximize the enstrophy dissipation χ_ν over the prescribed time window $[0, T]$. These flows are found using methods of numerical optimization to solve PDE-constrained optimization problems in which the enstrophy dissipation (2.6) is maximized with respect to the initial condition ϕ in (2.1) subject to certain constraints. This is a nonconvex optimization problem and we demonstrate that for every pair ν and T it admits several branches of locally maximizing solutions, each corresponding to a distinct dynamic mechanism for amplification of palinstrophy (which, as is evident from

(2.5), drives the dissipation of enstrophy). Finally, by assessing the dependence of the maximum enstrophy dissipation determined in this way for fixed T on the viscosity for decreasing values of ν, we arrive at interesting new insights about the sharpness of the different a priori estimates discussed above.

The structure of the chapter is as follows: in the next section we introduce the optimization problem formulated to maximize the enstrophy dissipation whereas in Section 2.3 we outline our gradient-based approach to finding families of local maximizers of that problem; computational results are presented in Section 2.4 whereas discussion and final conclusions are deferred to the last section.

2.2 Optimization Problem

Given a fixed viscosity ν and length T of the time window, we aim to construct flows maximizing the enstrophy dissipation χ_ν which will be accomplished by finding suitable optimal initial conditions $\breve{\phi}_\nu^T$ in system (2.1). Since the enstrophy dissipation is given in terms of a time integral of the palinstrophy, cf. (2.6), we will restrict our attention to initial data with bounded palinstrophy $\mathcal{P}_0 := \mathcal{P}(\phi)$, even though system (2.1) admits classical solutions for a much broader class of initial data (Kreiss and Lorenz 2004). We thus have the following optimization problem.

Problem 1. *Given $\mathcal{P}_0, \nu, T > 0$ in system (2.1) and the objective functional (2.6), find*

$$\breve{\phi}_\nu^T = \arg\max_{\phi \in \mathcal{S}} \chi_\nu(\phi), \quad \textit{where} \quad \mathcal{S} := \left\{ \phi \in H^1(\Omega) \ : \ \int_\Omega \phi(\mathbf{x})\, d\mathbf{x} = 0, \ \mathcal{P}(\phi) = \mathcal{P}_0 \right\}.$$

The Sobolev space $H^1(\Omega)$ is endowed with the inner product

$$\forall\ p_1, p_2 \in H^1(\Omega) \qquad \langle p_1, p_2 \rangle_{H^1(\Omega)} = \int_\Omega p_1 p_2 + \ell^2\, \boldsymbol{\nabla} p_1 \cdot \boldsymbol{\nabla} p_2 \, d\mathbf{x}, \tag{2.13}$$

where $\ell \in \mathbb{R}^+$ is a parameter. We note that the inner products in (2.13) corresponding to different values of ℓ are equivalent as long as $0 < \ell < \infty$. However, as will be shown in the next section, the choice of the parameter ℓ plays an important role in the numerical solution of Problem 1, cf. (2.22). With the initial palinstrophy $\mathcal{P}_0$ fixed, we will find families of locally maximizing solutions of Problem 1 parameterized by T for a range of viscosities ν. Our approach to finding such local maximizers is described next.

2.3 Solution Approach

2.3.1 Gradient-Based Optimization

Since Problem 1 is designed to test certain subtle mathematical properties of system (2.1), we choose to formulate the solution approach in the continuous ("optimize-then-discretize") setting, where the optimality conditions, constraints and gradient expressions are derived based on the original PDE before being discretized for the purpose of numerical evaluation, instead of the alternative "discretize-then-optimize" approach often used in applications (Gunzburger 2003). We first describe the discrete gradient flow focusing on computation of the gradient of the objective functional $\chi_\nu(\phi)$ with respect to the initial condition ϕ and then provide some details about numerical approximations.

For given values of $\mathcal{P}_0$, ν and T, a local maximizer $\breve{\phi}_\nu^T$ of Problem 1 can be found as $\breve{\phi}_\nu^T = \lim_{n\to\infty} \phi^{(n)}$ using the following iterative procedure representing a discretization of a gradient flow projected on $\mathcal{S}$

$$\begin{aligned} \phi^{(n+1)} &= \mathbb{P}_{\mathcal{S}} \left(\phi^{(n)} + \tau_n \nabla \chi_\nu \left(\phi^{(n)} \right) \right), \\ \phi^{(1)} &= \phi_0, \end{aligned} \tag{2.14}$$

where $\phi^{(n)}$ is an approximation of the maximizer obtained at the n-th iteration, ϕ_0 is the initial guess assumed to have zero mean and τ_n is the length of the step in the direction of the gradient $\nabla\chi_\nu(\phi^{(n)})$. The palinstrophy constraint is enforced by application of a

projection operator $\mathbb{P}_{\mathcal{S}} : H^1(\Omega) \to \mathcal{S}$ to be defined below. We defer discussion of the choice of the initial guess φ_0 to the end of this subsection.

A key step in procedure (2.14) is evaluation of the gradient $\nabla\chi_\nu(\phi)$ of the objective functional $\chi_\nu(\phi)$, cf. (2.6), representing its (infinite-dimensional) sensitivity to perturbations of the initial condition ϕ, and it is essential that the gradient be characterized by the required regularity, namely, $\nabla\chi_\nu(\phi) \in H^1(\Omega)$. This is, in fact, guaranteed by the Riesz representation theorem (Luenberger 1969) applicable because the Gâteaux (directional) differential $\chi'_\nu(\phi;\cdot) : H^1(\Omega) \to \mathbb{R}$, defined as $\chi'_\nu(\phi;\phi') := \lim_{\epsilon\to 0} \epsilon^{-1}\left[\chi_\nu(\phi+\epsilon\phi') - \chi_\nu(\phi)\right]$ for some perturbation $\phi' \in H^1(\Omega)$, is a bounded linear functional on $H^1(\Omega)$. The Gâteaux differential can be computed directly to give

$$\begin{aligned}\chi'_\nu(\phi;\phi') =& \frac{2\nu}{T}\int_0^T\int_\Omega \boldsymbol{\nabla} w_\nu(t,\mathbf{x};\phi)\cdot\boldsymbol{\nabla} w'_\nu(t,\mathbf{x};\phi,\phi')\,d\mathbf{x}dt,\\ =& -\frac{2\nu}{T}\int_0^T\int_\Omega \boldsymbol{\Delta} w_\nu(t,\mathbf{x};\phi)\, w'_\nu(t,\mathbf{x};\phi,\phi')\,d\mathbf{x}dt,\end{aligned} \tag{2.15}$$

where the last equality follows from integration by parts and the perturbation field $w'_\nu = w'_\nu(t,\mathbf{x};\phi,\phi')$ is a solution of the Navier-Stokes (2.1) system linearized around the trajectory corresponding to the initial data ϕ (Gunzburger 2003), i.e.,

$$\mathcal{K}\begin{bmatrix} w'_\nu \\ \psi'_\nu\end{bmatrix} := \begin{bmatrix} \frac{\partial w'_\nu}{\partial t} + \boldsymbol{\nabla}^\perp\psi'_\nu\cdot\boldsymbol{\nabla} w_\nu + \boldsymbol{\nabla}^\perp\psi_\nu\cdot\boldsymbol{\nabla} w'_\nu - \nu\boldsymbol{\Delta} w'_\nu \\ \boldsymbol{\Delta}\psi'_\nu + w'_\nu\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}, \tag{2.16a}$$

$$w'_\nu(t=0) = \phi', \tag{2.16b}$$

which is subject to the periodic boundary conditions and where ψ'_ν is the perturbation of the stream function ψ_ν. The Riesz representation theorem then allows us to write

$$\chi'_\nu(\phi;\phi') = \left\langle \nabla\chi_\nu(\phi),\phi'\right\rangle_{H^1(\Omega)} = \left\langle \nabla^{L^2}\chi_\nu(\phi),\phi'\right\rangle_{L^2(\Omega)}, \tag{2.17}$$

where the L^2 inner product is obtained by setting $\ell = 0$ in (2.13) and the Riesz representers $\nabla\chi_\nu(\phi)$ and $\nabla^{L^2}\chi_\nu(\phi)$ are the gradients of the objective functional computed with respect to the H^1 and L^2 topology, respectively. We remark that, while the H^1 gradient is used exclusively in the actual computations, cf. (2.14), the L^2 gradient is computed first as an intermediate step.

However, we note that expression (2.15) for the Gâteaux differential is not yet consistent with the Riesz form (2.17), because the perturbation ϕ' of the initial data (2.1c) does not appear in it explicitly as a factor, but is instead hidden as the initial condition in the linearized problem, cf. (2.16b). In order to transform (2.15) to the Riesz form, we introduce the *adjoint states* $w^*_\nu, \psi^*_\nu \ : \ [0,T]\times\Omega \to \mathbb{R}$ and the following duality-pairing relation

$$\left(\mathcal{K} \begin{bmatrix} w'_\nu \\ \\ \psi'_\nu \end{bmatrix}, \begin{bmatrix} w^*_\nu \\ \\ \psi^*_\nu \end{bmatrix} \right) := \int_0^T \int_\Omega \mathcal{K} \begin{bmatrix} w'_\nu \\ \\ \psi'_\nu \end{bmatrix} \cdot \begin{bmatrix} w^*_\nu \\ \\ \psi^*_\nu \end{bmatrix} d\mathbf{x}\, dt = 0. \tag{2.18}$$

Performing integration by parts with respect to both space and time in (2.18) and judiciously defining the *adjoint system* as (also subject to the period boundary conditions)

$$\mathcal{K}^* \begin{bmatrix} w^*_\nu \\ \\ \psi^*_\nu \end{bmatrix} := \begin{bmatrix} -\frac{\partial w^*_\nu}{\partial t} - \boldsymbol{\nabla}^\perp\psi_\nu \cdot \boldsymbol{\nabla} w^*_\nu + \psi^*_\nu - \nu\boldsymbol{\Delta} w^*_\nu \\ \\ \boldsymbol{\Delta}\psi^*_\nu - \boldsymbol{\nabla}^\perp \cdot (w^*_\nu \boldsymbol{\nabla} w_\nu) \end{bmatrix} = \begin{bmatrix} -\frac{2\nu}{T}\boldsymbol{\Delta} w_\nu \\ \\ 0 \end{bmatrix}, \tag{2.19a}$$

$$w^*_\nu(t = T) = 0, \tag{2.19b}$$

we arrive at

$$\begin{aligned}\left(\mathcal{K}\begin{bmatrix} w'_\nu \\ \\ \psi'_\nu \end{bmatrix}, \begin{bmatrix} w^*_\nu \\ \\ \psi^*_\nu \end{bmatrix}\right) &= \left(\begin{bmatrix} w'_\nu \\ \\ \psi'_\nu \end{bmatrix}, \mathcal{K}^*\begin{bmatrix} w^*_\nu \\ \\ \psi^*_\nu \end{bmatrix}\right) - \int_\Omega \phi'(\mathbf{x})\, w^*_\nu(0,\mathbf{x})\, d\mathbf{x} \\ &= \underbrace{-\frac{2\nu}{T}\int_0^T \int_\Omega w'_\nu \boldsymbol{\Delta} w_\nu \, d\mathbf{x}dt}_{\chi'_\nu(\phi;\phi')} - \int_\Omega \phi'(\mathbf{x})\, w^*_\nu(0,\mathbf{x})\, d\mathbf{x} = 0,\end{aligned} \tag{2.20}$$

where all boundary terms resulting from integration by parts with respect to the space variable vanish due to periodicity and one of the terms resulting from integration by parts with respect to time vanishes as well due to the terminal condition (2.19b). Identity (2.20) then implies $\chi'_\nu(\phi;\phi') = \int_\Omega \phi'(\mathbf{x})\, w^*_\nu(0,\mathbf{x})\, d\mathbf{x}$, from which we deduce the following expression for the L^2 gradient, cf. (2.17),

$$\nabla^{L^2}\chi_\nu(\mathbf{x}) = w^*_\nu(0,\mathbf{x}). \tag{2.21}$$

We note that the L^2 gradient does not possess the regularity required to solve Problem 1. Identifying the Gâteaux differential (2.15) with the H^1 inner product, cf. (2.13), integrating by parts and using (2.21), we obtain the required H^1 gradient $\nabla\chi$ as a solution of the elliptic boundary-value problem

$$\left[\mathrm{Id} - \ell^2\,\Delta\right]\nabla\chi_\nu = \nabla^{L^2}\chi_\nu \qquad \text{in } \Omega \tag{2.22}$$

subject to the periodic boundary conditions. As shown in Protas et al. (2004), extraction of gradients in spaces of smoother functions such as $H^1(\Omega)$ can be interpreted as low-pass filtering of the L^2 gradients with parameter ℓ acting as the cut-off length-scale. The value of ℓ can significantly affect the rate of convergence of the iterative procedure (2.14).

We define the inverse Laplacian on Ω such that it returns a zero-mean function. This ensures that the solution w_ν^* of the adjoint system (2.19) preserves the zero-mean property which is then also inherited by the L^2 and H^1 gradients, cf. (2.21)–(2.22). The projection operator in (2.14) is then defined in terms of the normalization (retraction)

$$\mathbb{P}_{\mathcal{S}}(\phi) = \sqrt{\frac{\mathcal{P}_0}{\mathcal{P}\left(\phi\right)}}\,\phi. \tag{2.23}$$

An optimal step size τ_n can be determined by solving the minimization problem

$$\tau_n = \arg\max_{\tau>0} \left\{ \chi_\nu \left(\mathbb{P}_{\mathcal{S}} \left(\phi^{(n)} + \tau\, \nabla \chi_\nu(\phi^{(n)}) \right) \right) \right\}, \tag{2.24}$$

which can be interpreted as a modification of a standard line search problem with optimization performed following an arc (a geodesic) lying on the constraint manifold $\mathcal{S}$, rather than a straight line.

To summarize, a single iteration of the gradient algorithm (2.14) requires solution of the Navier-Stokes system (2.1) followed by the solution of the adjoint system (2.19), which is a terminal-value problem and hence needs to be integrated backward in time, whereas its coefficients are determined by the solution of the Navier-Stokes system obtained before. These two solves allow one to evaluate the L^2 gradient via (2.21) which is then "lifted" to the space H^1 by solving (2.22). Finally, the approximation of the optimal initial condition $\breve{\phi}_\nu^T$ is updated using (2.14) with the step size τ_n determined in (2.24). As a first initial guess ϕ_0 in (2.14) we use the initial condition constructed in Jeong and Yoneda (2021) and then, to ensure the maximizers $\breve{\phi}_\nu^T$ obtained for the same viscosity ν but different lengths T of the time window lie on the same maximizing branch, we use a continuation approach where the maximizer $\breve{\phi}_\nu^T$ is employed as the initial guess ϕ_0 to compute $\breve{\phi}_\nu^{T+\Delta T}$ for some sufficiently small ΔT. In the same spirit, when searching for branches corresponding to different values of ν, the maximizer $\breve{\phi}_\nu^T$ is employed as the initial guess φ_0 to find $\breve{\varphi}_{\nu+\Delta\nu}^T$ for some sufficiently small $\Delta\nu$. We refer

the reader to (Ayala and Protas 2017) for further details of the continuation approach.

2.3.2 Computational Approach

Both the Navier-Stokes (2.1) and the corresponding adjoint system (2.19) are discretized in space using a standard Fourier pseudo-spectral method. Evaluation of nonlinear products and terms with nonconstant coefficients is performed using the 2/3 rule combined with a Gaussian filter defined by $\rho(\boldsymbol{k}) = e^{-36\left(\frac{|\boldsymbol{k}|}{K}\right)^{36}}$, where $\boldsymbol{k}$ is the wavenumber, $K = \frac{2N}{3}$ and N is the number of Fourier modes used in each direction (Hou 2009). Time integration is carried out using a four-step, globally third-order accurate mixed implicit/explicit Runge-Kutta scheme with low truncation error (Alimo et al. 2021). The results presented in the next section were obtained using the spatial resolutions $N = 512, 1024$ in each direction and the time-steps $\Delta t \approx 4.4721 \times 10^{-5}, 2.2361 \times 10^{-5}, 8.9443 \times 10^{-6}$, with finer resolutions employed for problems with smaller values of the viscosity ν. In system (2.22) defining the Sobolev gradients we set $\ell = 1$ and a spectral method is used to solve this system. The line-search problem (2.24) is solved with Brent's derivative-free algorithm (Press et al. 2007). Due to its large computational cost, a massively parallel implementation of the approach presented has been developed in FORTRAN 90 using the Message Passing Interface (MPI). Validation of the discretization techniques is discussed in Appendices A and B.

2.4 Results

In this section we present the results obtained by solving Problem 1 with $\mathcal{P}_0 = 1$ fixed and both ν and T varying over a broad range of values. In addition to understanding the structure of the flows maximizing the enstrophy dissipation and how it changes when the parameters are varied, our goal is also to provide insights which of the estimates (2.7)–(2.11) best describe the behavior of the maximum enstrophy dissipation $\chi_\nu(\breve{\phi}_\nu^T)$ in the limit of vanishing viscosity.

Problem 1 is nonconvex and as such admits multiple local maximizers at least for some values of ν and T. The results are organized in terms of "branches" defined as families of optimal initial conditions $\breve{\phi}_\nu^T$ obtained with fixed values of ν and varying T such that the maximum enstrophy dissipation $\chi_\nu(\breve{\phi}_\nu^T)$ is a smooth function of the length T of the time window. For each value of ν and each branch, the time windows considered are then chosen to capture the local maximum of $\chi_\nu(\breve{\phi}_\nu^T)$ and its neighbourhood. Information about the local maximizers found for $\nu = 2.2361 \times 10^{-6}$ and $T = 0.1789$ on six distinct branches is collected in Table 2.1 where we show the corresponding palinstrophy evolutions $\mathcal{P}(t)$, optimal initial conditions $\breve{\phi}_\nu^T(\mathbf{x})$ and the vorticity fields realizing the maximum palinstrophy $w_\nu(\mathbf{x}, \arg\max_{0<t\leq T} \mathcal{P}(t))$. These branches were determined for the given value of ν using the continuation approach described in Section 2.3.1 where χ_ν is regarded as a smooth function of T. When searching for branches corresponding to different viscosity values, continuation with respect to ν with T fixed was also used. The time evolution of the vorticity fields corresponding to all six branches is visualized in Video 1. This video offers insights about the different physical mechanisms involving the stretching of thin vorticity filaments which are responsible for the growth of palinstrophy and hence also increased enstrophy dissipation. It is noteworthy that all these flow evolutions feature very thin filaments which however do not undergo the Kelvin-Helmholtz instability as they are stabilized by the shear induced by the large vortices also present in the flow field. This effect is well understood in the idealized setting of inviscid flows with straight vortex sheets (Kiya and Arie 1979), see also (Sakajo and Okamoto 1996), and was observed in 2D turbulence (Kevlahan and Farge 1997). Flows on branches 3 and 4, which feature multiple palinstrophy maxima, employ a mechanism similar to the continuous baker's map to amplify the palinstrophy. Moreover, we see that, interestingly, in some cases seemingly very similar optimal initial conditions $\breve{\phi}_\nu^T$ give rise to quite different flow evolutions featuring different numbers of local palinstrophy maxima (one or two) in the considered time window $[0, T]$, see, e.g., the maximizers from Branches

2 and 3 in Table 2.1. This makes classifying local optimizers into branches a rather difficult task and the classification presented in Table 2.1 is tentative only, which will however not affect the main findings of our study. Video 2 and Video 3 show the flow evolutions and representative palinstrophy histories corresponding to the locally optimal initial conditions $\breve{\phi}_\nu^T$ obtained, respectively, on Branch 1 with $T = 0.1207$ and on Branch 5 with $T = 0.2683$ for five different values of the viscosity ν. In both cases we see that even though the optimal initial conditions $\breve{\phi}_\nu^T$ obtained for different values of ν are quite similar, qualitative changes occur in the flows evolutions as the viscosity is reduced. We attribute these changes to either possible bifurcations of the branches (understood as functions of ν) or to the possibility that the flow evolutions corresponding to smaller viscosity values belong to some unclassified branch, underpinning the difficulty mentioned above.

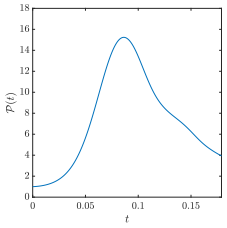
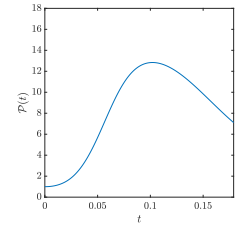
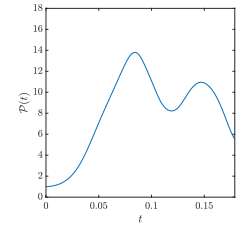
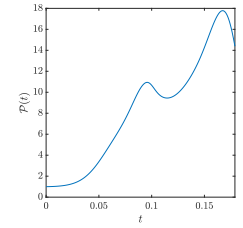
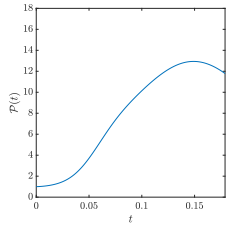
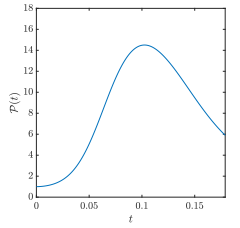
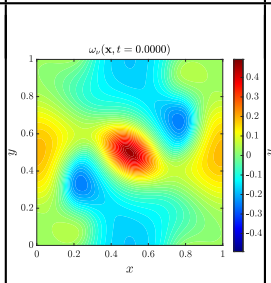
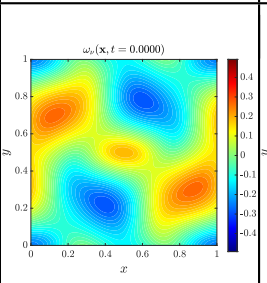
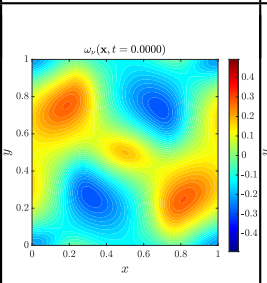
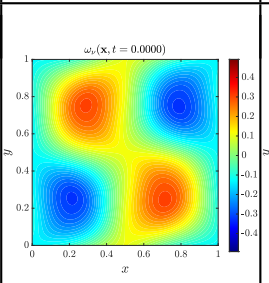
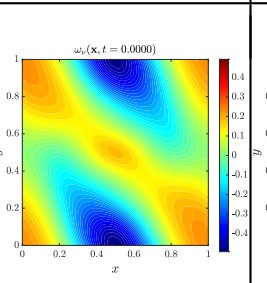
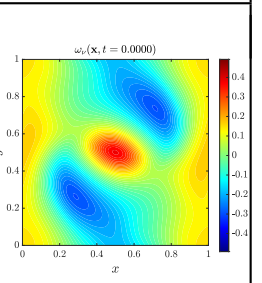
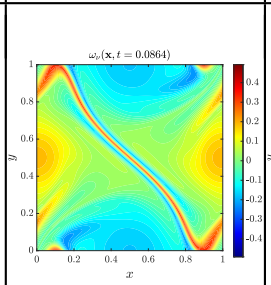
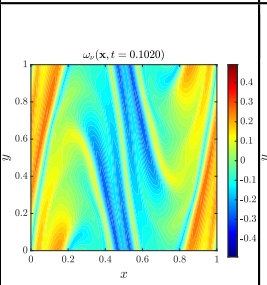
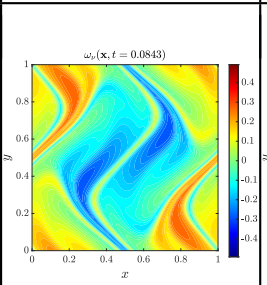
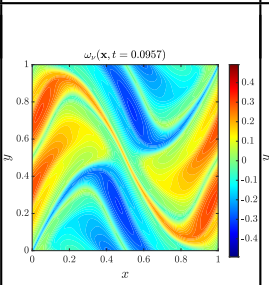
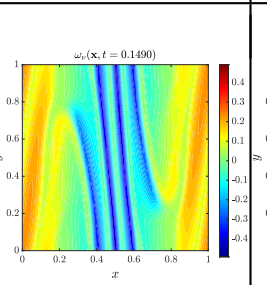
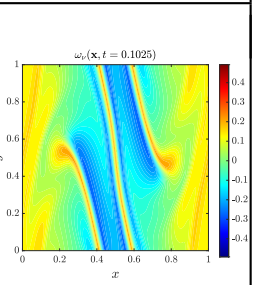
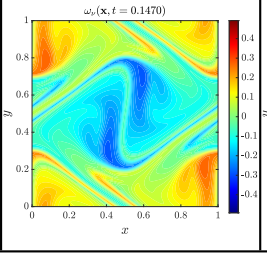
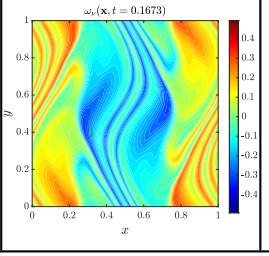
Branch	1	2	3	4	5	6
Palinstrophy						
Initial Condition						
Palinstrophy Peak 28						
Palinstrophy Peak 2	N/A	N/A			N/A	N/A

TABLE 2.1: Summary information about the local maximizers obtained by solving Problem 1 with $\nu = 2.2361\times10^{-6}$ and $T = 0.1789$. The time evolution of the vorticity fields is visualized in Video 1. "N/A" indicates that palinstrophy attains a single maximum only during the time evolution corresponding to the given branch.

Next, in Figure 2.1a we show the dependence of the maximum enstrophy dissipation $\chi_\nu(\breve{\phi}_\nu^T)$ on the length T of the time window for five values of viscosity spanning more than one order of magnitude. We carefully distinguish branches of distinct local maximizers, where by a "branch" we mean a family of optimal initial data $\breve{\phi}_\nu^T$ parametrized by T and such that the enstrophy dissipation $\chi_\nu(\breve{\phi}_\nu^T)$ changes smoothly as T is varied while ν remains fixed. We remark that for certain combinations of ν and T only a subset of the local maximizers described in Table 2.1 could be found. In Figure 2.1a we observe that along each branch the maximum enstrophy dissipation $\chi_\nu(\breve{\phi}_\nu^T)$ admits a well-defined maximum with respect to T. We add that the values of $\chi_\nu(\breve{\phi}_\nu^T)$ shown in Figure 2.1a are for each value of ν at least an order of magnitude larger than the enstrophy dissipation corresponding to the initial conditions constructed in Jeong and Yoneda (2021), which realize the behavior given in (2.12).

As these are the quantities needed to make quantitative comparisons with estimates (2.7)–(2.11), in Figure 2.1b we plot the "envelopes", defined as $\breve{\chi}_\nu^T := \max_{\text{branches}} \chi_\nu(\breve{\phi}_\nu^T)$, of the branches obtained at fixed values of ν. "Singularities" evident in these curves correspond to values of T where different branches become dominant as T varies.

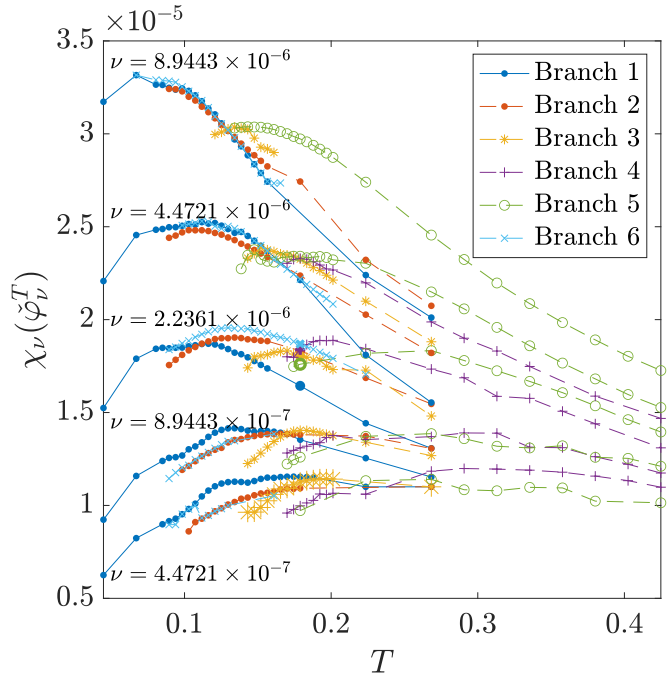
Next, we move on to identify quantitative connections between the data presented in Figure 2.1b and estimates (2.7)–(2.11) describing the vanishing of the enstrophy dissipation in the inviscid limit $\nu \to 0$. These estimates also depend on the length T of the time window, but this dependence is in some cases less explicit and we will therefore consider T as a fixed parameter. We thus introduce the following ansätze

$$f_1(\nu) = C\ [-\ln(\nu)]^{-\frac{1}{2}}\,, \tag{2.25a}$$

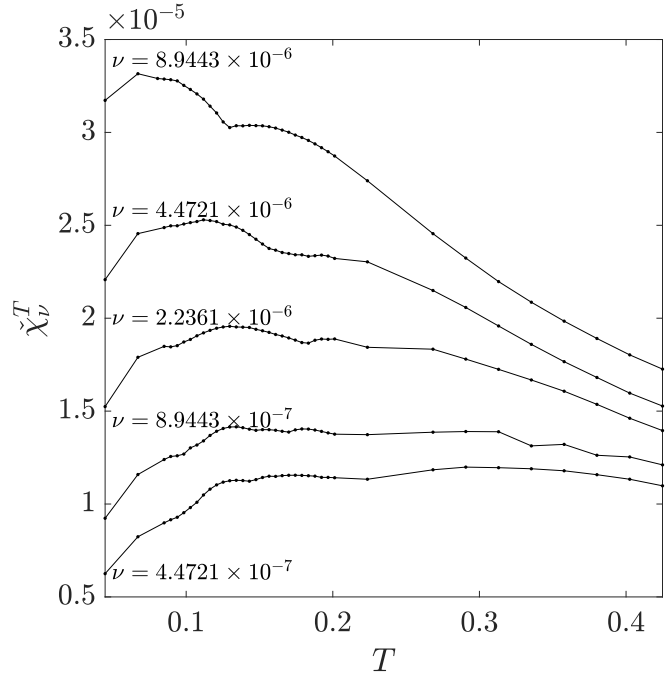
$$f_2(\nu) = C\,\nu^\beta, \tag{2.25b}$$

$$f_3(\nu) = C\ \left[\frac{\nu}{|\ln(\nu)|}\right]^\beta, \tag{2.25c}$$

$$f_4(\nu) = C\,\nu\ [-\ln(\nu)]^{\frac{1}{2}}\,, \tag{2.25d}$$



(a)



(b)

FIGURE 2.1: Dependence of (a) the maximum enstrophy dissipation $\chi_\nu(\breve{\phi}_\nu^T)$ for maximizers on the different branches, cf. Table 2.1, and (b) its envelope $\breve{\chi}_\nu^T$ on the length T of the time window for different viscosities ν. In panel (a) the local maximizers illustrated in Table 2.1 are marked with larger symbols.

motivated by the structure of the different estimates. More specifically, (2.25a) is the expression from Conjecture 1, cf. (2.7), (2.25b) has the general form of the upper bounds in (2.9)–(2.10), where in the latter case we only consider the second argument of the function $\max(\cdot)$ since the function ϕ_{ϕ,p,M_∞} appearing in the first argument is not given explicitly enough to allow for quantitative comparisons, (2.25c) is motivated by the form of estimate (2.11) whereas (2.25d) is the bound from Theorem 1, cf. (2.12).

We want to find out which of the functions (2.25a)–(2.25d) best describes the dependence of the data shown in Figure 2.1b on ν for different fixed values of T. For each discrete value of T (marked with solid symbols in Figure 2.1b) we determine the constant $C = C(T)$ in each of the ansatz functions (2.25a)–(2.25d) by solving the problem

$$\dot{C}(T) = \underset{C\in\mathbb{R}^+}{\arg\min}\, \mu_i^T(C), \qquad i = 1, 2, 3, 4, \tag{2.26}$$

with the fitting error defined as

$$\mu_i^T(C) := \frac{1}{5}\sum_{j=1}^{5} \left| \breve{\chi}_{\nu_j}^T - f_i(\nu_j; C) \right|, \tag{2.27}$$

where $\nu_j \in \{8.9443\times10^{-6}, 4.4721\times10^{-6}, 2.2361\times10^{-6}, 8.9443\times10^{-7}, 4.4721\times10^{-7}\}$ are the considered values of the viscosity. The fitting error (2.27) will be used to measure the accuracy with which the different ansatz functions (2.25a)–(2.25d) represent the data. In addition, we note that ansatz functions (2.25b)–(2.25c) also involve a priori undefined exponents $\alpha \in (0, 1)$. To determine this additional parameter in f_2 and f_3, we embed problem (2.26) in a bracketing procedure which finds the exponent $\dot{\beta} = \dot{\beta}(T)$ producing the smallest fitting error (2.27) for a given value of T. This bracketing procedure is performed by first determining $\mu_i^T(\dot{C}(T))$, by solving problem (D.4), for a range of discrete values of $\beta \in [0, 1]$ and then using bisection to iteratively improve the approximation of $\dot{\beta} = \dot{\beta}(T)$ which produces the smallest error (2.27). We emphasize that even though the ansätze (2.25a)–(2.25d) involve different numbers of parameters (one

or two), they are all fitted to the data in Figure 2.1b in the same way (i.e., by adjusting $C = C(T)$), which is done independently for different discrete exponents β in the case of relations (2.25b)–(2.25c).

In order to assess now well the different ansatz functions (2.25a)–(2.25d) capture the dependence of the maximum enstrophy dissipation $\breve{\chi}_\nu^T$ on ν, cf. Figure 2.1b, we define the ratios $\breve{\chi}_\nu^T / f_i(\nu)$, $i = 1, 2, 3, 4$, and plot them as functions of ν for different T in Figures 2.2a–d using the values of $\dot{C} = \dot{C}(T)$ and $\dot{\beta} = \dot{\beta}(T)$ determined as above. Thus, if $\breve{\chi}_\nu^T / f_i(\nu)$ is close to unity over the entire range of ν, this signals that the ansatz function $f_i(\nu)$ accurately captures the dependence of $\breve{\chi}_\nu^T$ on ν for the given value of T. We see that this is what indeed happens for $f_2(\nu)$ and $f_3(\nu)$ for most values of T, cf. Figures 2.2b,c. On the other hand, we note that relations $f_1(\nu)$ and $f_4(\nu)$, respectively, overestimate and underestimate the actual dependence of $\breve{\chi}_\nu^T$ on ν, cf. Figures 2.2a,d. This observation is consistent with the fact that (2.25a) represents estimate (2.7), which is more conservative than bounds (2.9)–(2.11), and (2.25d) has the form of the lower bound (2.12).

Hereafter we will focus on the fits given in terms of ansatz functions (2.25b)–(2.25c). In order to decide which of these relations more accurately represents the dependence of $\breve{\chi}_\nu^T$ on ν, in Figure 2.3 we show the corresponding fitting errors (2.27) as functions of T. Smaller fitting errors $\mu_i^T(\dot{C}(T))$ indicate a better fit between the ansatz function f_2 or f_3 and the data $\breve{\chi}_\nu^T$ for the given value of T. We see that relation $f_2(\nu)$ generally leads to smaller errors for shorter time windows (with $T \leq 0.147$), whereas relation $f_3(\nu)$ tends to better predict the dependence of $\breve{\chi}_\nu^T$ on ν for longer time windows. Finally, the optimal exponents $\dot{\beta} = \dot{\beta}(T)$ determined for ansatz functions (2.25b)–(2.25c) are shown in Figure 2.4 where an overall decreasing trend with T is evident. As regards the "dip" occurring for $0.0894 \lessapprox T \lessapprox 0.1342$, we speculate that it may be the result of some branches not being captured in Figure 2.1a. We note that, remarkably, the dependence of the exponent $\dot{\beta}$ on T reveals an approximately exponential form consistent with the structure

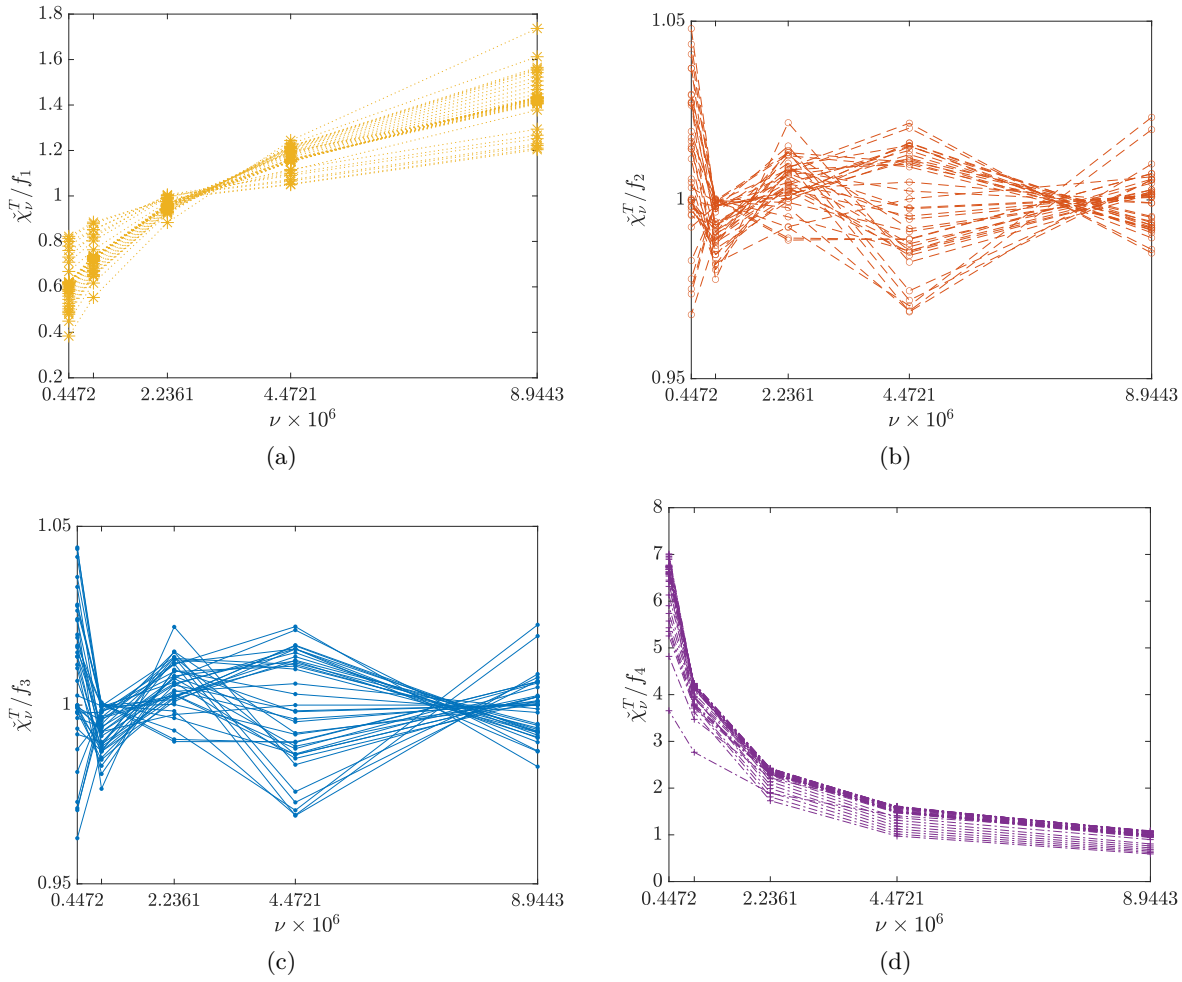


FIGURE 2.2: Dependence of (a) $\breve{\chi}_\nu^T/f_1(\nu)$, (b) $\breve{\chi}_\nu^T/f_2(\nu)$, (c) $\breve{\chi}_\nu^T/f_3(\nu)$ and (d) $\breve{\chi}_\nu^T/f_4(\nu)$, with optimal constants $\dot{C} = \dot{C}(T)$ and exponents $\dot{\beta} = \dot{\beta}(T)$, on the viscosity ν for different T.

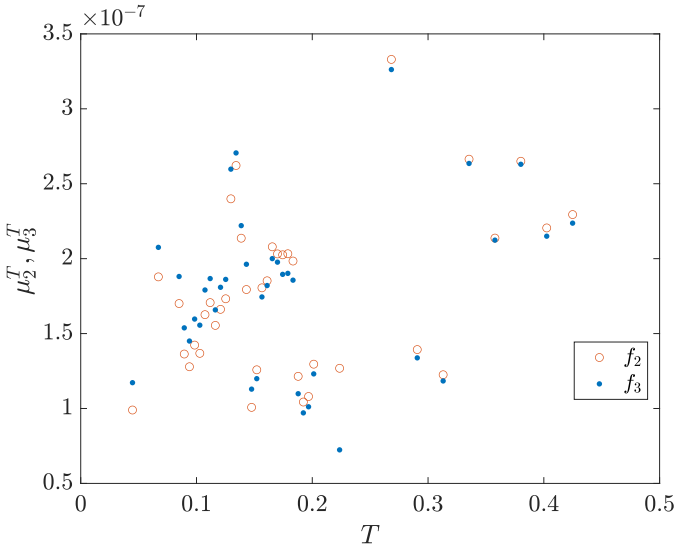


FIGURE 2.3: Dependence of the fitting errors $\mu_i^T(\dot{C}(T))$, $i = 2, 3$, cf. (2.27), corresponding to the fits of ansatz functions (red circles) $f_2(\nu)$ and (blue dots) $f_3(\nu)$ to $\widetilde{\chi}_\nu^T$ for different T.

of the upper bounds in (2.10)–(2.11), more specifically, the exponential dependence of the exponents of ν in these bounds on T. Moreover, the limit $\lim_{T\to 0}\dot{\beta}(T)$ is also quantitatively consistent with predictions of estimate (2.11).

2.5 Summary and Conclusions

In this study we provide a quantitative characterization of the behaviour of the enstrophy dissipation in 2D Navier-Stokes flows in the limit of vanishing viscosity. Unlike the case of Burgers flows in 1D and Navier-Stokes flows in 3D, where the energy dissipation anomaly is well documented, 2D Navier-Stokes flows are known not to exhibit anomalous behavior of enstrophy dissipation. As discussed in Section 2.1, the vanishing of enstrophy dissipation in the inviscid limit is subject to various estimates, some ad-hoc and some rigorous, providing lower and upper bounds on this quantity as viscosity vanishes. In our investigation we have probed the sharpness of these estimates by constructing families of Navier-Stokes flows designed to locally maximize the enstrophy dissipation subject to certain constraints. This was done by solving Problem 1 where locally optimal initial

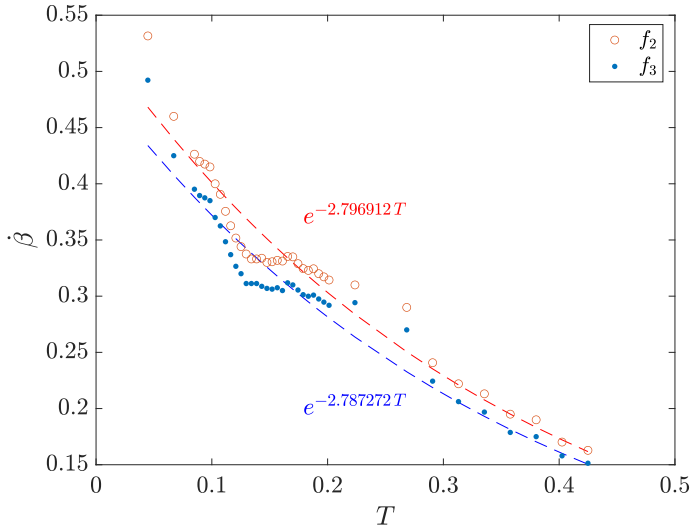


FIGURE 2.4: Dependence of the optimal exponents $\dot{\beta} = \dot{\beta}(T)$ in the ansatz functions (red circles) $f_2(\nu)$ and (blue dots) $f_3(\nu)$ on the length T of the time window. The dashed lines represent exponential fits, in the forms indicated, to the values of $\dot{\beta} = \dot{\beta}(T)$ for the ansatz function (red) $f_2(\nu)$ and (blue) $f_3(\nu)$.

data $\breve{\phi}_\nu^T$ with fixed palinstrophy $\mathcal{P}_0$ was found such that the corresponding flow with the given viscosity ν maximizes the enstrophy dissipation χ_ν over the time window $[0, T]$. Problem 1 was solved numerically using a state-of-the-art adjoint-based gradient ascent method described in Section 2.3. This optimization problem is nonconvex and we have found six distinct branches of local maximizers, each associated with a different mechanism for palinstrophy amplification, cf. Table 2.1. As is evident from Video 1, while in all cases palinstrophy amplification involves stretching of thin vorticity filaments, there are multiple ways to design flows maximizing this process on a periodic domain Ω and which of these different mechanisms produces the largest enstrophy dissipation depends on the value of viscosity ν and the length T of the time window, cf. Figure 2.1a.

Branches of local maximizers found by solving Problem 1 for different values of ν and T reveal how the extreme behaviour of the enstrophy dissipation they realize compares with the available estimates on this process discussed in Section 2.1. We conclude that the dependence of the maximum enstrophy dissipation $\breve{\chi}_\nu^T$ in the extreme flows we

found on ν with fixed T is quantitatively consistent with the upper bound in estimate (2.10), cf. Figure 2.2b, which is the sharpest estimate available to date. Remarkably, the exponential dependence of the exponent in this upper bound on T is also quantitatively consistent with our results, cf. Figure 2.4 (we attribute the deviation from the exponential decrease evident around $T \approx 0.1342$ in this figure to the likely possibility that, despite our efforts, not all branches of maximizing solutions have been found).

As regards estimate (2.10), we note that it depends on the quantity $\|\breve{\phi}_\nu^T\|_{L^\infty(\Omega)}$ (via the constant M_∞). Since our optimal initial conditions are sought in the space $H^1(\Omega)$, we do not have an a priori control over this quantity, however, in our computations we did not find any evidence for $\|\breve{\phi}_\nu^T\|_{L^\infty(\Omega)}$ to attain large values. Thus, these caveats notwithstanding, we conclude that estimate (2.10) is sharp and does not offer any room for improvement, other than perhaps a logarithmic correction analogous to the one appearing in (2.11). Relation (2.11) was found to describe the dependence of the maximum enstrophy dissipation $\breve{\chi}_\nu^T$ on viscosity in the limit $\nu \to 0$ with similar accuracy to estimate (2.10). However, we reiterate that, as discussed in Section 2.1, relation (2.11) does not represent a rigorous upper bound on the enstrophy dissipation. Improving this estimate, so that the $\dot{H}^{-1}(\Omega)$ norm on the left-hand side in (2.11) is strengthened to $L^2(\Omega)$, appears to be an open question in mathematical analysis.

Among other open problems, it would be interesting to better understand the bifurcation structure of the different optimal solution branches shown in Figure 2.1a. Another open question is what new insights about the problem considered here could be deduced based on the kinetic theory, i.e., by considering an optimization problem analogous to Problem 1 in the context of the Boltzmann equation or some of its variants. Some efforts in this direction are already underway, which is briefly discussed in Chapter 5. Finally, there is the question about what can be said about the energy dissipation anomaly in 3D Navier-Stokes flows using the approach developed in the present study.

Chapter 3

Optimal Eddy Viscosity in Closure Models for 2D Turbulent Flows

3.1 Introduction

The closure problem is arguably one of the most important outstanding open problems in turbulence research. It touches upon some of the key basic questions concerning turbulent flows and at the same time has far-reaching consequences for many applications, most importantly, for how we simulate turbulent flows in numerous geophysical, biological and engineering settings. Given the extreme spatio-temporal complexity of turbulent flows, accurate numerical solutions of the Navier-Stokes system even at modest Reynolds numbers requires resolutions exceeding the capability of commonly accessible computational resources. To get around this difficulty, one usually relies on various simplified versions of the Navier-Stokes system obtained through different forms of averaging and/or filtering, such as the Reynolds-Averaged Navier-Stokes (RANS) system and the Large-Eddy Simulation (LES). However, such formulations are not closed, because these

systems involve nonlinear terms representing the effect of unresolved subgrid stresses on the resolved variables. The "closure problem" thus consists in expressing these quantities in terms of resolved variables such that the RANS or LES system is closed.

In general, closure models in fluid mechanics are of two main types: algebraic, where there is an algebraic relationship expressing the subgrid stresses in terms of the resolved quantities, and differential, where this relationship involves an additional partial-differential equation (PDE) which needs to be solved together with the RANS or LES system. Most classical models are usually formulated based on some ad-hoc, albeit well-justified, physical assumptions. There exists a vast body of literature concerning the design, calibration and performance of such models in various settings. Since it is impossible to offer an even cursory survey of these studies here, we refer the reader to the well-known monographs (Lesieur 1993; Pope 2000; Davidson 2015) for an overview of the subject. Recently, there has been a lot of activity centered on learning new empirical closure models from data using methods of machine learning (Kutz 2017; Gamahara and Hattori 2017; Jimenez 2018; Duraisamy et al. 2019; Duraisamy 2021; Pawar and San 2021). It is however fair to say that the field of turbulence modelling has been largely dominated by empiricism and there is a consensus that the potential and limitations of even the most common models are still not well understood. Our study tackles this fundamental question, more specifically, how well certain common closure models can in principle perform if they are calibrated in a optimal way. We will look for an optimal, in a mathematically precise sense, form of a certain closure model and will conclude that, somewhat surprisingly, it does not in fact exist.

On the other hand, from the physical point of view, turbulence closure models are not meant to capture nonlinear transfer processes with pointwise accuracy, but rather to represent them in a certain average sense, for example, ensuring the inertial range of the energy spectrum is well represented, or that a quantity of interest such as drag, is

correctly captured. The ill-posedness of the problem of optimally calibrating a closure model signalled above can thus be viewed as a consequence of the inability of the closure model to match the original solution pointwise in space and in time. More precisely, the optimal eddy viscosity exhibits unphysical high-frequency oscillations. In the present study we will use a novel and mathematically systematic approach to illustrate this physical intuition and demonstrate how the ill-posedness arises. We will also show that the model calibration problem is in fact well-behaved when the LES with a closure model is required to match quantities defined in the statistical rather than pointwise sense (here, we consider the enstrophy spectrum).

We are going to focus on an example from a class of widely used algebraic closure models, namely, the Smagorinsky-type eddy-viscosity models (Smagorinsky 1963) for LES. More specifically, we will consider the Leith model (Leith 1968; Leith 1971; Leith 1996) for two-dimensional (2D) turbulent flows. Like all eddy-viscosity closure models, the Leith model depends on one key parameter which is the eddy viscosity, typically taken to be a function of some flow variable. Needless to say, performance of such models critically depends on the form of this function. One specific question we are interested in is how accurately the LES equipped with such an eddy-viscosity closure model can at best reproduce solutions of the Navier-Stokes system obtained via Direct Numerical Simulation (DNS). Another related question we will consider concerns reproducing certain statistical properties of Navier-Stokes flows in LES. We will address these questions by formulating them as PDE-constrained optimization problems where we will seek an optimal functional dependence of the eddy viscosity on the state variable. In the first problem we will require the corresponding LES to match the filtered DNS pointwise in space over a time window of several eddy turnover times, whereas in the second problem the LES will be required to match the time-averaged enstrophy spectrum of the Navier-Stokes flow for small wavenumbers. By framing these questions in terms of optimization problems we will be able to find the best (in a mathematically

precise sense) eddy viscosities, and this will in turn allow us to establish ultimate performance limitations for this class of closure models. We emphasize that the novelty of our approach is that by finding an optimal functional form of the eddy viscosity we identify, subject to minimum assumptions, an optimal structure of the nonlinearity in the closure model, which is fundamentally different, and arguably more involved, than calibrating one or more constants in a selected ansatz for the eddy viscosity. This formulation is also more general than common dynamic closure models and some formulations employing machine learning to deduce information about local properties of closure models from the DNS (see, e.g., (Maulik et al. 2020)). Our goal is to understand what form the eddy viscosity needs to take in order to maximize the performance of the closure model in achieving a prescribed objective. The emphasis will be on methodology rather than on specific contributions to subgrid modeling.

The optimization problem in question has a non-standard structure, but an elegant solution can be obtained using a generalization of the adjoint-based approach developed by Bukshtynov et al. (2011) and Bukshtynov and Protas (2013). In being based on methods of the calculus of variations, this approach thus offers a mathematically rigorous alternative to machine-learning methods which have recently become popular (Kutz 2017; Gamahara and Hattori 2017; Jimenez 2018; Duraisamy et al. 2019; Duraisamy 2021; Pawar and San 2021). As a proof of the concept applicable to the problem considered here, this approach was recently adapted to find optimal closures in a simple one-dimensional (1D) model problem by Matharu and Protas (2020). Importantly, this approach involves a regularization parameter controlling the "smoothness" of the obtained eddy viscosity.

In the first problem, which involves matching the filtered DNS solution in the pointwise sense, we find optimal eddy viscosities for the Leith closure model in the LES systems with different filter cutoff wavenumbers k_c. As this wavenumber increases and

the filter width vanishes, the optimal eddy viscosity is close to zero and the match between the predictions of the LES and the filtered DNS is nearly exact, as expected. On the other hand, for smaller cutoff wavenumbers k_c the optimal eddy viscosity becomes highly irregular whereas the match between the LES and DNS deteriorates, although it still remains much better than the match involving the LES with the standard Leith model or with no closure model at all. Interestingly, the optimal eddy viscosity reveals highly oscillatory behavior with alternating positive and negative values as the state variable increases. When the regularization in the solution of the optimization problems is reduced and the numerical resolution is refined at a fixed cutoff wavenumber, the frequency and amplitude of these oscillations are amplified which results in an improved match against the DNS. Thus, in this limit the optimal eddy viscosity becomes increasingly oscillatory as a function of the state variable which suggests that in the absence of regularization the problem of finding an optimal eddy viscosity does not in fact have a solution as the limiting eddy viscosity is not well defined. On the other hand, an arbitrarily regular eddy viscosity can be found when sufficient regularization is used in the solution of the optimization problem, but at the price of reducing the match against the DNS. While such smooth eddy viscosities may be more useful in practice, the corresponding LES models will not achieve their theoretical performance limits. In addition to this observation, our results also demonstrate how the best accuracy achievable by the LES with the considered closure model depends on the cutoff wavenumber of the filter, which sheds new light on the fundamental performance limitations inherent in this closure model.

In our second problem, which involves matching the time-averaged vorticity spectrum of the filtered DNS, the obtained optimal eddy viscosity is more regular and its key features remain essentially unchanged as the regularization in the solution of the optimization problems is reduced and the numerical resolution is refined. This demonstrates that the problem of optimally calibrating the closure model is better behaved

when a suitable statistical quantity is used as the target. This is not surprising as such a formulation is in fact closer to the spirit of turbulence modelling.

The structure of the chapter is as follows: in the next section we formulate our LES model and state the optimization problem defining the optimal eddy viscosity; in Section 3.3 we introduce an adjoint-based approach to the solution of the optimization problem and in Section 3.4 discuss computational details; our results are presented in Section 3.5 whereas final conclusions are deferred to Section 3.6 ; some additional technical material is provided in Appendix C.

3.2 Large-Eddy Simulation and Optimal Eddy Viscosity

We consider 2D flows of viscous incompressible fluids on a periodic domain $\Omega := [0, 2\pi]^2$ over the time interval $[0, T]$ for some $T > 0$ (":=" means "equal to by definition"). Assuming the fluid is of uniform unit density $\rho = 1$, its motion is governed by the Navier-Stokes system written here in the vorticity form

$$\partial_t w + \boldsymbol{\nabla}^\perp \psi \cdot \boldsymbol{\nabla} w = \nu_N \boldsymbol{\Delta} w - \alpha w + f_\omega \quad \text{in} \quad (0, T] \times \Omega, \tag{3.1a}$$

$$\boldsymbol{\Delta} \psi = -w \quad \text{in} \quad (0, T] \times \Omega, \tag{3.1b}$$

$$w(t = 0) = w_0 \quad \text{in} \quad \Omega, \tag{3.1c}$$

where $w = -\boldsymbol{\nabla}^\perp \cdot \boldsymbol{u}$, with $\boldsymbol{\nabla}^\perp = [\partial_{x_2}, -\partial_{x_1}]^T$ and $\boldsymbol{u}$ the velocity field, is the vorticity component perpendicular to the plane of motion, ψ is the streamfunction, ν_N is the coefficient of the kinematic viscosity (for simplicity, we reserve the symbol ν for the eddy viscosity), and w_0 is the initial condition. System (3.1) is subject to two forcing mechanisms: a time-independent forcing f_ω which ensures that the flow remains in a statistical equilibrium and the Ekman friction $-\alpha w$ describing large-scale dissipation due to, for example, interactions with boundary layers arising in geophysical fluid phenomena. The

forcing term is defined to act on Fourier components of the solution with wavenumbers in the range $[k_a, k_b]$ for some $0 < k_a < k_b < \infty$, i.e.,

$$\left[\widehat{f_\omega}\right]_{\boldsymbol{k}} := \begin{cases} F, & k_a \leq |\boldsymbol{k}| \leq k_b, \\ 0, & \text{otherwise}, \end{cases} \tag{3.2}$$

where $\left[\widehat{f_\omega}\right]_{\boldsymbol{k}}$ is the Fourier component of f_ω with the wavevector $\boldsymbol{k}$ (hereafter hats " $\widehat{\cdot}$ " will denote Fourier coefficients) and $F > 0$ is a constant parameter.

The phenomenology of 2D forced turbulence is described by the Kraichnan-Batchelor-Leith theory (Kraichnan 1967; Batchelor 1969; Leith 1968) which makes predictions about various physical characteristics of such flows. Their prominent feature, distinct from turbulent flows in three dimensions (3D), is the presence of a forward enstrophy cascade and an inverse energy cascade (Boffetta 2007; Boffetta and Musacchio 2010; Bracco and McWilliams 2010; Vallgren and Lindborg 2011; Boffetta and Ecke 2012). Here we will chose k_a and k_b such that the forcing term (3.2) will act on a narrow band of Fourier coefficients to produce a well-developed enstrophy cascade towards large wavenumbers and a rudimentary energy cascade towards small wavenumbers. The parameters ν_N, α and F will be adjusted to yield a statistically steady state with enstrophy $\mathcal{E}(t) := \int_\Omega w^2(t, \mathbf{x})\, d\Omega$ fluctuating around a well-defined mean value $\mathcal{E}_0$. The initial condition ω_0 in (3.1c) will be chosen such that the evolution begins already in this statistically steady state at time $t = 0$.

3.2.1 The Leith Closure Model

The LES is obtained by applying a suitable low-pass filter G_δ, where $\delta > 0$ is its width, to the Navier-Stokes system (3.1) and defining the filtered variables $\widetilde{w} = G_\delta * w$ and $\widetilde{\psi} = G_\delta * \psi$ (" $*$ " denotes the convolution operation and hereafter we will use tilde " $\widetilde{\cdot}$ " to represent filtered variables). For simplicity, we will employ a sharp low-pass spectral

filter defined in terms of its Fourier-space representation as

$$\left[\widehat{G}_\delta\right]_{\boldsymbol{k}} := \begin{cases} 1, & |\boldsymbol{k}| \leq k_c, \\ 0, & \text{otherwise}, \end{cases} \tag{3.3}$$

where k_c is the largest resolved wavenumber such that the filter width is $\delta = 2\pi/k_c$. Since we normally have $k_b < k_c$, it follows that $\widetilde{f_\omega} = f_\omega$. Application of filter (3.3) to the vorticity equation (3.1a) yields $\partial_t \widetilde{w} + \widetilde{\boldsymbol{\nabla}^\perp \widetilde{\psi} \cdot \boldsymbol{\nabla} w} = \nu_N \boldsymbol{\Delta} \widetilde{w} - \alpha \widetilde{w} + f_\omega + M$, where the term M represents the effect of the unresolved subgrid quantities

$$M = \widetilde{\boldsymbol{\nabla}^\perp \widetilde{\psi} \cdot \boldsymbol{\nabla} w} - \widetilde{\boldsymbol{\nabla}^\perp \psi \cdot \boldsymbol{\nabla} w}. \tag{3.4}$$

Since expression (3.4) depends on the original unfiltered variables w and ψ, to close the filtered system the term M must be modelled in terms of an expression involving the filtered variables only. We will do this using the Leith model (Leith 1968; Leith 1971; Leith 1996), which has a similar structure to the Smagorinsky model (Smagorinsky 1963) widely used as a closure for 3D flows, but is derived considering the forward enstrophy cascade as the dominant mechanism in 2D turbulent flows. There is evidence for good performance of the Leith model in such flows (Graham and Ringler 2013; Maulik and San 2017). Its preferred form is

$$M \approx \widetilde{M} = \boldsymbol{\nabla} \cdot (\widetilde{\nu_L \boldsymbol{\nabla} \widetilde{\omega}}), \tag{3.5}$$

in which $\widetilde{\omega}$ is the solution to the LES system, cf. (3.8), and the eddy viscosity is assumed to be a linear function of the magnitude of the vorticity gradient, i.e.,

$$\nu_L(s) := (C_L \delta)^3 \sqrt{s} \qquad \text{with} \quad s := |\boldsymbol{\nabla} \widetilde{\omega}|^2 \in \mathcal{I} := [0, s_{\max}], \tag{3.6}$$

where the Leith constant $C_L = 1$ and $s_{\max} > 0$ is a sufficiently large number to be specified later. We will refer to $\mathcal{I}$ as the "state space" domain.

While in the original formulation of the Leith model the eddy viscosity is taken to be a linear function of $|\boldsymbol{\nabla}\widetilde{\omega}|$ as in (3.6) (Graham and Ringler 2013; Maulik and San 2017), here we consider a general dependence of the eddy viscosity on $|\boldsymbol{\nabla}\widetilde{\omega}|$ in the form

$$\nu(s) = [\nu_L(s) + \nu_0]\, \varphi\left(\frac{s}{s_{\max}}\right), \tag{3.7}$$

where $\nu_0 > 0$ and $\varphi : [0,1] \to \mathbb{R}$ is a dimensionless function subject to some minimum only assumptions to be specified below. The parameter ν_0 is introduced to allow the eddy viscosity $\nu(s)$ to take nonzero values at $s = 0$, in contrast to Leith's original model (3.6). We remark that defining the eddy viscosity in terms of such a function φ ensures that ansatz (3.7) is dimensionally consistent. Making φ and ν functions of $|\boldsymbol{\nabla}\widetilde{\omega}|^2$, rather than of $|\boldsymbol{\nabla}\widetilde{\omega}|$, in (3.7) will simplify subsequent calculations. We add that ansatz (3.7) is used here to illustrate the approach and in principle one could also consider other formulations parametrized by nondimensional functions. With the Leith model (3.5)–(3.7), the LES version of the 2D Navier-Stokes system (3.1) takes the form

$$\partial_t \widetilde{\omega} + \widetilde{\boldsymbol{\nabla}^{\perp}\widetilde{\psi} \cdot \boldsymbol{\nabla}\widetilde{\omega}} = \boldsymbol{\nabla} \cdot \widetilde{([\nu_N + \nu(s)]\, \boldsymbol{\nabla}\widetilde{\omega})} - \alpha\widetilde{\omega} + f_\omega \quad \text{in} \quad (0,T] \times \Omega, \tag{3.8a}$$

$$\boldsymbol{\Delta}\widetilde{\psi} = -\widetilde{\omega} \quad \text{in} \quad (0,T] \times \Omega, \tag{3.8b}$$

$$\widetilde{\omega}(t=0) = \widetilde{\omega}_0 := \widetilde{w}_0 \quad \text{in} \quad \Omega, \tag{3.8c}$$

where the initial condition is given as the filtered initial condition (3.1c) from the DNS system.

An equivalent form of equation (3.8a) can be obtained noting that with the form of the filter given in (3.3), the decomposition of the subgrid stresses (3.4) reduces to $M = \boldsymbol{\nabla}^{\perp}\widetilde{\psi} \cdot \boldsymbol{\nabla}\widetilde{\omega} - \widetilde{\boldsymbol{\nabla}^{\perp}\psi \cdot \boldsymbol{\nabla}\omega}$ (Pope 2000). As discussed in Appendix D, the advection

term in (3.8a) can be replaced with $\boldsymbol{\nabla}^{\perp}\widetilde{\psi} \cdot \boldsymbol{\nabla}\widetilde{\omega}$. In addition, in Appendix E we show both the 2D vorticity equation (3.1) and the corresponding LES equation with the Leith model (3.8) are Galilean invariant in frames of reference moving with a fixed velocity. The Galilean invariance indicates that the equations and closure models remain unchanged from one inertial frame to another and are not sensitive to translations, which allows these models to be utilized in various physical settings. While our numerical solution will be based on (3.8a), this second form will facilitate the derivations presented in Section 3.3. We will assume that for all times $t \in [0, T]$ the filtered vorticity field $\widetilde{\omega}$ is in the Sobolev space $H_0^2(\Omega)$ of zero-mean functions with square-integrable second derivatives (Adams and Fournier 2003). We stress the distinction between the fields w, $\widetilde{w}$, $\widetilde{\omega}$ which represent, respectively, the solution of the DNS system (3.1), its filtered version and the solution of the LES system (3.8).

3.2.2 Optimization Formulation for Eddy Viscosity

We consider two formulations with the DNS field matched pointwise in space and in time, and in a certain statistical sense. First, the optimal eddy viscosity will be found as a minimizer of an error functional representing the mean-square error between observations of the filtered DNS, i.e., of the filtered solution $\widetilde{w}(t, \mathbf{x})$ of the Navier-Stokes system (3.1), and observations of the corresponding prediction $\widetilde{\omega}(t, \mathbf{x}; \varphi)$ of the LES model (3.8) with eddy viscosity ν. These observations are acquired at points $\mathbf{x}_i$, $i = 1, \ldots, M^2$, forming a uniform $M \times M$ grid in Ω with operators $H_i : H^2(\Omega) \longrightarrow \mathbb{R}$ defined as

$$\left(H_i\widetilde{\omega}\right)(t) := \int_{\Omega} \delta(\mathbf{x} - \mathbf{x}_i)\widetilde{\omega}(t, \mathbf{x})\, d\Omega = \widetilde{\omega}(t, \mathbf{x}_i), \quad i = 1, \ldots, M^2, \tag{3.9}$$

where $\delta(\cdot)$ is the Dirac delta distribution and observations $\left(H_i\widetilde{\omega}(\varphi)\right)(t)$ of the LES solution are defined analogously (an integral representation of the observation operators will be convenient for the derivation of the solution approach for the optimization problem presented in Section 3.3). The number of the observations points M^2 will be chosen

such that $M \gtrapprox k_c$, i.e., the observations will resolve all flow features with wavenumbers slightly higher than the cutoff wavenumber k_c in (3.3). The error functional then takes the form

$$\mathcal{J}_1(\varphi) := \frac{1}{2}\int_0^T \sum_{i=1}^{M^2} \left[\left(H_i\widetilde{w}\right)(t) - \left(H_i\widetilde{\omega}(\varphi)\right)(t)\right]^2 \, dt, \tag{3.10}$$

and is understood as depending on the function φ parametrizing the eddy viscosity $\nu = \nu(s)$ via ansatz (3.7).

In the second formulation, the optimal eddy viscosity will be found by minimizing the error between the time-averaged vorticity spectra in the filtered DNS and predicted by the LES. For simplicity and with a slight abuse of notation, we will treat the wavenumber k as a continuous variable, i.e., we will assume that $\boldsymbol{k} \in \mathbb{R}^2$ rather than $\boldsymbol{k} \in \mathbb{Z}^2$; in the actual implementation one needs to account for the discrete nature of the wavevector $\boldsymbol{k}$. The vorticity spectrum predicted by the LES is then defined as

$$E_{\widetilde{\omega}}(t,k) := \frac{1}{2}\int_{\mathscr{C}(k)} |\widehat{\widetilde{\omega}}(t,\boldsymbol{k})|^2 \, dS(\boldsymbol{k}), \qquad \forall t, k \geq 0, \tag{3.11}$$

where $\widehat{\widetilde{\omega}}(t,\boldsymbol{k})$ is the Fourier transform of $\widetilde{\omega}(t,\mathbf{x})$ and $\mathscr{C}(k) := \{\boldsymbol{k} \in \mathbb{R}^2 : |\boldsymbol{k}| = k\}$ a circle with radius k in the 2D plane. The vorticity spectrum $E_w(t,k)$ in the (filtered) DNS is defined analogously. Denoting $[f]_T := (1/T)\int_0^T f(t)\,dt$ the time average of a function $f : [0,T] \to \mathbb{R}$, the error functional is defined as

$$\mathcal{J}_2(\varphi) := \frac{1}{4}\int_{k=0}^{k_c} \left([E_{\widetilde{\omega}}(\cdot,k;\varphi)]_T - [E_w(\cdot,k)]_T\right)^2 \, dk, \tag{3.12}$$

with matching performed up to the cutoff wavenumber k_c.

The form of equation (3.8a) suggests that $\nu = \nu(s)$, and hence also $\varphi = \varphi(s/s_{\max})$, must be at least piecewise C^1 functions on $\mathcal{I}$ and $[0,1]$, respectively. However, as will become evident in Section 3.3, our solution approach imposes some additional regularity requirements, namely, $\nu = \nu(s)$ needs to be piecewise C^2 on $\mathcal{I}$ (to satisfy the adjoint

of the linearization of (3.8), defined below, in a classical sense) with the first and third derivatives vanishing at $s = 0, s_{\max}$. Moreover, we will require $\varphi \in C^3([0,1])$ in order to satisfy an elliptic boundary-value problem needed to define the Sobolev gradient, which is described in Section 3.3. Since gradient-based solution approaches to PDE-constrained optimization problems are preferably formulated in Hilbert spaces (Protas et al. 2004), we shall look for an optimal function φ parametrizing the optimal eddy viscosity as an element of the following linear space which is a subspace of the Sobolev space $H^2(\mathcal{I})$

$$\mathcal{S} := \left\{ \varphi \in C^3([0,1]) \ : \ \frac{d}{d\xi}\varphi(\xi) = \frac{d^3}{d\xi^3}\varphi(\xi) = 0 \text{ at } \xi = 0, 1 \right\}. \tag{3.13}$$

Then, we have the following two formulations of the optimization problem defining the optimal eddy viscosity.

Problem 2. *For the system (3.8), and objective functionals (3.10) or (3.12), find*

$$\breve{\varphi} := \arg\min_{\varphi \in \mathcal{S}} \mathcal{J}_j(\varphi), \qquad j = 1, 2, \tag{3.14}$$

where the optimal eddy viscosity $\breve{\nu}$ is deduced from $\breve{\varphi}$ via ansatz (3.7).

Our approach to solving this problem is outlined in the next section.

3.3 Adjoint-based Optimization

To fix attention, we focus here on solution of the optimization problem in the first formulation, i.e., for $j = 1$ in (3.14), with the error functional given in (3.10). Essentially the same approach can also be used to solve the second optimization problem with the error functional (3.12) and required modifications are discussed in Appendix C. We formulate our approach in the continuous ("optimize-then-discretize") setting (Gunzburger 2003) and adopt the strategy developed and validated by Matharu and Protas (2020). Here we only summarize its key steps and refer the reader to that study for further details. A

local solution of Problem 2 can be found using an iterative gradient-based minimization approach as $\breve{\varphi} = \lim_{n\to\infty} \varphi^{(n)}$, where

$$\begin{cases} \varphi^{(n+1)} & = \varphi^{(n)} - \tau^{(n)}\, \nabla_{\varphi} \mathcal{J}_1(\varphi^{(n)}), \qquad n = 0, 1, \ldots, \\ \varphi^{(0)} & = \varphi_0, \end{cases} \tag{3.15}$$

in which $\varphi^{(n)}$ is the approximation of the optimal function $\breve{\varphi}$ at the nth iteration (which can be used to construct the corresponding approximation $\nu^{(n)}$ of the optimal eddy viscosity), $\nabla_{\varphi}\mathcal{J}_1(\varphi)$ is the gradient of the error functional (3.10) with respect to φ, $\tau^{(n)}$ is the step length along the descent direction and φ_0 is an initial guess usually suggested by some form of the eddy viscosity.

A central element of algorithm (3.15) is the gradient $\nabla_{\varphi}\mathcal{J}_1(\varphi)$. In many problems of PDE-constrained optimization it can be conveniently expressed using solutions of suitably-defined adjoint equations (Gunzburger 2003). However, the present optimization Problem 2 has a nonstandard structure because the control variable $\varphi(s/s_{\max})$ is a function of the *dependent* variable $s = |\boldsymbol{\nabla}\widetilde{\omega}|^2$ in system (3.8). On the other hand, in its standard formulation adjoint analysis allows one to obtain expressions for gradients depending on the *independent* variables in the problem (here, t and $\mathbf{x}$). This difficulty was overcome by Bukshtynov et al. (2011) and Bukshtynov and Protas (2013) who generalized adjoint analysis of PDE systems to problems of the type given in Problem 2 by introducing a suitable change of variables. For convenience we will denote $\sigma := s/s_{\max}$.

The Gâteaux (directional) differential of the error functional (3.10) with respect to φ, defined by $\mathcal{J}_1{}'(\varphi;\varphi') := \lim_{\epsilon\to 0} \epsilon^{-1}\left[\mathcal{J}_1(\varphi+\epsilon\varphi') - \mathcal{J}_1(\varphi)\right]$, is defined as

$$\mathcal{J}_1{}'(\varphi;\varphi') = \int_0^T \int_{\Omega} \sum_{i=1}^{M^2} H_i^* \left[\left(H_i\widetilde{\omega}(\varphi)\right)(t) - \left(H_i\widetilde{w}\right)(t)\right] \widetilde{\omega}'(t,\mathbf{x};\varphi,\varphi')\, d\mathbf{x}\, dt, \tag{3.16}$$

where $\varphi' \in \mathcal{S}$ is an arbitrary perturbation of the control variable φ, $\widetilde{\omega}'(t, \mathbf{x}; \varphi, \varphi')$ satisfies the system

$$\mathcal{K}\begin{bmatrix} \widetilde{\omega}' \\ \\ \widetilde{\psi}' \end{bmatrix} := \begin{bmatrix} \partial_t \widetilde{\omega}' + \boldsymbol{\nabla}^{\perp}\widetilde{\psi}' \cdot \boldsymbol{\nabla}\widetilde{\omega} + \boldsymbol{\nabla}^{\perp}\widetilde{\psi} \cdot \boldsymbol{\nabla}\widetilde{\omega}' + \alpha\widetilde{\omega}' \\ -\boldsymbol{\nabla} \cdot \Big(2(\boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}')(\frac{d\nu}{ds}\,\varphi\boldsymbol{\nabla}\widetilde{\omega} + \frac{\nu_L+\nu_0}{s_{\max}}\,\frac{d\varphi}{d\sigma}\,\boldsymbol{\nabla}\widetilde{\omega}) + (\nu_N + \nu)\boldsymbol{\nabla}\widetilde{\omega}'\Big) \\ \boldsymbol{\Delta}\widetilde{\psi}' + \widetilde{\omega}' \end{bmatrix} \tag{3.17a}$$

$$= \begin{bmatrix} \boldsymbol{\nabla} \cdot ((\nu_L + \nu_0)\varphi'\boldsymbol{\nabla}\widetilde{\omega}) \\ \\ \\ 0 \end{bmatrix}, \tag{3.17b}$$

$$\widetilde{\omega}'(t = 0, \mathbf{x}) = 0, \tag{3.17c}$$

obtained as linearization of the LES system (3.8) and $H_i^* : \mathbb{R} \longrightarrow H^{-2}(\Omega)$, $i = 1, \ldots, M^2$, are the adjoints of the observation operators H_i, cf. (3.9), given by

$$\forall \xi \in \mathbb{R}, \qquad (H_i^*\xi) := \delta(\mathbf{x} - \mathbf{x}_i)\xi, \quad i = 1, \ldots, M^2. \tag{3.18}$$

In order to extract the gradient $\nabla_\varphi \mathcal{J}_1$ from the Gâteaux differential (3.16), we note that this derivative is a bounded linear functional when viewed as a function of φ' and invoke the Riesz representation theorem (Berger 1977) to obtain

$$\mathcal{J}_1{}'(\varphi; \varphi') = \left\langle \nabla_\varphi^{H^2} \mathcal{J}_1, \varphi' \right\rangle_{H^2([0,1])} = \left\langle \nabla_\varphi^{L^2} \mathcal{J}_1, \varphi' \right\rangle_{L^2([0,1])}, \tag{3.19}$$

where the inner product in the space $H^2([0, 1])$ is defined as

$$\left\langle p_1, p_2 \right\rangle_{H^2([0,1])} = \int_0^1 p_1\, p_2 + \ell_1^2\, \frac{dp_1}{d\sigma}\,\frac{dp_2}{d\sigma} + \ell_2^4\, \frac{d^2p_1}{d\sigma^2}\,\frac{d^2p_2}{d\sigma^2}\, d\sigma, \tag{3.20}$$

in which ℓ_1 and ℓ_2 are length-scale parameters. While for all values of $\ell_1, \ell_2 \in (0, \infty)$ the inner products (3.20) are equivalent (in the sense of norm equivalence), these two parameters play a very important role in regularization of solutions to the optimization Problem 2. In (3.15) we require the gradient in the space $H^2([0,1])$, i.e., $\nabla_\varphi \mathcal{J}_1 = \nabla_\varphi^{H^2} \mathcal{J}_1$, but it is convenient to first derive the gradient with respect to the L^2 topology.

Introducing *adjoint fields* $\widetilde{\omega}^*$ and $\widetilde{\psi}^*$, we can define the following duality-pairing relation

$$\begin{aligned}\left(\mathcal{K} \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \begin{bmatrix} \widetilde{\omega}^* \\ \psi^* \end{bmatrix} \right) &:= \int_0^T \int_\Omega \mathcal{K} \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix} \cdot \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} d\mathbf{x}\, dt \\ \int_0^T \int_\Omega \begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix} \cdot \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \widetilde{\psi}^* \end{bmatrix} d\mathbf{x}\, dt &= \left(\begin{bmatrix} \widetilde{\omega}' \\ \widetilde{\psi}' \end{bmatrix}, \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \psi^* \end{bmatrix} \right),\end{aligned} \tag{3.21}$$

where integration by parts was performed with respect to both space and time (noting the periodic boundary conditions and the initial condition (3.17c)) and the *adjoint system* has the form

$$\mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \\ \widetilde{\psi}^* \end{bmatrix} := \begin{bmatrix} -\partial_t \widetilde{\omega}^* - \boldsymbol{\nabla}^\perp \widetilde{\psi} \cdot \boldsymbol{\nabla} \widetilde{\omega}^* + \alpha \widetilde{\omega}^* + \widetilde{\psi}^* \\ -\boldsymbol{\nabla} \cdot \left(2 \left(\boldsymbol{\nabla} \widetilde{\omega} \cdot \boldsymbol{\nabla} \widetilde{\omega}^* \right) \left(\frac{d\nu}{ds}\, \varphi \boldsymbol{\nabla} \widetilde{\omega} + \frac{\nu_L + \nu_0}{s_{\max}} \frac{d\varphi}{d\sigma} \boldsymbol{\nabla} \widetilde{\omega} \right) + (\nu_N + \nu) \boldsymbol{\nabla} \widetilde{\omega}^* \right) \\ \boldsymbol{\Delta} \widetilde{\psi}^* - \boldsymbol{\nabla}^\perp \cdot \left(\widetilde{\omega}^* \boldsymbol{\nabla} \widetilde{\omega} \right) \end{bmatrix} = \begin{bmatrix} W \\ \\ 0 \end{bmatrix}, \tag{3.22a}$$

$$\widetilde{\omega}^*(t = T, \mathbf{x}) = 0, \tag{3.22b}$$

with the source term $W(t,\mathbf{x}) := \sum_{i=1}^{M^2} H_i^* \left[(H_i\widetilde{\omega}(\varphi))(t) - (H_i\widetilde{w})(t) \right]$. Combining (3.17), (3.21) and (3.22), we then arrive at

$$\begin{aligned} \left(\begin{bmatrix} \widetilde{\omega}' \\ \\ \widetilde{\psi}' \end{bmatrix}, \mathcal{K}^* \begin{bmatrix} \widetilde{\omega}^* \\ \\ \widetilde{\psi}^* \end{bmatrix} \right) &= \overbrace{\int_0^T \int_\Omega W(t,\mathbf{x})\, \widetilde{\omega}'\, d\mathbf{x}dt}^{\mathcal{J}_1{}'(\varphi;\varphi')} \\ &= -\int_0^T \int_\Omega (\nu_L + \nu_0)\, (\boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}^*)\, \varphi'\, d\mathbf{x}\, dt, \end{aligned} \tag{3.23}$$

from which we obtain an expression for the Gâteaux differential

$$\mathcal{J}_1{}'(\varphi;\varphi') = -\int_0^T \int_\Omega (\nu_L + \nu_0)\, (\boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}^*)\, \varphi'\, d\mathbf{x}\, dt,$$

with the perturbation φ' now appearing explicitly as a factor. However, this expression is still not consistent with the Riesz form (3.19), which requires integration with respect to s over $[0,1]$. In order to perform the required change of variables, we make the substitution $\varphi'(\boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}) = \int_0^1 \delta\left(\frac{\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\widetilde{\omega}}{s_{\max}} - \sigma \right) \varphi'(\sigma)\, d\sigma$. Fubini's theorem then allows us to swap the order of integration such that the Gâteaux differential (3.16) is finally recast in the Riesz form (3.19) as an integral with respect to σ

$$\mathcal{J}_1{}'(\varphi;\varphi') = \int_0^1 \left[-\int_0^T \int_\Omega \delta\left(\frac{\boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}}{s_{\max}} - \sigma \right) (\nu_L + \nu_0)\, \boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}^*\, d\mathbf{x}\, dt \right] \varphi'(\sigma)\, d\sigma. \tag{3.24}$$

The gradient defined with respect to the L^2 topology is then deduced from this expression as

$$\nabla_\varphi^{L^2} \mathcal{J}_1(\sigma) = -\int_0^T \int_\Omega \delta\left(\frac{\boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}}{s_{\max}} - \sigma \right) (\nu_L + \nu_0)\, \boldsymbol{\nabla}\widetilde{\omega} \cdot \boldsymbol{\nabla}\widetilde{\omega}^*\, d\mathbf{x}\, dt. \tag{3.25}$$

The L^2 gradient given in (3.25) may in principle be discontinuous as a function of s

and hence will not ensure the regularity required of the optimal eddy viscosity, cf. Section 3.2.2. To circumvent this problem, we define a Sobolev gradient using the Riesz relations (3.19) to identify the H^2 inner product (3.20) with expression (3.24) for the Gâteaux differential. Integrating by parts with respect to σ and noting that the perturbation $\varphi' \in \mathcal{S}$ is arbitrary, we obtain the Sobolev gradient $\nabla_\varphi^{H^2}\mathcal{J}$ as a solution of the elliptic boundary-value problem

$$\left[\mathrm{Id} - \ell_1^2\,\frac{d^2}{d\sigma^2} + \ell_2^4\frac{d^4}{d\sigma^4}\right]\nabla_\varphi^{H^2}\mathcal{J}_1(\sigma) = \nabla_\varphi^{L^2}\mathcal{J}_1(\sigma), \qquad \sigma \in [0,1], \tag{3.26a}$$

$$\frac{d^{(1)}\left(\nabla_\varphi^{H^2}\mathcal{J}_1\right)}{d\sigma^{(1)}}\Big|_{\sigma=0,1} = \frac{d^{(3)}\left(\nabla_\varphi^{H^2}\mathcal{J}_1\right)}{d\sigma^{(3)}}\Big|_{\sigma=0,1} = 0. \tag{3.26b}$$

The choice of the boundary conditions in (3.26b) ensures the vanishing of all the boundary terms resulting from the integration by parts. There is in fact some freedom in how to cancel these terms and the choice in (3.26b) is arguably the least restrictive. As argued in Section 3.2.1, we allow the eddy viscosity $\nu(s)$ to take nonzero values at $s = 0$ so the corresponding Sobolev gradient should not vanish at $\sigma = 0$ such that it can modify the value of $\varphi(0)$, which turns out to be important in practice, cf. Section 3.5. Thus, the choice of boundary conditions at $\sigma = 0$ provided in (3.26b) is necessary. On the other hand, the choice of the boundary conditions at $\sigma = 1$ has been found to have little effect on the gradient and on the obtained results provided $s_{\max}$ is sufficiently large. Therefore, the form of these boundary conditions given in (3.26b) is justified by simplicity. The boundary conditions (3.26b) are the reason for the presence of additional constraints in the definition of space $\mathcal{S}$ in (3.13).

Determination of the Sobolev gradients $\nabla_\varphi^{H^2}\mathcal{J}_1$ based on the L^2 gradients $\nabla_\varphi^{L^2}\mathcal{J}_1$ by solving system (3.26) can be viewed as low-pass filtering of the latter gradient using a non-sharp filter (as discussed by Protas et al. (2004), this can be seen representing the operator $\left[\mathrm{Id} - \ell_1^2\,(d^2/d\sigma^2) + \ell_2^4(d^4/d\sigma^4)\right]^{-1}$ in the Fourier space). The parameters ℓ_1 and ℓ_2 serve as cutoff length scales representing the wavelengths of the finest features

retained in the gradients $\nabla_{\varphi}^{H^2}\mathcal{J}_1$ such that increasing ℓ_1 and ℓ_2 has the effect of making the Sobolev gradient "smoother" and vice versa. Thus, ℓ_1 and ℓ_2 are "knobs" which can be tuned to control the regularity of the optimal eddy viscosities obtained as solutions of the Problem 2.

Since by construction $\nabla_{\varphi}^{H^2}\mathcal{J}_1 \in \mathcal{S}$, choosing the initial guess in (3.15) such that $\varphi_0 \in \mathcal{S}$ will ensure that $\varphi^{(0)}, \varphi^{(1)}, \ldots, \breve{\varphi} \in \mathcal{S}$. At each step in (3.15) an optimal step size $\tau^{(n)}$ can be found by solving the following line-minimization problem (Nocedal and Wright 2006)

$$\tau^{(n)} = \arg\min_{\tau>0} \mathcal{J}_1(\varphi^{(n)} - \tau\, \nabla_{\varphi}\mathcal{J}_1(\varphi^{(n)})). \tag{3.27}$$

Numerical implementation of the approach outlined above is discussed in the next section.

3.4 Computational Approach

The evaluation of the Sobolev gradient $\nabla_{\varphi}^{H^2}\mathcal{J}_1$ requires the numerical solutions of the LES system (3.8) and the adjoint system (3.22) followed by the solution of problem (3.26). For the first two systems we use a standard Fourier pseudo-spectral method in combination with a CN/RKW3 time-stepping technique introduced by Le and Moin (1991) which give spectrally accurate results in space and a globally second-order accuracy in time; details are provided in Appendix A. The spatial domain is discretized using $N_x = 256$ equispaced grid points in each direction. Since the eddy viscosity $\nu = \nu(s)$ and the function $\varphi(s/s_{\max})$ are state-dependent, we also need to discretize the state domain $\mathcal{I}$, cf. (3.7), which is done using N_s Chebyshev points (values of N_s are provided in Table 3.1). We use Chebyshev differentiation matrices to perform differentiation with respect to s and the eddy viscosity $\nu(s)$ and its derivatives are interpolated from state space $\mathcal{I}$ to the spatial domain Ω using the barycentric formulas (Trefethen 2013). The

boundary-value problem (3.26) is solved using a method based on ultraspherical polynomials available in the `chebop` feature of Chebfun (Driscoll et al. 2014). Solution of the 2D Navier-Stokes system (3.1) is dealiased using Gaussian filtering based on the 3/2 rule (Hou 2009), however, this is unnecessary for the LES system (3.8) due to the aggressive filtering applied. To ensure that aliasing errors resulting from the presence of an a priori unknown nonlinearity in the state-dependent viscosity are eliminated, the adjoint system (3.22) is solved using twice as many grid points $2N_x$ in each direction.

Evaluation of the L^2 gradient (3.25) requires non-standard integration over level sets as described by Bukshtynov and Protas (2013). Due the complexity of the determining this gradient, its evaluation is thoroughly validated in Appendix A2. While for simplicity a simple gradient approach was presented in (3.15), in practice we use the Polak-Ribière variant of the conjugate-gradient method to accelerate convergence. For the line minimization problem (3.27), the standard Brent's algorithm is used (Press et al. 2007). The consistency and accuracy of the formulation and of the entire computational approach was validated using a standard suite of tests as was done by Matharu and Protas (2020).

3.5 Results

The results obtained by solving optimization problem (3.14) with error functionals (3.10) and (3.12) are presented in Sections 3.5.1 and 3.5.2 below. Our computations are based on a flow problem defined by the following parameters $\nu_N = 1\times10^{-2}$, $\alpha = 1\times10^{-3}$, $F = 5$, and $k_a = k_b = 4$, which results in a turbulent flow with a fully developed spectrum. In the first optimization problem we fix $M = 32$ in (3.10), which is slightly larger than the largest cutoff wavenumber k_c we consider (cf. Table 3.1) and therefore ensures that the optimal eddy viscosity is determined based on all available flow information, and $T = 20 \approx 30t_e$, where $t_e := \left[\int_0^T \mathcal{E}(t)\,dt/(8\pi^2 T)\right]^{-1/2}$ is the eddy turnover time (Bracco and McWilliams 2010). We emphasize that the key insights provided by our

computations do not depend on the particular choice of T, as long as it remains of comparable magnitude to the value given above.

3.5.1 Matching the DNS Pointwise in Space and Time — Results for the Optimization Problem with Error Functional (3.10)

Our first set of results addresses the effect of the cutoff wavenumber k_c. They are obtained by solving problem (3.14) with $j = 1$ for decreasing values of $k_c = 30, 25, 20$ while retaining fixed values of the regularization parameters ℓ_1, ℓ_2 and a fixed resolution N_s in the state space $\mathcal{I}$, cf. cases A, B and C in Table 3.1. In each case the optimization problem is solved using the initial guess $\varphi_0(s/s_{\max}) \equiv 0$ corresponding to no closure model at all. The dependence of the error functional $\mathcal{J}_1(\nu^{(n)})$ on iterations n in the three cases is shown in Figure 3.1a, where we see that the mean-square errors between the DNS and the optimal LES increase as the cutoff wavenumber k_c is decreased and the largest relative reduction of the error is achieved in case C with the smallest k_c. While minimization in problem (3.14) is performed with respect to the nondimensional function φ, cf. (3.7), we focus here on the corresponding optimal eddy viscosities $\breve{\nu} = \breve{\nu}(s)$ shown in Figure 3.1b. Since small values of s are attained more frequently in the flow, cf. the probability density function (PDF) of $\sqrt{s}$ embedded in the figure, the horizontal axis is scaled as $\sqrt{s}$ which magnifies the region of small values of s. We see that for the largest cutoff wavenumber $k_c = 30$ the optimal eddy viscosity is close to zero over the entire range of s. However, for decreasing k_c the optimal eddy viscosity exhibits oscillations of increasing magnitude. We note that values of $s \gtrapprox 50$ occur very rarely in the flow and hence the gradient (3.25) provides little sensitivity information for s in this range. Thus, the behavior of $\breve{\nu}(s)$ for $s \gtrapprox 50$ is an artifact of the regularization procedure defined in (3.26) and is not physically relevant.

In order to provide additional insights about the properties of the optimal eddy viscosity, our second set of results is obtained as solutions of problem (3.14) with $j = 1$

Case	k_c	N_s	ℓ_1	ℓ_2	φ_0	$\mathcal{J}_1(\varphi_0)$	$\mathcal{J}_1(\varphi^{(\infty)})$	r
A	30	64	10^4	10^3	No Closure	4.398×10^{-7}	1.492×10^{-7}	8.999×10^{-8}
B	25	64	10^4	10^3	No Closure	1.951×10^{-5}	2.450×10^{-6}	1.572×10^{-6}
C	20	64	10^4	10^3	No Closure	3.635×10^{-4}	6.217×10^{-5}	4.468×10^{-5}
D	20	128	10^3	10^2	Case C	6.217×10^{-5}	2.001×10^{-5}	1.239×10^{-5}
E	20	256	10^1	10^0	Case D	2.333×10^{-5}	1.450×10^{-5}	8.723×10^{-6}

TABLE 3.1: Summary information about the different cases considered when solving optimization problem (3.14) with $j = 1$.

using a fixed $k_c = 20$ and progressively reduced regularization achieved by decreasing the parameters ℓ_1, ℓ_2 while simultaneously refining the resolution N_s in the state space $\mathcal{I}$, cf. cases C, D and E in Table 3.1. Optimization problems with weaker regularization are solved using the optimal eddy viscosity obtained with stronger regularization as the initial guess. From the normalized error functionals shown as functions of iterations in Figure 3.2a, we see that as regularization is reduced, the mean-square errors between the optimal LES and the DNS become smaller and approach a certain nonzero limit, cf. Table 3.1. As is evident from Figure 3.2b, this is achieved with the corresponding optimal eddy viscosities developing oscillations with an ever increasing frequency. More precisely, each time the regularization parameters ℓ_1, ℓ_2 are reduced and the resolution N_s is refined, a new oscillation with a higher frequency appears in the optimal eddy viscosity $\breve{\nu}(s)$ (in fact, in each case, this is the highest-frequency oscillation which can be represented on a grid with N_s points).

In order to assess how well the solutions of the LES system (3.8) with the optimal eddy viscosities $\breve{\nu}$ shown in Figures 3.1b and 3.2b approximate the solution of the Navier-Stokes system (3.1), in Figures 3.3a and 3.3b we show the time evolution of the quantity

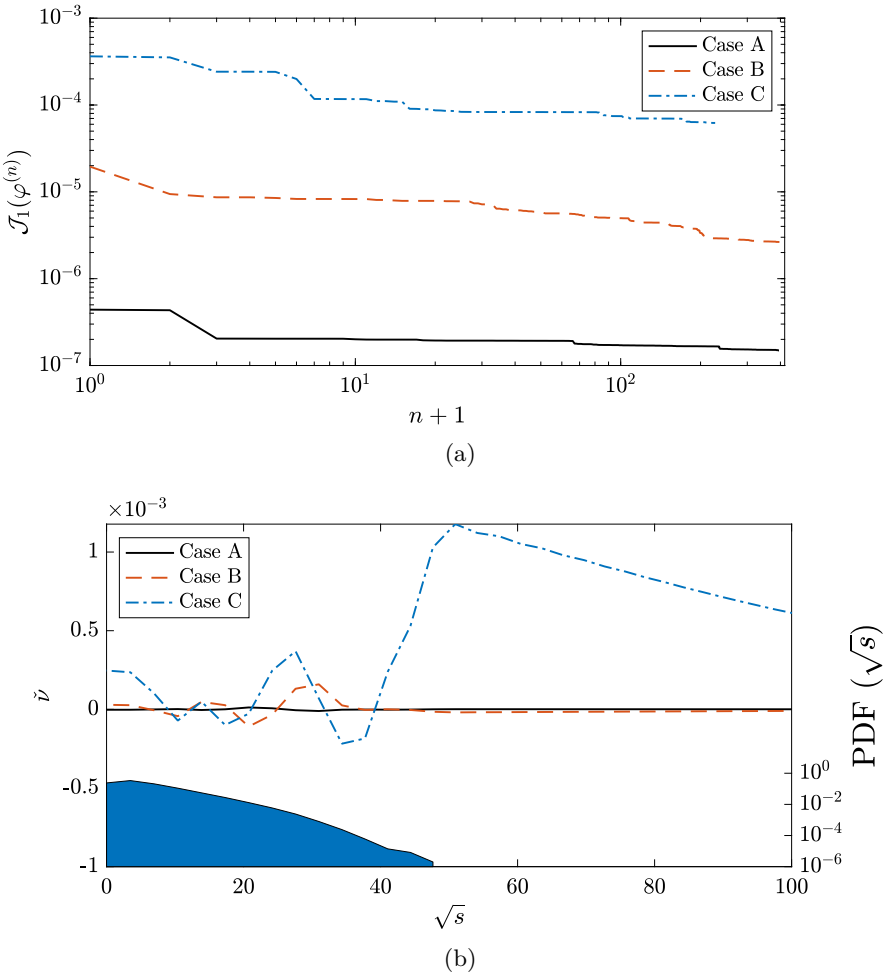
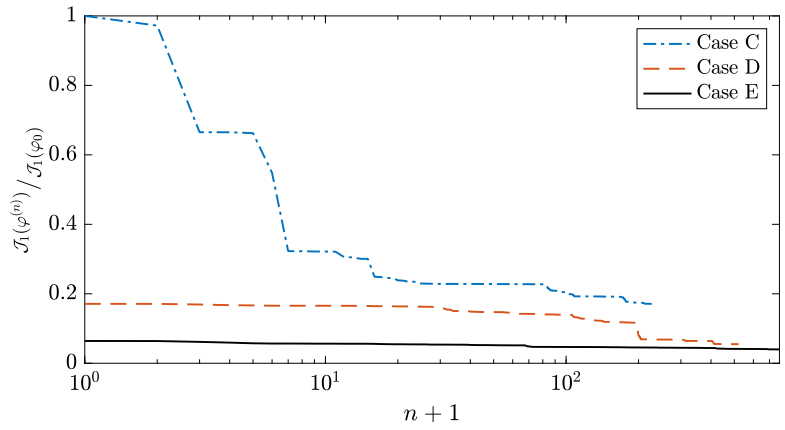
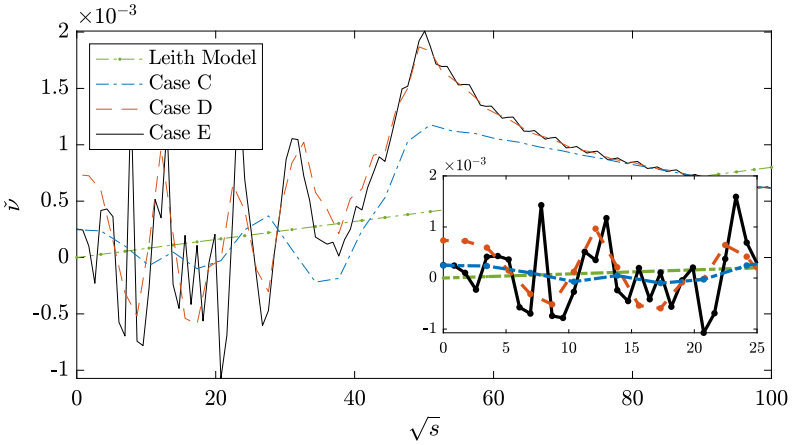


FIGURE 3.1: (a) Dependence of the functional $\mathcal{J}_1(\varphi^{(n)})$ on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\breve{\nu}$ on $\sqrt{s}$ for cases A, B and C, cf. Table 3.1. Panel (b) also shows the PDF of $\sqrt{s}$ in case C.

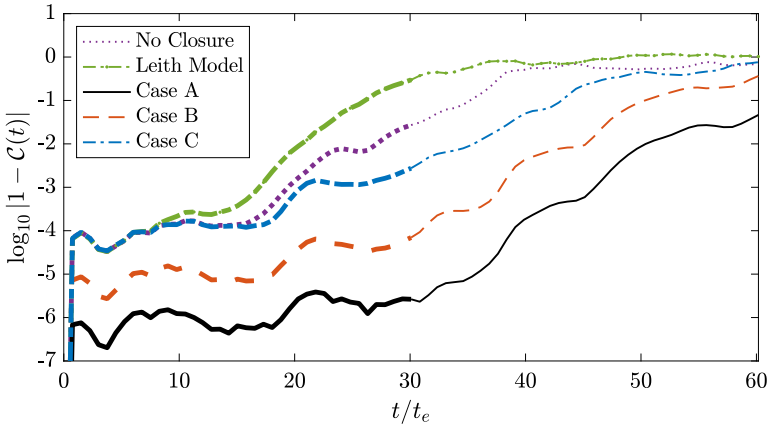


(a)

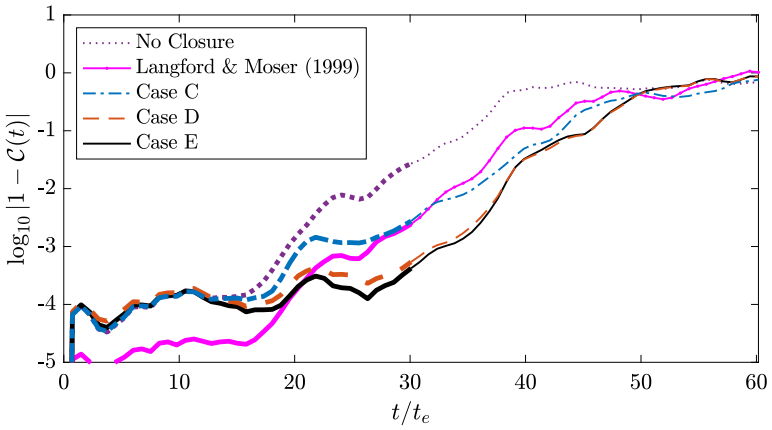


(b)

FIGURE 3.2: (a) Dependence of the normalized functional $\mathcal{J}_1(\varphi^{(n)})/\mathcal{J}_1(\varphi_0)$, with $\mathcal{J}_1(\varphi_0)$ from case C, on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\breve{\nu}$ on $\sqrt{s}$ for cases C, D and E, cf. Table 3.1. The inset in panel (b) shows magnification of the region $\sqrt{s} \in [0, 25]$. Panel (b) also shows the Leith model with $k_c = 20$ and the eddy viscosity $\nu_L(s)$, cf. (3.6).



(a)



(b)

FIGURE 3.3: Adjusted normalized correlations (3.28) for the LES with (a) no closure and the optimal eddy viscosity in cases A, B and C, and (b) no closure and the optimal eddy viscosity in cases C, D and E. The correlation is also shown for the Leith model with $k_c = 20$ and the eddy viscosity $\nu_L(s)$, cf. (3.6), in (a) and for an optimal closure model based on the stochastic estimator (Langford and Moser 1999) in (b). Thick and thin lines correspond to, respectively, time in the "training window" ($t \in [0, T]$) and beyond this window ($t \in (T, 2T]$).

$\log_{10}|1-\mathcal{C}(t)|$ where

$$\mathcal{C}(t) := \frac{1}{||\widetilde{w}(t)||_{L^2(\Omega)}\,||\widetilde{\omega}(t)||_{L^2(\Omega)}} \int_\Omega \widetilde{w}(t,\mathbf{x})\,\widetilde{\omega}(t,\mathbf{x})\; d\Omega \tag{3.28}$$

is the normalized correlation between the two flows. For a more comprehensive assessment, these results are shown for $t \in [0, 2T]$, i.e., for times up to twice longer than the "training window" $[0, T]$ used in the optimization problem (3.14). In Figure 3.3b we also present the results obtained for $k_c = 20$ with an optimal closure model based on the linear stochastic estimator introduced by Langford and Moser (1999). Since at early times correlation $\mathcal{C}(t)$ reveals exponential decay corresponding to the exponential divergence of the LES flow from the DNS, this effect can be quantified by approximating the correlation as $\mathcal{C}(t) \approx \bar{\mathcal{C}}(t) := \mathcal{C}_0 e^{-rt}$, where $\mathcal{C}_0 = 1$ follows the fact that $\widetilde{\omega}_0 \equiv \widetilde{w}_0$, whereas the decay rate r is obtained from a least-squares fit over the time window $[0, T]$. The decay rates r obtained in this way are collected in Table 3.1.

Finally, in order to provide insights about how the closure model with the optimal eddy viscosity acts in the physical space, in Figures 3.4a, 3.4b and 3.4d we show the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, the corresponding state variable $s(T, \mathbf{x})$ cf. (3.7), and the spatial distribution $\breve{\nu}(s(T, \mathbf{x}))$ of the optimal eddy viscosity obtained in case E; for comparison, the spatial distribution of the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model, cf. (3.6) with $\delta = 0.02$, is shown in Figure 3.4c (the fields are shown in the entire domain, i.e., for $\mathbf{x} \in \Omega$, at the end of the training window). We see that while the vorticity and state-variable fields vary smoothly, this is also the case for the spatial distribution of the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model. On the other hand, the spatial distribution of the optimal eddy viscosity $\breve{\nu}(s(T, \mathbf{x}))$ exhibits rapid variations, which is consistent with the results presented in Figure 3.2b. In particular, positive and negative values of $\breve{\nu}(s(T, \mathbf{x}))$, corresponding to localized dissipation and injection of enstrophy, tend to form concentric bands in some low-vorticity regions of the flow domain. The time evolution

of the vorticity field in the DNS, LES with no closure model and LES with the optimal eddy viscosity (case E) are available together with an animated version of Figure 3.3b in Video 4. An animation representing the time evolution of the fields shown in Figure 3.4 for $t \in [0, 2T]$ is shown in Video 5.

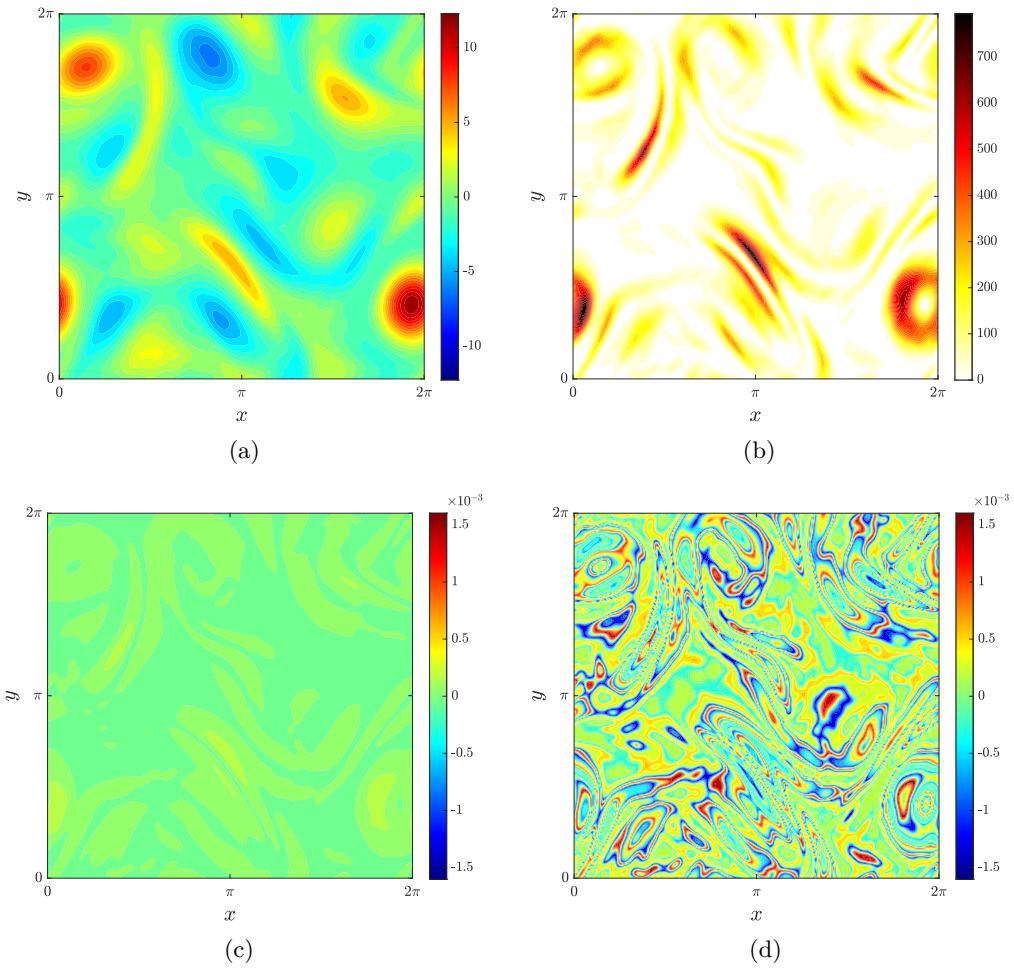


FIGURE 3.4: For case E we show: (a) the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, $\mathbf{x} \in \Omega$, (b) the corresponding state variable $s(T, \mathbf{x})$, cf. (3.7), and the spatial distribution of (c) the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model, cf. (3.6) with $\delta = 0.02$, and (d) the optimal eddy viscosity $\breve{\nu}(s(T, \mathbf{x}))$, cf. Figure 3.2b, all shown at the end of the training window for $t = T$. For better comparison the same color scale is used in panels (c) and (d).

3.5.2 Matching the DNS in an Average Sense — Results for the Optimization Problem with Error Functional (3.12)

Now we review the results obtained by solving optimization problem (3.14) for $j = 2$ with a fixed cutoff wavenumber $k_c = 20$ and with two sets of parameters determining regularization (ℓ_1 and ℓ_2) and the resolution in the state space $\mathcal{I}$ (N_s), cf. cases F and G in Table 3.2. We remark that the regularization performed in the present problem is less aggressive than in the problem discussed in Section 3.5.1. As shown in Figure 3.5a, the normalized error functional converges to a local minimum in only a few iterations and, as the regularization is reduced, a larger reduction of the error functional is obtained. However, as is evident from Figure 3.5b, this is achieved with optimal eddy viscosities much better behaved than the optimal eddy viscosities found by solving the optimization problem discussed in Section 3.5.1, even though a weaker regularization is now applied, cf. Table 3.2 (the obtained optimal eddy viscosity exhibits more small-scale variability in case G than in case F, but the difference is not significant).

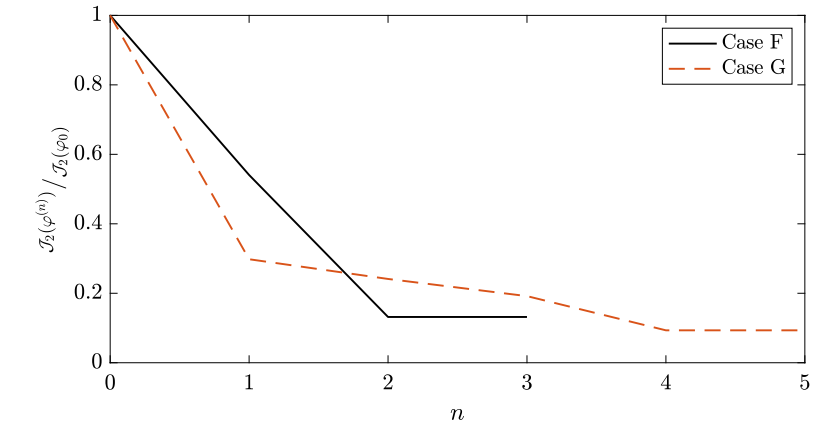
The difference between the time-averaged vorticity spectra (3.11) is the LES with no closure, LES with the optimal closure $\breve{\nu}$ (cases F and G) and in the filtered DNS is shown in Figure 3.6 as a function of the wavenumber k (this quantity is related to the integrand expression in the error functional (3.12)). We see that when the optimal eddy viscosity $\breve{\nu}$ is used in the LES, this error is reduced, especially at low wavenumbers k. On the other hand, the evolution of the quantity $\log_{10}|1 - \mathcal{C}(t)|$, cf. (3.28), shown for the same cases in Figure 3.7 demonstrates that, in contrast to Figure 3.3, in the present problem the LES flows equipped with the optimal eddy viscosity do not achieve a better pointwise-in-space accuracy with respect to the DNS than the LES flow with no closure model.

Finally, we show the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, the corresponding state variable $s(T, \mathbf{x})$, cf. (3.7), the spatial distribution $\breve{\nu}(s(T, \mathbf{x}))$ of the optimal eddy viscosity obtained in

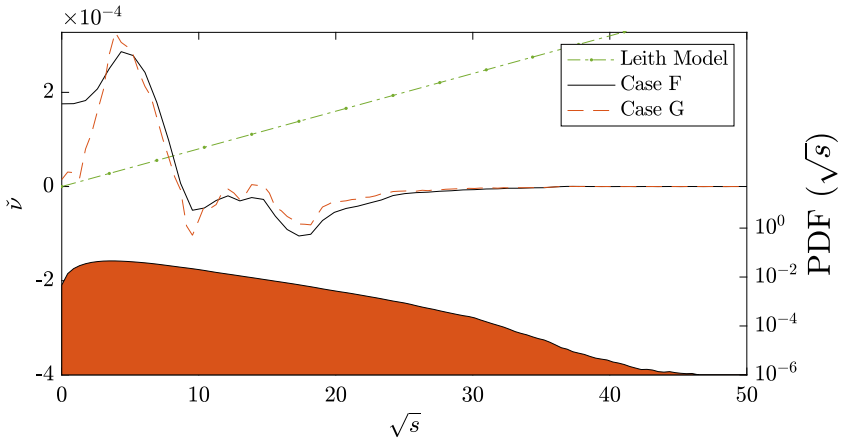
Case	k_c	N_s	ℓ_1	ℓ_2	φ_0	$\mathcal{J}_2(\varphi_0)$	$\mathcal{J}_2(\varphi^{(\infty)})$	r
F	20	256	10^1	10^0	No Closure	6.736×10^{-2}	8.876×10^{-3}	2.882×10^{-4}
G	20	512	10^{-1}	10^{-2}	No Closure	6.736×10^{-2}	6.286×10^{-3}	1.685×10^{-4}

TABLE 3.2: Summary information about the different cases considered when solving optimization problem (3.14) with $j = 2$.

case G, and for comparison, the spatial distribution of the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model, cf. (3.6), in Figures 3.8a, 3.8b, 3.8d, and 3.8c, respectively. We remark that the spatial distribution of the optimal eddy viscosity in Figure 3.8d is now significantly smoother than the distribution of the optimal eddy viscosity obtained in the first formulation by solving optimization problem (3.14) with $j = 1$, cf. Figure 3.4d. An animated version of Figure 3.8 illustrating the evolution of the fields for $t \in [0, 2T]$ is shown in Video 6.



(a)



(b)

FIGURE 3.5: (a) Dependence of the normalized functional $\mathcal{J}_2(\varphi^{(n)})/\mathcal{J}_2(\varphi_0)$ on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\breve{\nu}$ on $\sqrt{s}$ for cases F and G, cf. Table 3.2. Panel (b) also shows the PDF of $\sqrt{s}$ in case G.

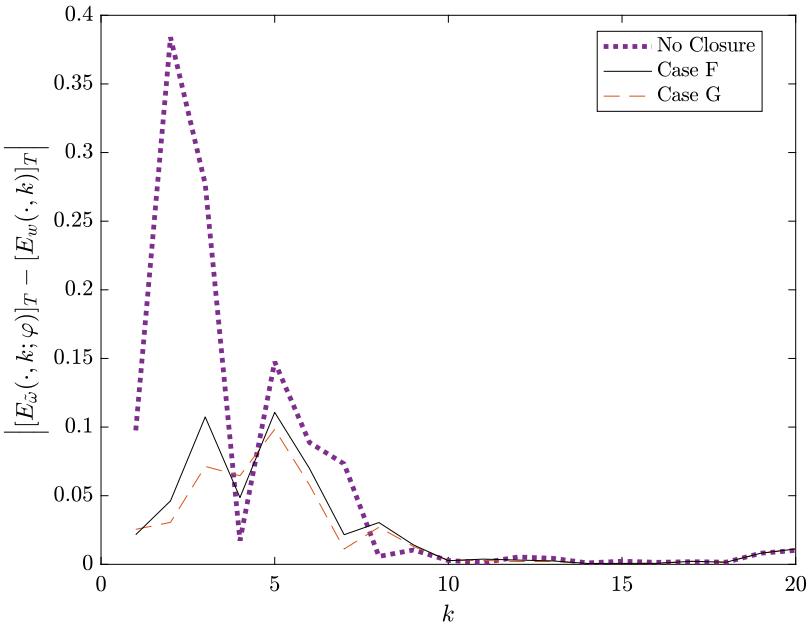


FIGURE 3.6: The difference between time-averaged vorticity spectra (3.11) in the filtered DNS and in the LES with no closure and with the optimal eddy viscosity $\breve{\nu}$ obtained in cases F and G, cf. Table 3.2, as function of the wavenumber k.

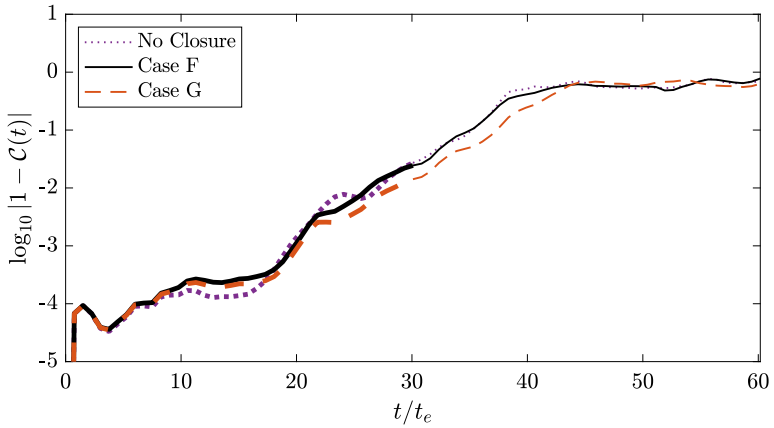


FIGURE 3.7: Adjusted normalized correlations (3.28) for the LES with no closure and the optimal eddy viscosity in cases F and G. Thick and thin lines correspond to, respectively, time in the "training window" ($t \in [0,T]$) and beyond this window ($t \in (T,2T]$).

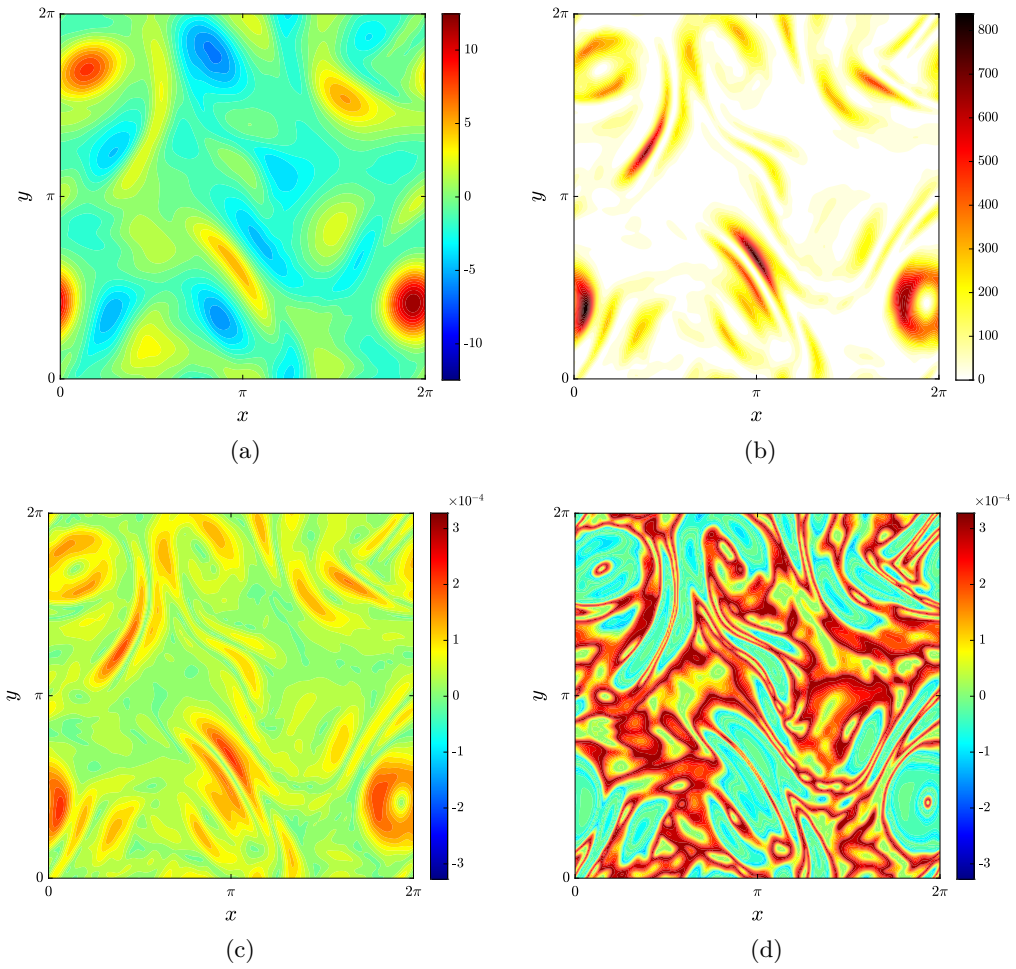


FIGURE 3.8: For case G we show: (a) the vorticity field $\widetilde{\omega}(T, \mathbf{x})$, $\mathbf{x} \in \Omega$, (b) the corresponding state variable $s(T, \mathbf{x})$, cf. (3.7), and the spatial distribution of (c) the eddy viscosity $\nu_L(s(T, \mathbf{x}))$ in the Leith model, cf. (3.6) with $\delta = 0.02$, and (d) the optimal eddy viscosity $\breve{\nu}(s(T, \mathbf{x}))$, cf. Figure 3.5b, all shown at the end of the training window for $t = T$. For better comparison the same color scale is used in panels (c) and (d).

3.6 Discussion and Conclusions

In this study we have considered the question of fundamental limitations on the performance of eddy-viscosity closure models for turbulent flows. We focused on the Leith model for 2D LES for which we sought optimal eddy viscosities that subject to minimum assumptions would result in the least mean-square error between the corresponding LES and the filtered DNS. Such eddy viscosities were found as minimizers of a PDE-constrained optimization problem with a nonstandard structure which was solved using a suitably adapted adjoint-based gradient approach (Matharu and Protas 2020). A key element of this approach was a regularization strategy involving the length-scale parameters ℓ_1 and ℓ_2 in the Sobolev gradients, cf. (3.26). The approach proposed is admittedly rather technically involved which may limit its practical applicability to construct new forms of the eddy viscosity, but its value is in making it possible to systematically characterize the best possible performance of different types of closure models.

Our main finding in Section 3.5.1 is that with a fixed cutoff wavenumber k_c the LES with an optimal eddy viscosity $\breve{\nu}$ matches the DNS increasingly well as the regularization in the solution of the optimization problem is reduced, cf. Figure 3.2a. This is quantified by a reduction of the rate of exponential decay of the correlation between the corresponding LES and the DNS, cf. Figure 3.3b and Table 3.1. This optimal performance of the closure model is achieved with eddy viscosities $\breve{\nu}(s)$ rapidly oscillating with a frequency increasing as the regularization parameters are reduced. From this we conclude that in the limit of vanishing regularization parameters and an infinite numerical resolution the optimal eddy viscosity would be undefined as it would exhibit oscillations with an unbounded frequency. Thus, from the mathematical point of view, the problem of finding an optimal eddy viscosity in the absence of regularization is ill-posed. In practical terms, this means that the "best" eddy viscosity for the Leith model does not exist.

The optimal performance of the LES is realized by a rapid variation of the eddy viscosity $\breve{\nu}(s)$ which oscillates between positive and negative values as s changes, cf. Figure 3.2b, resulting in the dissipation and injection of the enstrophy occurring in the physical domain in narrow alternating bands, cf. Figure 3.4d. We note that a somewhat similar behavior was also observed in (Maulik et al. 2020) where the authors used machine learning methods to determine pointwise estimates of eddy viscosity which exhibited oscillations between positive and negative values. This behavior can be understood in physical terms based on relations (3.4)–(3.5) which can be interpreted as defining the eddy viscosity in terms of the space- and time-dependent DNS field, but the problem is severely overdetermined. Thus, some form of relaxation is needed to determine ν and the proposed optimization approach with its inherent regularization strategy is one possibility.

In addition, the optimal eddy viscosities found here have the property that $\breve{\nu}(0) > 0$, in contrast to what is typically assumed in the Leith model where $\nu(0) = 0$ (Maulik and San 2017). In contrast to the behavior observed in Figure 3.4d, standard eddy viscosity closure models are usually assumed to be strictly dissipative (Rodi et al. 2013), which is reflected in the fact that the eddy viscosity is non-negative as in Figure 3.4c. We add that we have also considered finding optimal eddy viscosities by matching against the unfiltered DNS field, i.e., using $w(t, \mathbf{x})$ in the error functional (3.10) instead of $\widetilde{w}(t, \mathbf{x})$, however, this approach produced results very similar to the ones reported above. As is evident from Figure 3.3b, the performance of the LES with optimal eddy viscosities compares favourably to the LES with an optimal closure model proposed by Langford and Moser (1999) based on a stochastic estimator, which has a less restrictive structure than the Leith model.

The optimal eddy viscosities constructed in Section 3.5.1 to maximize the pointwise match against the filtered DNS are unlikely to be useful in practice due to their highly

irregular behaviour which is difficult to resolve using finite numerical precision. On the other hand, the second formulation studied in Section 3.5.2 where optimal eddy viscosities were determined by matching predictions of the LES against the time-averaged vorticity spectrum of the DNS for small wavenumbers lead to a much better behaved optimization problem and produced results easier to interpret physically. In particular, the general form of the optimal eddy viscosity obtained in this case was found to have little dependence on regularization, cf. Figure 3.5b.

The main question left open by the results reported here is whether the optimal eddy viscosity for the Smagorinsky model in 3D turbulent flows would exhibit similar properties. It can be studied by solving an optimization problem analogous to (3.14), a task we will undertake in the near future. In addition, it is also interesting to analyze the optimal performance of other closure models using the framework developed here. Furthermore, extending this work to understand the effect of the filter cutoff on the optimal eddy viscosities and its placement in the inverse energy cascade or forward enstrophy cascade ranges for a high resolution, fully developed turbulent simulation would provide a more complete picture of closure models for 2D turbulence.

Chapter 4

Enforcing Constraints Via Adjoint Analysis

4.1 Introduction

Now, we consider an extension of the problem in Chapter 3, cf. (3.14), where we will include an additional constraint to our nonstandard optimization problem. The optimization problem of determining eddy viscosities has a nonstandard structure due to the fact that our control variable has a functional dependence on the dependent variable. The optimization problem becomes further complicated in the presence of additional nonlinear constraints imposed on the eddy viscosities as considered here.

There are various forms in which constraints arise in optimization, with different approaches that can be used. Linear constraints typically define linear subspaces, in which constraints can be enforced via orthogonal projections. When nonlinear constraints are homogeneous, such as in Problem 1 in Chapter 2, then a simple rescaling or normalization of the control variable can be used to enforce the constraint, cf. (2.23). For general nonlinear constraints defining complex constraint manifolds, the best one can do is define a tangent space which can be projected onto. In the present problem, we

consider constraining the eddy viscosity such that the corresponding flow preserves an energy-type quantity in a certain sense, which defines a complex manifold in the space of eddy viscosities. Since characterization of this manifold through a suitable retraction operator is not feasible, we will need to resort to defining a tangent space. When dealing with intricate problems with constraints on energy quantities, methods such as Newton iterations can be employed (Farazmand and Sapsis 2017). Although this root finding algorithm can be used, this method is computationally very expensive for enforcing constraints on PDE optimization problems. Adjoint calculus has been used to solve PDE-constrained optimization problems, and has a long history beginning with (Lions 1968). Despite this, to the best of our knowledge, it has not previously been utilized for enforcing complicated nonhomogeneous constraints imposed on the control variable. In Figure 4.1, we illustrate the main idea behind projecting on the tangent space to approximately enforce constraints on a complex manifold, to be implemented in this work.

In this work, we thus introduce the use of adjoint calculus to characterize the tangent space to a constraint manifold, which allows one to approximately enforce constraints. Using this technique, the constraint can be directly enforced to ensure the optimal control remains "close" to the constraint manifold for each iteration of a gradient descent approach. We illustrate this novel technique by considering the problem from Chapter 3, cf. (3.14), in which the objective functional is modified to consider the time-averaged, least-squares error of the palinstrophy for a filtered DNS and LES field, with the additional constraint that the time-averaged enstrophy in the LES and in the filtered DNS be equal. We demonstrate that this constraint can be effectively enforced, though in an approximate sense, and also provide insights about the functional form of the optimal eddy viscosity depending on whether or not the constraint is used.

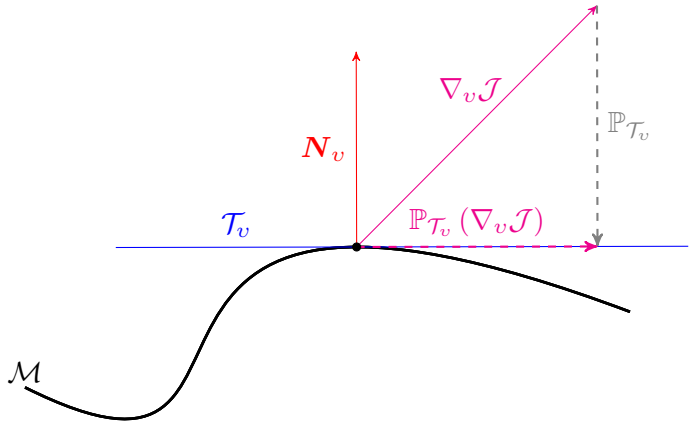


FIGURE 4.1: Schematic representation of the projection of the gradient $\nabla_v \mathcal{J}$ onto the manifold $\mathcal{T}_v$ tangent to a general manifold $\mathcal{M}$. The gradient $\nabla_v \mathcal{J}$ is projected onto the tangent space $\mathcal{T}_v$ via the normal components $\boldsymbol{N}_v$, to approximately enforce constraints represented by the general manifold $\mathcal{M}$, cf. (4.15)

The structure of this chapter is as follows: the next section formulates the general optimization problem, whereas Section 4.3 outlines the solution approach and Section 4.4 introduces the adjoint-based method to enforcing the additional constraint on the problem; our results are presented in Section 4.5, and a final discussion is provided in Section 4.6.

4.2 Optimization Problem

Here we are interested in the long-time average of energy-type quantities, whereas previously in Chapter 3 we considered pointwise match in space and in time to the DNS field, as well as in a certain statistical sense. This present problem is of particular importance in engineering applications, as time averages of an output of certain quantities are often of interest. We are interested in the palinstrophy

$$\mathcal{P}(t) = \frac{1}{2} \int_{\Omega} |\boldsymbol{\nabla} \widetilde{w}(t, \mathbf{x})|^2 \, d\mathbf{x}, \tag{4.1}$$

where here we define the palinstrophy of the filtered DNS field, i.e., the filtered solution of system (3.1). Using this quantity, we define the new objective functional $\mathcal{J}$ in the form of the least-squares error between the palinstrophy evolution $\mathcal{P}(t)$ and the corresponding palinstrophy evolution in the solution $\widetilde{\omega}$ of the LES problem (3.8) obtained for the given eddy viscosity $\nu = \nu(\upsilon)$ parameterized by the function $\upsilon \ : \ [0,1] \rightarrow \mathbb{R}$ (hence, υ plays a role analogous to the function φ in Chapter 3), i.e.,

$$\begin{aligned} \mathcal{J}(\upsilon) &:= \frac{1}{2\,D} \int_0^T \left[\mathcal{P}(t) - \frac{1}{2} \int_\Omega \left(\boldsymbol{\nabla} \widetilde{\omega}(t, \mathbf{x}; \upsilon) \right)^2 \, d\mathbf{x} \right]^2 dt, \\ &= \frac{1}{2\,D} \int_0^T \left[\mathcal{P}(t) - \widetilde{\mathcal{P}}(t; \upsilon) \right]^2 dt, \end{aligned} \tag{4.2}$$

where for simplicity we denote the palinstrophy in the LES as $\widetilde{\mathcal{P}}(t; \upsilon)$ and

$$D := \left[\int_0^T \mathcal{P}(t)^2 \, dt \right]^{\frac{3}{2}},$$

is a normalization factor. From Poincaré's inequality, we have that enstrophy is bounded by palinstrophy, thus it is natural to consider enstrophy as another quantity of interest. Now, suppose we wish to impose a constraint on the problem to ensure that as the eddy viscosity is updated, the evolution of the LES system (3.8) preserves the enstrophy, defined as

$$\widetilde{\mathcal{E}}(t, \upsilon) = \frac{1}{2} \int_\Omega \widetilde{\omega}(t, \mathbf{x}; \upsilon)^2 \, d\mathbf{x}, \tag{4.3}$$

in a time-averaged sense. As used in Chapter 3, we denote the time average of a function as $[f]_T := (1/T) \int_0^T f(t)\, dt$. Thus, we consider eddy viscosities parameterized by the function υ, such that the time-averaged enstrophy in the corresponding LES flow is given by $\mathcal{E}_0$, i.e.,

$$\left[\widetilde{\mathcal{E}}(\cdot; \upsilon) \right]_T = \frac{1}{2\,T} \int_0^T \int_\Omega \widetilde{\omega}(t, \mathbf{x}; \upsilon)^2 \, d\mathbf{x}\, dt = \mathcal{E}_0. \tag{4.4}$$

In this problem we set $\mathcal{E}_0 = \frac{1}{2T}\int_0^T \int_\Omega \widetilde{w}(t,\mathbf{x})^2\, d\mathbf{x}\, dt$, which is the time averaged enstrophy of the filtered DNS solution. Relation (4.4) thus defines a manifold in the subspace $\mathcal{S}$, cf. (3.13), namely,

$$\mathcal{M} := \left\{ \upsilon \in C^3([0,1]) \ : \ \left[\widetilde{\mathcal{E}}(\cdot;\upsilon)\right]_T = \mathcal{E}_0,\ \frac{d}{d\xi}\upsilon(\xi) = \frac{d^3}{d\xi^3}\upsilon(\xi) = 0 \text{ at } \xi = 0,1 \right\}. \tag{4.5}$$

We emphasize that this manifold in the space of eddy viscosities is defined implicitly through a property of the corresponding LES flows, namely, the requirement that their time-averaged enstrophy be equal to $\mathcal{E}_0$. Thus, we have the following problem for finding an optimal eddy viscosity.

Problem 3. *For the system (3.8) and objective functional (4.2), find*

$$\breve{\upsilon} := \underset{\upsilon\in\mathcal{M}}{\arg\min}\, \mathcal{J}(\upsilon), \tag{4.6}$$

where the optimal eddy viscosity $\breve{\nu}$ is obtained from the ansatz given in (3.7), with φ replaced with $\breve{\upsilon}$.

4.3 Solution Approach

As before, we iteratively determine a local maximizer of Problem 3 using the discrete gradient flow (3.15), suitably modified such that $\breve{\upsilon} = \lim_{n\to\infty} \upsilon^{(n)}$, where

$$\begin{cases} \upsilon^{(n+1)} &= \upsilon^{(n)} - \tau^{(n)}\, \mathbb{P}_{\mathcal{T}_\upsilon}\left(\nabla_\upsilon \mathcal{J}(\upsilon^{(n)})\right), \qquad n = 0,1,\ldots, \\ \upsilon^{(0)} &= \upsilon_0, \end{cases} \tag{4.7}$$

in which $\upsilon^{(n)}$ is the approximation of the optimal function $\breve{\upsilon}$ at the nth iteration (which lies "close" to the constraint manifold $\mathcal{M}$), $\nabla_\upsilon \mathcal{J}(\upsilon)$ is the gradient of the functional (4.2) with respect to υ, and $\tau^{(n)}$ is the step length along the descent direction, and υ_0

is an initial guess that lies on the constraint manifold $\mathcal{M}$. The projection operator $\mathbb{P}_{\mathcal{T}_\upsilon}$ projects onto the space tangent to the manifold $\mathcal{M}$ at υ, which makes it possible to approximately enforce the constraint (cf. Figure 4.1). It plays a key role in our approach and is characterized below. The gradient of the objective functional (4.2) is determined in an analogous way to what was done in Section 3.3, cf. (3.25). Namely, we first compute the Gâteaux differential, for which we obtain

$$\begin{aligned}
\mathcal{J}'(\upsilon;\upsilon') &:= \lim_{\epsilon\to 0}\frac{\mathcal{J}(\upsilon+\epsilon\upsilon')-\mathcal{J}(\upsilon)}{\epsilon} = \frac{d}{d\epsilon}\,\mathcal{J}\Big(\upsilon+\epsilon\upsilon'\Big)\Big|_{\epsilon=0}\\
&= -\int_0^T [\mathcal{P}(t)-\widetilde{\mathcal{P}}(t;\upsilon)]\int_\Omega \boldsymbol{\nabla}\widetilde{\omega}(t,\mathbf{x};\upsilon)\cdot\boldsymbol{\nabla}\widetilde{\omega}'(t,\mathbf{x};\upsilon,\upsilon')\,d\mathbf{x}\,dt,\\
&= -\int_0^T\int_\Omega [\mathcal{P}(t)-\widetilde{\mathcal{P}}(t;\upsilon)]\boldsymbol{\nabla}\widetilde{\omega}(t,\mathbf{x};\upsilon)\cdot\boldsymbol{\nabla}\widetilde{\omega}'(t,\mathbf{x};\upsilon,\upsilon')\,d\mathbf{x}\,dt,\\
&= \int_0^T\int_\Omega [\mathcal{P}(t)-\widetilde{\mathcal{P}}(t;\upsilon)]\boldsymbol{\Delta}\widetilde{\omega}(t,\mathbf{x};\upsilon)\,\widetilde{\omega}'(t,\mathbf{x};\upsilon,\upsilon')\,d\mathbf{x}\,dt, \qquad (4.8)
\end{aligned}$$

where υ' is an arbitrary perturbation of $\upsilon\in\mathcal{M}$, and $\widetilde{\omega}'(t,\mathbf{x};\upsilon,\upsilon')$ satisfies the corresponding linear perturbation system (3.17). Following the same procedure as in Chapter 3, we combine (3.17), (3.21), (3.22), (4.8), and (3.23) with the source term redefined as $W(t,\mathbf{x}) = \Big(\mathcal{P}(t)-\widetilde{\mathcal{P}}(t)\Big)\,\boldsymbol{\Delta}\widetilde{\omega}$ to obtain an expression for the L^2 gradient of the objective functional. Then, solving (3.26) provides us with a H^2 Sobolev gradient in the space $\mathcal{S}$, which however still needs to be projected onto the space tangent to the manifold $\mathcal{M}$ at υ, cf. Figure 4.1.

4.4 Enstrophy Constraint

In order to approximately enforce our enstrophy constraint, we need to characterize the manifold $\mathcal{M}$ locally by constructing a tangent subspace $\mathcal{T}_\upsilon$ at a given $\upsilon\in\mathcal{M}$. Since both the manifold and its tangent bundle have co-dimension 1, this will be done by constructing a "normal vector" $\boldsymbol{N}_\upsilon$, which can be used to project onto the tangent subspace of the constraint manifold (see Figure 4.1). As we are unable to construct a

computable retraction $R_{\mathcal{M}} : \mathcal{T}_{\upsilon} \to \mathcal{M}$ from the tangent subspace $\mathcal{T}_{\upsilon}$ to the constraint manifold $\mathcal{M}$, at each iteration in (4.7) the constraint will be satisfied only approximately with errors $\mathcal{O}(\tau^{(n)})$. This tangent subspace is constructed utilizing an innovative use of adjoint calculus, which we will now describe.

Consider the Gâteaux differential of the constraint (4.4) with respect to υ

$$\begin{aligned}\Big[\widetilde{\mathcal{E}}'(\cdot;\upsilon,\upsilon')\Big]_T &:= \lim_{\epsilon\to 0}\frac{\Big[\widetilde{\mathcal{E}}(\cdot;\upsilon+\epsilon\upsilon')-\widetilde{\mathcal{E}}(\cdot;\upsilon)\Big]_T}{\epsilon} = \frac{d}{d\epsilon}\ \Big[\widetilde{\mathcal{E}}(\cdot;\upsilon+\epsilon\upsilon')\Big]_T\Big|_{\epsilon=0}\\ &= \frac{1}{T}\int_0^T\int_{\Omega}\widetilde{\omega}(t,\mathbf{x};\upsilon)\,\widetilde{\omega}'(t,\mathbf{x};\upsilon,\upsilon')\,d\mathbf{x}\,dt = 0,\end{aligned} \tag{4.9}$$

noting this is equal to zero since the Gateaux differential of $\mathcal{E}_0$ vanishes. As before, the perturbation υ' does not explicitly appear as a factor in the Gâteaux differential (4.9) but is hidden in the source term in the linear system governing the evolution of $\widetilde{\omega}'$. We note that (4.9) defines a bounded linear functional of $\upsilon' \in \mathcal{S}$ and in order to transform it to the Riesz form with υ' appearing explicitly as a factor, we introduce new scalar *adjoint fields* $\breve{\omega}^*$ and $\breve{\psi}^*$ which satisfy the judiciously chosen *adjoint system*

$$\mathcal{F}^*\begin{bmatrix}\breve{\omega}^*\\ \\ \breve{\psi}^*\end{bmatrix} := \begin{bmatrix}-\partial_t\breve{\omega}^* - \boldsymbol{\nabla}^{\perp}\widetilde{\psi}\cdot\boldsymbol{\nabla}\breve{\omega}^* + \alpha\breve{\omega}^* + \breve{\psi}^*\\ -\boldsymbol{\nabla}\cdot\Big(2\,(\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\breve{\omega}^*)\,(\frac{d\nu}{ds}\,\upsilon\boldsymbol{\nabla}\widetilde{\omega} + \frac{\nu_L+\nu_0}{s_{\max}}\,\frac{d\upsilon}{d\sigma}\,\boldsymbol{\nabla}\widetilde{\omega}) + (\nu_N+\nu)\boldsymbol{\nabla}\breve{\omega}^*\Big)\\ \boldsymbol{\Delta}\breve{\psi}^* - \boldsymbol{\nabla}^{\perp}\cdot(\breve{\omega}^*\,\boldsymbol{\nabla}\widetilde{\omega})\end{bmatrix} = \begin{bmatrix}\\ \frac{1}{T}\,\widetilde{\omega}\\ \\ 0\end{bmatrix}, \tag{4.10a}$$

$$\breve{\omega}^*(t=T) = 0, \tag{4.10b}$$

also subject to periodic boundary conditions. With the new systems introduced here, we use the following duality-pairing relation

$$\left(\mathcal{K}\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix},\begin{bmatrix}\breve{\omega}^*\\ \breve{\psi}^*\end{bmatrix}\right) := \int_0^T\int_\Omega \mathcal{K}\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}\cdot\begin{bmatrix}\breve{\omega}^*\\ \breve{\psi}^*\end{bmatrix}\,d\mathbf{x}\,dt = \int_0^T\int_\Omega \boldsymbol{\nabla}\cdot\left((\nu_L+\nu_0)\upsilon'\boldsymbol{\nabla}\widetilde{\omega}\right)\,\breve{\omega}^* d\mathbf{x}\,dt,$$

$$\int_0^T\int_\Omega \begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}\cdot\mathcal{F}^*\begin{bmatrix}\breve{\omega}^*\\ \breve{\psi}^*\end{bmatrix}\,d\mathbf{x}\,dt = \underbrace{\int_0^T\int_\Omega \frac{1}{T}\,\widetilde{\omega}\,\widetilde{\omega}'\,d\mathbf{x}dt}_{\left[\widetilde{\mathcal{E}}'(\cdot;\upsilon,\upsilon')\right]_T}, \tag{4.11}$$

where integration by parts with respect to both space and time was performed with all the boundary terms vanishing due to periodicity, and the initial and terminal conditions (3.17c) and (4.10b). Thus, we have obtained an expression for the Gâteaux differential

$$\left[\widetilde{\mathcal{E}}'(\cdot;\upsilon,\upsilon')\right]_T = -\int_0^T\int_\Omega (\nu_L+\nu_0)\,(\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\breve{\omega}^*)\,\upsilon'\,d\mathbf{x}\,dt, \tag{4.12}$$

with the perturbation υ' now appearing as a factor. However this is still not consistent with the Riesz form since it requires integration with respect to s rather than space and time. Again, we make an appropriate change of variables by representing our perturbation as $\upsilon'(\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\widetilde{\omega}) = \int_0^1 \delta\left(\frac{\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\widetilde{\omega}}{s_{\max}}-\sigma\right)\upsilon'(\sigma)\,d\sigma$. This, along with Fubini's theorem, which justifies swapping the order of integration, allows us to transform the Gâteaux differential (4.12) into the required Riesz form

$$\begin{aligned}\left[\widetilde{\mathcal{E}}'(\cdot;\upsilon,\upsilon')\right]_T &= \int_0^1\left[-\int_0^T\int_\Omega \delta\left(\frac{\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\widetilde{\omega}}{s_{\max}}-\sigma\right)(\nu_L+\nu_0)\,\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\breve{\omega}^*\,d\mathbf{x}\,dt\right]\upsilon'(\sigma)\,d\sigma,\\ &= \left\langle \boldsymbol{N}_\upsilon^{L^2},\upsilon'\right\rangle_{L^2([0,1])} = \left\langle \boldsymbol{N}_\upsilon^{H^2},\upsilon'\right\rangle_{H^2([0,1])} = 0.\end{aligned} \tag{4.13}$$

Relation (4.13) defines a subspace $\mathcal{T}_\upsilon$ tangent to the constraint manifold $\mathcal{M}$ at υ, cf. Figure 4.1. To determine the normal vector $\boldsymbol{N}_\upsilon^{H^2}$ characterizing this subspace, we first

obtain the normal vector defined with respect to the L^2 topology

$$\boldsymbol{N}_{\upsilon}^{L^2}(\sigma) = -\int_0^T \int_\Omega \delta\left(\frac{\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\widetilde{\omega}}{s_{\max}} - \sigma\right)(\nu_L + \nu_0)\,\boldsymbol{\nabla}\widetilde{\omega}\cdot\boldsymbol{\nabla}\breve{\omega}^*\,d\mathbf{x}\,dt. \tag{4.14}$$

Then $\boldsymbol{N}_{\upsilon}^{H^2}$ can be determined by solving system (3.26). Projection on the tangent subspace $\mathcal{T}_\upsilon$ is then achieved using the projection operator $\mathbb{P}_{\mathcal{T}_\upsilon} : \mathcal{S} \to \mathcal{T}_\upsilon$ defined as

$$\mathbb{P}_{\mathcal{T}_\upsilon}\left(\nabla_{\upsilon}^{H^2}\mathcal{J}\right) = \nabla_{\upsilon}^{H^2}\mathcal{J} - \zeta\,\boldsymbol{N}_{\upsilon}^{H^2}, \tag{4.15}$$

where ζ

$$\zeta = \frac{\left\langle \nabla_{\upsilon}^{H^2}\mathcal{J}, \boldsymbol{N}_{\upsilon}^{H^2}\right\rangle_{H^2([0,1])}}{\left\langle \boldsymbol{N}_{\upsilon}^{H^2}, \boldsymbol{N}_{\upsilon}^{H^2}\right\rangle_{H^2([0,1])}}, \tag{4.16}$$

can be interpreted as a Lagrange multiplier. Thus, assuming that the initial guess υ_0 satisfies the constraint given in (4.4), the projected gradient flow (4.7) ensures the iterates $\upsilon^{(n+1)}$ remain close the constraint manifold $\mathcal{M}$, with an error in the satisfaction of the constraint of the order $\mathcal{O}(\tau^{(n)})$.

4.5 Results

Here we illustrate the solution of Problem 3 using the projected gradient flow (4.7) with the projection operator defined in (4.15). We emphasize that we only use a standard gradient approach, with no conjugate-gradient method utilized, as these methods are harder to implement in the presence of general nonlinear constraints. The numerical approach is the same as outlined in Section 3.4, except that we use a third-order IMEX method introduced by Alimo et al. (2021) (described in Appendix A) for time integration of the LES system (3.8) and the adjoint system (3.22), instead of the CN/RKW3 previously used. As a target we consider solutions of system (3.1) with the following

parameters: $\nu_N = 6 \times 10^{-3}$, $\alpha = 5 \times 10^{-3}$, $F = 5$, and $k_a = k_b = 4$. For computational parameters, use $N_x = 256$ equispaced points in each direction to discretize the spatial domain, set the wavenumber cutoff $k_c = 5$, fix the number of discretization points in the state space $N_s = 128$, and $T = 50 \approx 30t_e$, where again we define the eddy turnover time as $t_e := \left[\int_0^T \mathcal{E}(t)\, dt/(8\pi^2 T)\right]^{-1/2}$ (Bracco and McWilliams 2010). We set the regularization parameters $\ell_1 = 100$ and $\ell_2 = 10$, to be used to determine the gradient $\nabla_\upsilon^{H^2}\mathcal{J}$ and "normal vector" $\boldsymbol{N}_\upsilon^{H^2}$, for the H^2 topology via (3.26). Although we solve our minimization problem with respect to the nondimensional function υ, we will focus here on the corresponding optimal eddy viscosities. In (4.7), we use a standard Leith-type initial guess for the eddy viscosity

$$\nu_0(s) = (C_l\, k_c)^3 \sqrt{s + \gamma}, \tag{4.17}$$

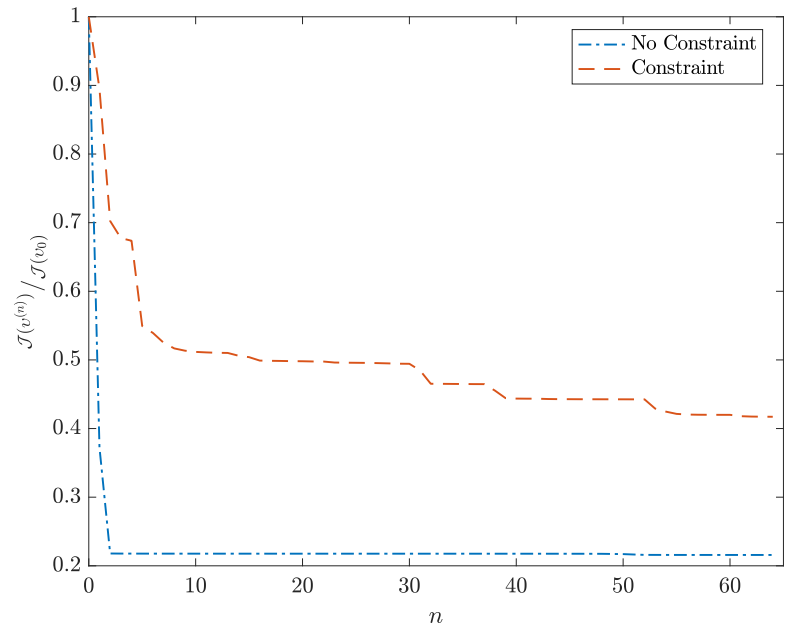
where the Leith constant $C_l = 1.0992 \times 10^{-2}$ was chosen such that the eddy viscosity is on the constraint manifold $\mathcal{M}$. Due to the complex and nonstandard structure nature of this problem, we demonstrate the validity our gradient $\nabla_\upsilon^{H^2}\mathcal{J}$ and of the "normal vector" $\boldsymbol{N}_\upsilon^{H^2}$ in Appendix A3.

To illustrate how the constraint is enforced, we solve the minimization problem with and without the enstrophy constraint. That is, we solve (4.6) with $\upsilon \in \mathcal{S}$ and $\upsilon \in \mathcal{M}$, defined in (3.13) and (4.5), respectively. The decrease of the error functional $\mathcal{J}(\upsilon^{(n)})$ is shown in Figure 4.2a, where we notice less of a decrease when the enstrophy constraint is imposed. As $\mathcal{M}$ is a manifold imbedded the subspace $\mathcal{S}$, this restricts the choice of eddy viscosities so it is expected that the eddy viscosities constrained to the manifold $\mathcal{M}$ produce a smaller decrease of the objective functional. The eddy viscosities obtained from both optimization problems are shown in Figure 4.2b, where it is clear that the eddy viscosity constrained to the manifold $\mathcal{M}$ reveals oscillations which are required in order to enforce the constraint. Interestingly, we note that the qualitative features of the

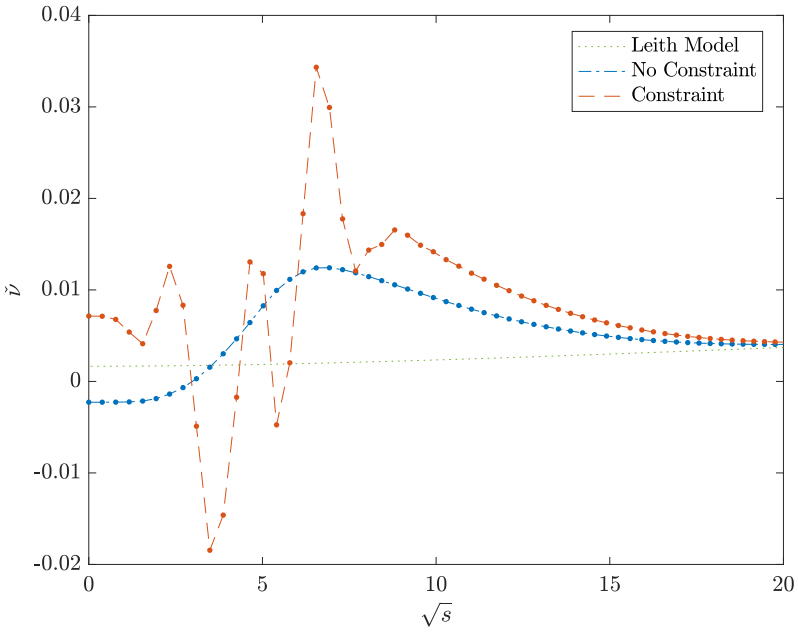
optimal eddy viscosity are different when the constraint is imposed. Most notably, for small values of s the optimal eddy viscosity achieves positive values when the enstrophy constraint is imposed, and negative values when no constraint is applied. As shown in Figure 3.1b and Figure 3.5b, small values of s are attained more frequently in the flow. Thus, this indicates that in order to match the flows palinstrophy accurately, the eddy viscosity must inject energy into the system for small values of s, while dissipating energy for small values of s will result in a better match of the time-averaged enstrophy.

To showcase the effectiveness of the method, we consider the time-averaged enstrophy $\left[\widetilde{\mathcal{E}}(\cdot;\upsilon)\right]_T$, cf. (4.4), over the "training window" T, and normalized by $\mathcal{E}_0$ in Figure 4.3a. To observe the effectiveness of eddy viscosities beyond the "training window", we consider the time-averaged enstrophy over twice the "training window", $\left[\widetilde{\mathcal{E}}(\cdot;\upsilon)\right]_{2T}$ and normalized by $\mathcal{E}_{2T} = \frac{1}{2\,(2T)}\int_0^{2T}\int_\Omega \widetilde{w}(t,\mathbf{x})^2\,d\mathbf{x}\,dt$ as a function of iteration in Figure 4.3b. It is evident that imposing the enstrophy constraint via the projected gradient flow (4.7) results in a value of the time-averaged enstrophy over T that remains close to target value of $\mathcal{E}_0$. This approach performs noticeably better than the "no constraint" optimal eddy viscosity, over both time windows. In addition, we visualize the enstrophy (4.3) and palinstrophy (4.1) as functions of time in Figure 4.4a and Figure 4.4b, respectively, with the filtered DNS provided on these plots for comparison to the target field. The optimal eddy viscosity determined with constraint, on average, maintains a value of enstrophy closer to the target of $\mathcal{E}_0$ compared to when no constraint is used. On the other hand, the optimal eddy viscosity trained with no constraint provides closer palinstrophy values to the filtered DNS. These comparisons of the enstrophy and palinstrophy are both as expected, as the constraint is enforced on the enstrophy, however no constraint achieves a better decrease of the error functional. Finally, we compare the time evolution of the normalized correlation $\mathcal{C}(t)$ from (3.28), and the adjusted correlation $\log_{10}|1-\mathcal{C}(t)|$ in Figure 4.5a and Figure 4.5b, respectively, to assess how well the optimal eddy viscosities

combat the exponential divergence of trajectories for the flow. The divergence of trajectories occur roughly at the same rate for each case over the "training window". However beyond this window, the eddy viscosity obtained using the constraint takes longer to become uncorrelated.

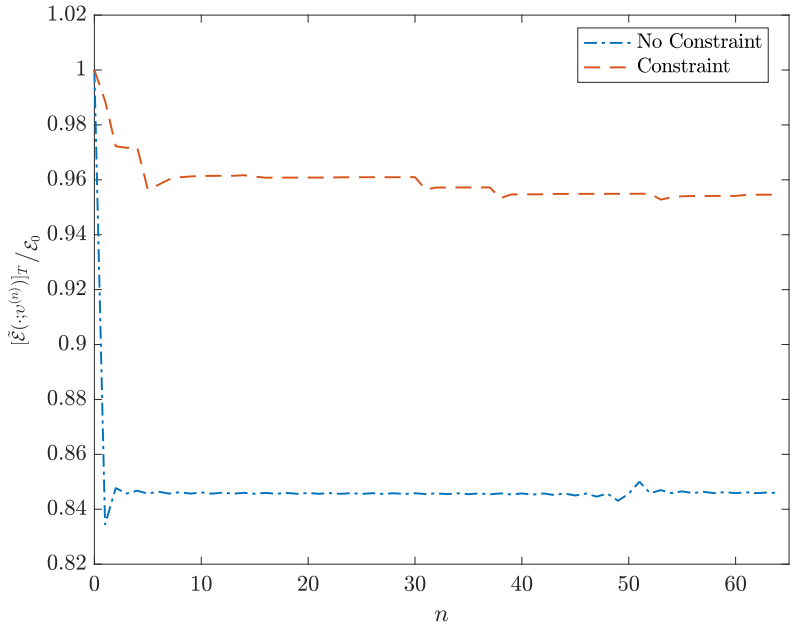


(a)

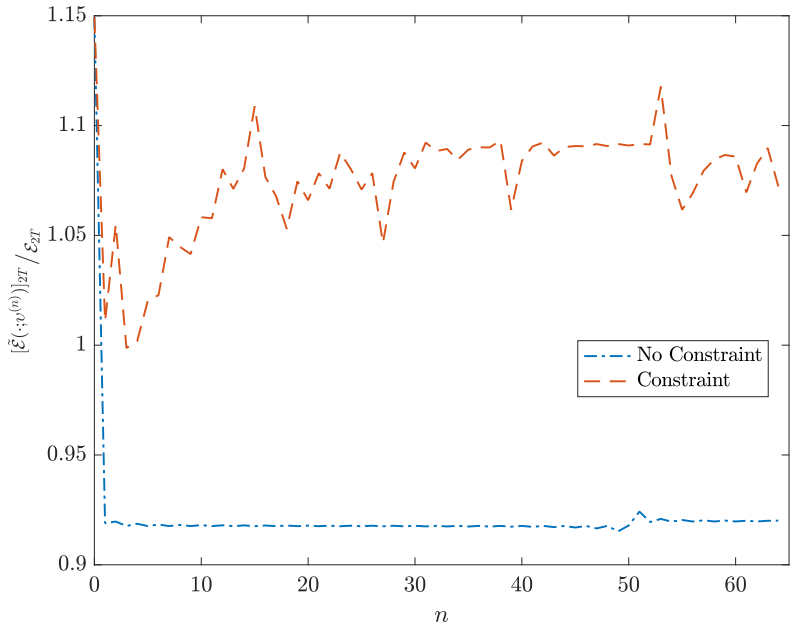


(b)

FIGURE 4.2: (a) Dependence of the normalized functional $\mathcal{J}(\upsilon^{(n)})/\mathcal{J}(\upsilon_0)$, on the iteration n and (b) dependence of the corresponding optimal eddy viscosity $\check{\nu}$ on $\sqrt{s}$ for (blue, dashed-dot line) $\upsilon \in \mathcal{S}$ and (red, dashed line) $\upsilon \in \mathcal{M}$. Panel (b) also shows the standard Leith-type initial guess, cf. (4.17).

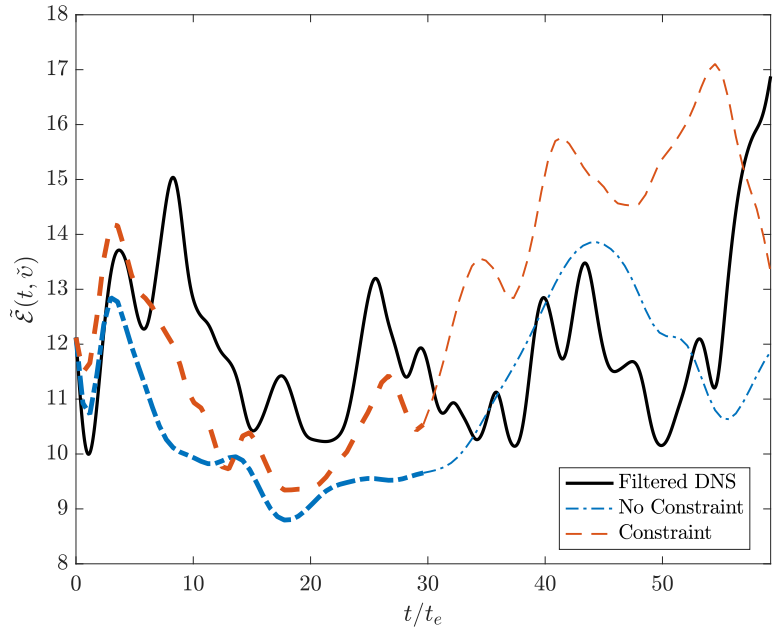


(a)

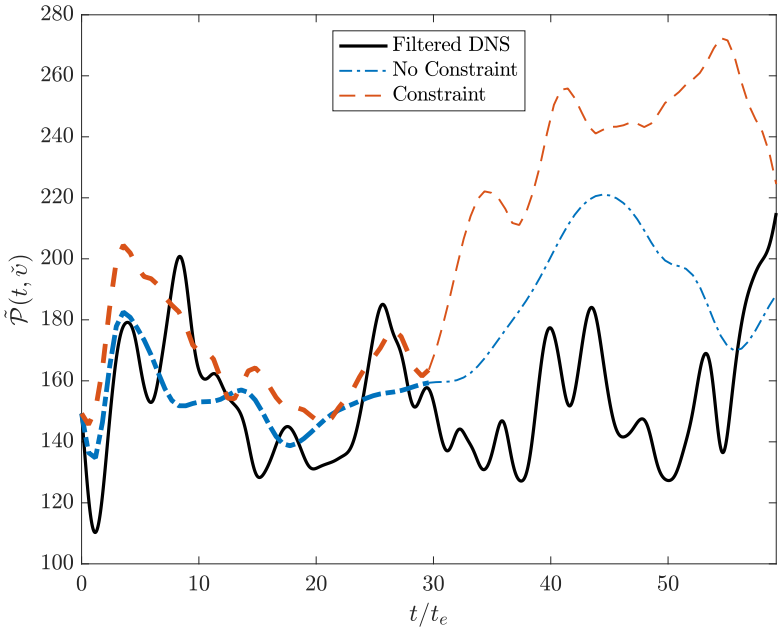


(b)

FIGURE 4.3: Normalized time-averaged enstrophy (4.4) for (a) the "training window" (T) and (b) twice the "training window" ($2T$). The time-averaged enstrophy is shown for (blue, dashed-dot line) $v \in \mathcal{S}$ and (red, dashed line) $v \in \mathcal{M}$.

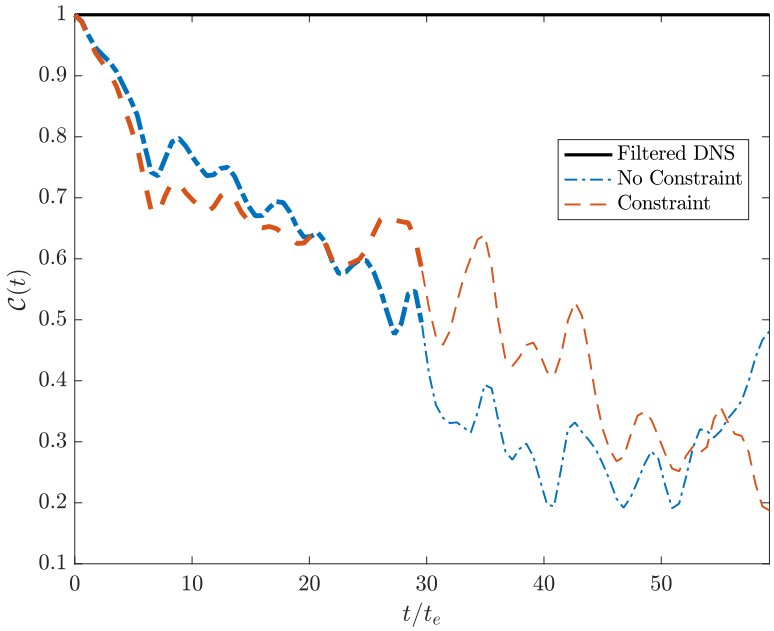


(a)

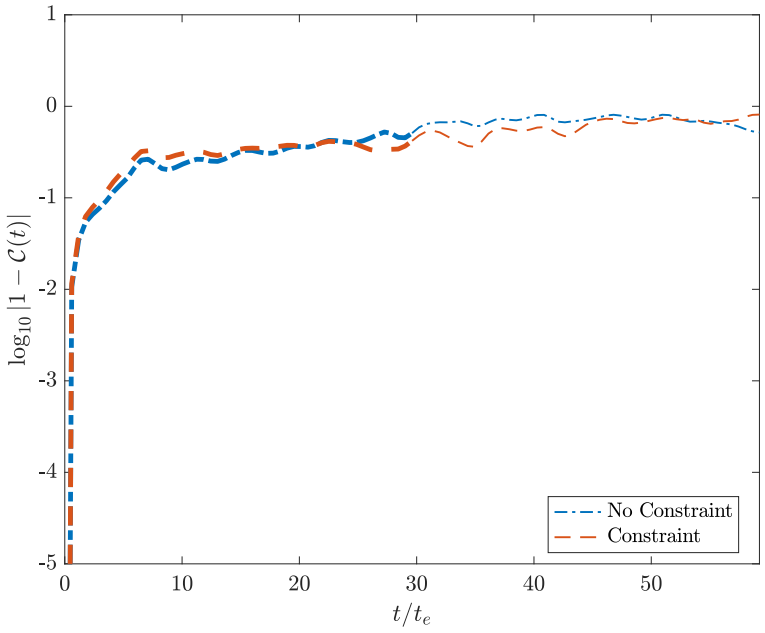


(b)

FIGURE 4.4: (a) Enstrophy (4.3) and (b) palinstrophy (4.1) for (black, solid line) the filtered DNS, (blue, dashed-dot line) optimal $v \in \mathcal{S}$, and (red, dashed line) optimal $v \in \mathcal{M}$. Thick and thin lines correspond to, respectively, time in the "training window" ($t \in [0, T]$) and twice this window ($t \in (T, 2T]$).



(a)



(b)

FIGURE 4.5: (a) Normalized correlation (3.28), $\mathcal{C}(t)$, and (b) adjusted correlation $\log_{10}|1-\mathcal{C}(t)|$, for (black, solid line) the filtered DNS, (blue, dashed-dot line) optimal $\upsilon \in \mathcal{S}$, and (red, dashed line) optimal $\upsilon \in \mathcal{M}$. Thick and thin lines correspond to, respectively, time in the "training window" ($t \in [0, T]$) and twice this window ($t \in (T, 2T]$).

4.6 Discussion

We introduced a computationally flexible framework for determining optimal eddy-viscosity closure models subject to nonlinear constraints that define complex manifolds. Enforcing these complex constraints requires searching for eddy viscosities on a constraint manifold, which was implemented utilizing novel tools based on adjoint calculus. To illustrate these concepts, we consider an error functional that represents the least-squares error between a target field (palinstrophy $\mathcal{P}(t)$ of a filtered DNS field) and the palinstrophy of the LES with eddy-viscosity closure model, averaged over all time. Although this is already a nonstandard problem, we add the additional nonlinear constraint that requires the eddy viscosities to produce LES flows with the same value of the time-averaged enstrophy, hence restricting our search of eddy viscosities to a constraint manifold.

First, we compute the gradient of the error functional, by using an adjoint-based gradient approach outlined in Chapter 3, suitably modified. As a key innovation, the constraint on enstrophy, which can only be approximately enforced, requires the introduction of a subspace tangent to the constraint manifold described in terms of a suitable "normal vector". It was determined via an adjoint-based method, and allowed us to define a projected gradient flow (4.7) where each iterate $\upsilon^{(n+1)}$ was constructed in the subspace tangent to the manifold, thus making it possible to satisfy the nonlinear constraint approximately.

It was shown in Section 4.5 that approximately enforcing this constraint via projections onto the tangent subspace does indeed maintain a more consistent value of the time-averaged enstrophy, compared to when no constraint is used. The eddy-viscosity determined by approximately enforcing the constraint also clearly provided a better match to the filtered DNS for the enstrophy over time, but not for the palinstrophy, as is expected considering that the constraint is defined in terms of enstrophy values.

Interestingly, a comparison of the eddy viscosities determined with and without the constraint enforcement also offers insights on the functional form of the optimal eddy viscosity to better match the palinstrophy or enstrophy, depending on which is a more crucial quantity in an application.

Although the application in this study is for determining eddy-viscosity closure models, the main contribution of this work is the implementation of a complex nonlinear constraint in a nonstandard optimization problem. The introduction of this novel method allows one to approximately enforce such constraints using a generalization of the adjoint approach. We emphasize that the adjoint system (4.10), which needs to be solved in order to determine the normal vector $\boldsymbol{N}_\upsilon$ and the associated projection operator (4.15), is defined in terms of the same adjoint operator, but with a different source term, as the adjoint system (3.22), which has to be solved in order to determine the gradient of the objective functional. This greatly simplifies the implementation of the proposed approach.

Chapter 5

Extensions and Open Problems

In this chapter, we describe some exciting and interesting extensions of the problems considered in the previous sections. These problems are formulated to probe some outstanding questions in turbulence, and pathways towards solutions are outlined, as results for these problems have not been finalized yet.

5.1 Enstrophy Dissipation in 2D from the Perspective of the Kinetic Theory

5.1.1 Overview

The problem of "anomalous dissipation" has traditionally been addressed from the point of view of macroscopic fluids, by considering the Navier-Stokes flow in the inviscid limit. To study fluids at the microscopic level one usually adopts the perspective of the kinetic theory which relies on classical models such as the Boltzmann equation. In the hydrodynamic limit, the solutions of the Boltzmann equation have been shown to converge to solutions of the Navier-Stokes equations (Bardos et al. 1991), including the case when a model collisional kinetic equation is considered in place of the Boltzmann equation with a

general collision kernel (Saint-Raymond 2003). In this work, we aim to study the enstrophy dissipation problem considered in Chapter 2 from the perspective of kinetic theory. Below, we formulate an optimization problem allowing one to study maximum enstrophy dissipation, at the microscopic level. To do this, we consider a fluid governed by the Bhatnagar–Gross–Krook (BGK) Boltzmann equation, and set up a problem analogous to the previous work in Chapter 2, which considered the 2D Navier-Stokes equation. We then outline a proposed solutions approach for this problem. When completed, this work will provide a novel and complementary perspective on this problem and the driving mechanisms to produce the largest enstrophy dissipation. Furthermore, we hope to further probe the optimality of the initial data discovered in Chapter 2.

5.1.2 Introduction to the BGK Boltzmann Equation

Here, we provide an informal introduction to the simplified BGK collision operator in the Boltzmann equation, rather than a rigorous introduction to the classical Boltzmann equation. Although this model is simple compared to the Boltzmann equation with binary collisions, and is not truly a collisional model (Saint-Raymond 2003), it is still widely used and is typically considered for numerical simulations. The primary reason to consider a BGK collision operator is due to its simplification of numerical simulations, where a $D2Q9$ scheme can easily be implemented (Krüger et al. 2017). To introduce the BGK Boltzmann equation we denote the Knudsen number, $\varepsilon > 0$, defined as the ratio of Mach number over the Reynolds number. This acts as a small parameter in our system that describes the ratio of the mean free path of a particle before a collision and the characteristic length scale in the flow. Then a cloud of particles can be described via its number density f_ε, which we consider as a function of space over a 2D periodic domain $\Omega := \mathbb{T}^2 = [0,1]^2$ and of velocity $\boldsymbol{v} \in \Theta := \mathbb{R}^2$. Thus, the microscopic interactions of fluid particles with an initial condition $f_\varepsilon^0(\mathbf{x}, \boldsymbol{v})$ can be described by the Boltzmann equation,

with the BGK collision operator,

$$\varepsilon\partial_t f_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} f_\varepsilon = \frac{1}{\nu\varepsilon}\Big(f_\varepsilon^{\text{eq}} - f_\varepsilon\Big) \qquad \text{in } (0,T]\times\Omega\times\Theta, \tag{5.1a}$$

$$f_\varepsilon(t=0) = f_\varepsilon^0, \qquad \text{in } \Omega\times\Theta, \tag{5.1b}$$

where

$$f_\varepsilon^{\text{eq}}(t,\mathbf{x},\boldsymbol{v}) := \frac{R_\varepsilon}{2\pi T_\varepsilon}\cdot e^{-\big||\boldsymbol{v}|-U_\varepsilon\big|^2/(2T_\varepsilon)},$$

ν is the relaxation parameter, and $R_\varepsilon, T_\varepsilon, U_\varepsilon$ satisfy

$$\begin{aligned} R_\varepsilon(t,\mathbf{x}) &:= \int_\Theta f_\varepsilon(t,\mathbf{x},\boldsymbol{v})\,d\boldsymbol{v},\\ R_\varepsilon(t,\mathbf{x})U_\varepsilon(t,\mathbf{x}) &:= \int_\Theta \boldsymbol{v} f_\varepsilon(t,\mathbf{x},\boldsymbol{v})\,d\boldsymbol{v},\\ R_\varepsilon(t,\mathbf{x})|U_\varepsilon(t,\mathbf{x})|^2 + 2R_\varepsilon(t,\mathbf{x})T_\varepsilon(t,\mathbf{x}) &:= \int_\Theta |\boldsymbol{v}|^2 f_\varepsilon(t,\mathbf{x},\boldsymbol{v})\,d\boldsymbol{v}. \end{aligned}$$

Let M denote the Maxwellian defined as

$$M(\boldsymbol{v}) := \frac{1}{2\pi}e^{-|\boldsymbol{v}|^2/2}, \quad \boldsymbol{v}\in\Theta.$$

In Saint-Raymond (2003), it is stated that if we consider the solution to (5.1) in the form of a fluctuation near the Maxwellian, hence $f_\varepsilon = M\,(1+\varepsilon g_\varepsilon)$, then as $\varepsilon\to 0$ the Boltzmann equation that g_ε satisfies converges to the Navier-Stokes equation in hydrodynamic limit. More precisely, the Boltzmann equation that the fluctuation g_ε satisfies is

$$\varepsilon\partial_t g_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} g_\varepsilon = \frac{1}{\nu\varepsilon}\Big(\Pi(g_\varepsilon) - g_\varepsilon\Big) + O(\varepsilon^2) \qquad \text{in } (0,T]\times\Omega\times\Theta,$$

where

$$\Pi(g_\varepsilon) := \langle g_\varepsilon \rangle_M + \boldsymbol{v} \cdot \langle g_\varepsilon \boldsymbol{v} \rangle_M + \left\langle g_\varepsilon \left(\frac{|\boldsymbol{v}|^2}{2} - 1 \right) \right\rangle_M \left(\frac{|\boldsymbol{v}|^2}{2} - 1 \right),$$

with $\langle \cdot \rangle_M$ denoting the average over measure $M\, d\boldsymbol{v}$, i.e.,

$$\langle \Xi \rangle_M = \int_\Theta \Xi(\boldsymbol{v}) M(\boldsymbol{v})\, d\boldsymbol{v},$$

for any integrable function $\Xi = \Xi(\boldsymbol{v})$. As $\varepsilon \to 0$,

$$\langle g_\varepsilon \boldsymbol{v} \rangle_M \to \boldsymbol{u}, \quad \left\langle g_\varepsilon (|\boldsymbol{v}|^2 - 4) \right\rangle_M \to 4\theta, \tag{5.2}$$

where u and θ satisfy the Navier-Stokes-Fourier equations

$$\partial_t \boldsymbol{u} + \boldsymbol{\nabla} \cdot (\boldsymbol{u} \otimes \boldsymbol{u}) + \boldsymbol{\nabla} \Pi = \nu \boldsymbol{\Delta} \boldsymbol{u},$$
$$\partial_t \theta + \boldsymbol{\nabla} \cdot (\theta \boldsymbol{u}) = \nu \boldsymbol{\Delta} \theta.$$

From above, we notice that the approach in Saint-Raymond (2003) is linear as it relies on truncation at $\mathcal{O}(\varepsilon^2)$. To retain some nonlinearity in our BGK model, we first simplify $f_\varepsilon^{\text{eq}}$ by using the Hermite expansion in $\boldsymbol{v}$ up to $\mathcal{O}(|\boldsymbol{v}|^2)$, i.e.,

$$f_\varepsilon^{\text{eq}} \approx M\, R_\varepsilon \left(1 + \boldsymbol{v} \cdot U_\varepsilon + (\boldsymbol{v} \cdot U_\varepsilon)^2 - \Big| U_\varepsilon \Big|^2 \right), \tag{5.3}$$

assuming the system to be isothermal. Substituting $f_\varepsilon = M\,(1 + \varepsilon g_\varepsilon)$ in (5.1) with (5.3) and truncating terms of order $\mathcal{O}(\varepsilon^2)$, which do not contribute to the hydrodynamic limit as $\varepsilon \to 0$, we obtain that the fluctuation g_ε satisfies the Boltzmann equation with a

modified BGK model

$$\varepsilon\partial_t g_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} g_\varepsilon = \frac{1}{\nu\varepsilon}\Big(g_\varepsilon^{\text{eq}} - g_\varepsilon\Big) \qquad \text{in } (0,T]\times\Omega\times\Theta, \tag{5.4a}$$

$$g_\varepsilon(t=0) = \eta, \qquad \text{in } \Omega\times\Theta, \tag{5.4b}$$

where η is the initial condition and

$$g_\varepsilon^{\text{eq}} := \rho_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{u}_\varepsilon + \varepsilon\Big((\boldsymbol{v}\cdot\boldsymbol{u}_\varepsilon)^2 - |u_\varepsilon|^2\Big), \tag{5.5}$$

with

$$\rho_\varepsilon := \langle g_\varepsilon\rangle_M, \tag{5.6}$$

$$\boldsymbol{u}_\varepsilon := \langle \boldsymbol{v} g_\varepsilon\rangle_M. \tag{5.7}$$

By following the derivation in Saint-Raymond (2003), it can be shown that as $\varepsilon\to 0$, $\boldsymbol{u}_\varepsilon \to \boldsymbol{u}$ in a weak sense where $\boldsymbol{u}$ satisfies the Navier-Stokes equation

$$\partial_t\boldsymbol{u} + (\boldsymbol{u}\cdot\boldsymbol{\nabla})\boldsymbol{u} - \boldsymbol{\Delta}\boldsymbol{u} + \boldsymbol{\nabla}\Pi = 0,$$

$$\boldsymbol{\nabla}\cdot\boldsymbol{u} = 0.$$

5.1.3 Optimization Problem

Let $S := -\boldsymbol{v}\cdot\boldsymbol{\nabla}^\perp = v_2\partial_{x_1} - v_1\partial_{x_2}$, where $\boldsymbol{v} = (v_1, v_2)^T$. Our main interest in this work is maximizing the functional

$$\mathcal{J}_\varepsilon(\eta) = \frac{\nu}{T}\int_0^T\int_\Omega \Big|\boldsymbol{\nabla}\int_\Theta S g_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta)\, M(\boldsymbol{v})\, d\boldsymbol{v}\Big|^2\, d\mathbf{x}\, dt. \tag{5.8}$$

Using (5.7), and that in the limit as $\varepsilon\to 0$, $\boldsymbol{u}_\varepsilon$ converges in a weak sense to the velocity field $\boldsymbol{u}$ in the Navier-Stokes equation, then we can conclude that the expression

$\int_\Theta Sg_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta)\,M(\boldsymbol{v})\,d\boldsymbol{v}$, which is a function of time t and space $\mathbf{x}$, is analogous to the vorticity field of the 2D Navier-Stokes equation, cf. (1.3). Therefore, expression (5.8) can be considered as the enstrophy dissipation per unit time, which depends on the initial distribution η in (5.4b).

The goal of this work is to study the connection between optimal initial distributions η maximizing functional (5.8) and the optimal initial conditions presented in Section 2.4. Thus, we consider an optimization problem analogous to Problem 1 to study the rate in which enstrophy dissipation vanishes in the limit as $\nu \to 0$, now in the setting of the kinetic theory. In order to study this problem in the setting of the kinetic theory, we must impose an additional constraint to ensure conservation of mass, where the initial mass is equal to a constant, m_0. In addition, let $L^2_M(\Theta)$ be the weighted L^2 space defined as

$$L^2_M(\Theta) := \left\{ \left(\int_\Theta |\Xi(\boldsymbol{v})|^2\, M(\boldsymbol{v}) d\boldsymbol{v} \right)^{\frac{1}{2}} < \infty \mid \Xi \in L^1_{loc}(\mathbb{R}^2) \right\}.$$

Now we state the optimization problem.

Problem 4. *Given $\mathcal{P}_0, \nu, \varepsilon, T > 0$ in system (5.4) and the objective functional (5.8), find*

$$\breve{\eta}_\varepsilon^{\nu,T} = \arg\max_{\eta\in\mathcal{S}} \mathcal{J}_\varepsilon(\eta), \quad \textit{where}$$

$$\mathcal{S} := \left\{ \eta \in H^4(\Omega)\cap L^2_M(\Theta) \;:\; \int_\Omega \left| \boldsymbol{\nabla} \int_\Theta S\eta(\mathbf{x},\boldsymbol{v})\, M(\boldsymbol{v})\, d\boldsymbol{v} \right|^2 d\mathbf{x} = \mathcal{P}_0\,, \right.$$
$$\left. \int_\Omega \int_\Theta \eta(\mathbf{x},\boldsymbol{v})\, M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x} = m_0 \right\}. \tag{5.9}$$

For Problem 4, we will use previously obtained optimal initial conditions for 2D Navier-Stokes investigated in Chapter 2. To obtain a velocity field from vorticity one can invert the Biot-Savart law, however determining η provided the velocity $\boldsymbol{u}_0$ does not

have a unique solution g_ε as in (5.7). As a result, a new function must be constructed to satisfy the equations around the equilibrium $f_\varepsilon^{\text{eq}}$ (Krüger et al. 2017). As previously done, the goal is to determine a family of locally maximizing solutions of Problem 4 for a range of ε and T, while considering a fixed initial palinstrophy $\mathcal{P}_0$. Unlike obtaining η from $\boldsymbol{u}_0$, using (5.6) and (5.7) we can obtain a unique velocity field $\boldsymbol{u}_0$ from η, and thus a unique vorticity field ω_0, which we can use to compare with the optimal initial conditions found in Chapter 2. The requirement that $\eta \in H^4(\Omega) \cap L^2_M(\Theta)$ ensures that (5.4) admits unique local solutions $g_\varepsilon \in L^\infty(0, T\,:\, H^4(\Omega) \cap L^2_M(\Theta))$ (Zhongyang 2022). Now we will suitably adapt the approach used in Chapter 2, to determine local maximizers of (5.8).

5.1.4 Gradient-Based Solution Approach

As we did in Chapter 2, we formulate our solution approach in the continuous setting for the objective functional $\mathcal{J}_\varepsilon(\eta)$, using an "optimize-then-discretize" approach (Gunzburger 2003). Thus, for fixed values of $\mathcal{P}_0$, ν, ε, and T, a local maximizer $\breve{\eta}_\varepsilon^{\nu,T}$ of Problem 4 can be determined as $\breve{\eta}_\varepsilon^{\nu,T} = \lim_{n\to\infty} \eta^{(n)}$ using an iterative gradient flow procedure projected on $\mathcal{S}$ (5.9) defined in Problem 4

$$\begin{aligned} \eta^{(n+1)} &= \mathbb{P}_{\mathcal{S}} \left(\eta^{(n)} + \tau_n \nabla \mathcal{J}_\varepsilon \left(\eta^{(n)} \right) \right), \\ \eta^{(1)} &= \eta_0, \end{aligned} \tag{5.10}$$

where $\eta^{(n)}$ is the maximizer obtained at the n-th iteration, η_0 is the initial guess (from Chapter 2), and τ_n is the length of the step in the direction of the gradient $\nabla \mathcal{J}_\varepsilon(\eta^{(n)})$.

To represent the infinite-dimensional sensitivity of the objective functional (5.8) to perturbations of the initial condition η, we require an expression for the gradient $\nabla \mathcal{J}_\varepsilon(\eta)$ in (5.10). We will determine the gradient by using the Riesz representation theorem (Luenberger 1969) and the fact that Gâteaux (directional) differential $\mathcal{J}_\varepsilon{}'(\eta;\cdot) : H^2(\Omega,\Theta) \to \mathbb{R}$, defined as $\mathcal{J}_\varepsilon{}'(\eta;\eta') := \lim_{\epsilon\to 0} \epsilon^{-1} \left[\mathcal{J}_\varepsilon(\eta + \epsilon\eta') - \mathcal{J}_\varepsilon(\eta)\right]$ for

some perturbation $\eta' \in H^2(\Omega \times \Theta)$, is a bounded linear functional on $H^2(\Omega, \Theta)$. The Gâteaux differential of $\mathcal{J}_\varepsilon(\eta)$ can be computed as

$$
\begin{aligned}
&\mathcal{J}_\varepsilon{}'(\eta;\eta') \\
&= \frac{2\nu}{T}\int_0^T\int_\Omega \left(\boldsymbol{\nabla}\int_\Theta \grave{S} g_\varepsilon(t,\mathbf{x},\grave{\boldsymbol{v}};\eta)\, M(\grave{\boldsymbol{v}})\, d\grave{\boldsymbol{v}}\right)\cdot\left(\boldsymbol{\nabla}\int_\Theta S g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')\, M(\boldsymbol{v})\, d\boldsymbol{v}\right) d\mathbf{x}\, dt, \\
&= -\frac{2\nu}{T}\int_0^T\int_\Omega \left(\boldsymbol{\Delta}\int_\Theta \grave{S} g_\varepsilon(t,\mathbf{x},\grave{\boldsymbol{v}};\eta)\, M(\grave{\boldsymbol{v}})\, d\grave{\boldsymbol{v}}\right)\left(\int_\Theta S g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')\, M(\boldsymbol{v})\, d\boldsymbol{v}\right) d\mathbf{x}\, dt, \\
&= -\frac{2\nu}{T}\int_0^T\int_\Omega\int_\Theta \left(\boldsymbol{\Delta}\int_\Theta \grave{S} g_\varepsilon(t,\mathbf{x},\grave{\boldsymbol{v}};\eta)\, M(\grave{\boldsymbol{v}})\, d\grave{\boldsymbol{v}}\right) S g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')\, M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x}\, dt, \\
&= \int_0^T\int_\Omega\int_\Theta \left(\frac{2\nu}{T} S\boldsymbol{\Delta}\int_\Theta \grave{S} g_\varepsilon(t,\mathbf{x},\grave{\boldsymbol{v}};\eta)\, M(\grave{\boldsymbol{v}})\, d\grave{\boldsymbol{v}}\right)\left(g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')\right) M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x}\, dt,
\end{aligned}
\tag{5.11}
$$

where the integration by parts with respect to the $\mathbf{x}$-coordinate was performed twice, $\grave{S} := -\grave{\boldsymbol{v}}\cdot\boldsymbol{\nabla}^\perp$, and $g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')$ satisfies the linearized Boltzmann equation. That is, perturbing the initial condition $\eta \leftarrow \eta + \epsilon\eta'$, for $0 < \epsilon \ll 1$, substituting this into (5.4), and collecting terms of order $\mathcal{O}(\epsilon)$ we obtain

$$
\varepsilon\partial_t g'_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} g'_\varepsilon = \frac{1}{\nu\varepsilon}\Big(P(g'_\varepsilon) - g'_\varepsilon\Big), \tag{5.12a}
$$

$$
g'_\varepsilon(t=0) = \eta', \tag{5.12b}
$$

which is also subject to periodic boundary conditions in $\mathbf{x}$ and where

$$
\begin{aligned}
P(g'_\varepsilon) &:= \rho'_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{u}'_\varepsilon + 2\varepsilon\big[\,[\boldsymbol{v}\cdot\boldsymbol{u}_\varepsilon]\;[\boldsymbol{v}\cdot\boldsymbol{u}'_\varepsilon] - \boldsymbol{u}_\varepsilon\cdot\boldsymbol{u}'_\varepsilon\big], \\
\rho'_\varepsilon &:= \langle g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')\rangle_M, \\
\boldsymbol{u}'_\varepsilon &:= \langle \boldsymbol{v} g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')\rangle_M.
\end{aligned}
$$

Viewing (5.11) as a bounded linear functional of η' and using the fact that it is a directional derivative, the Riesz representation theorem allows us to express (5.11) as

the inner product

$$\mathcal{J}_\varepsilon{}'(\eta;\eta') = \left\langle \nabla \mathcal{J}_\varepsilon(\eta), \eta' \right\rangle_{H^2(\Omega\times\Theta)} = \left\langle \nabla^{L^2} \mathcal{J}_\varepsilon(\eta), \eta' \right\rangle_{L^2(\Omega,\Theta)}, \tag{5.13}$$

where the Riesz representers $\nabla \mathcal{J}_\varepsilon(\eta)$ and $\nabla^{L^2} \mathcal{J}_\varepsilon(\eta)$ are the gradients of the objective functional computed with respect to the H^2 and L^2 topology, respectively. We will derive the L^2 gradient here, however in practice, the H^2 gradient will be used in computations.

The expression in (5.11) is inconsistent with the Riesz form (5.13) as the perturbation of the initial condition η' is hidden in the system determining the evolution of $g'_\varepsilon(t,\mathbf{x},\boldsymbol{v};\eta,\eta')$ and must be uncovered. To perform this, we introduce the adjoint field $g^*_\varepsilon : [0,T]\times\Omega\times\Theta \to \mathbb{R}$, which satisfies the judiciously defined adjoint system (subject to periodic boundary conditions in $\mathbf{x}$)

$$\varepsilon\partial_t g^*_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} g^*_\varepsilon + \frac{1}{\nu\varepsilon}Q(g^*_\varepsilon) - \frac{1}{\nu\varepsilon}g^*_\varepsilon = -\frac{2\nu}{T}\, S\left(\boldsymbol{\Delta}\int_\Theta \grave{S} g_\varepsilon(t,\mathbf{x},\grave{\boldsymbol{v}};\eta)\, M(\grave{\boldsymbol{v}})\, d\grave{\boldsymbol{v}}\right), \tag{5.14a}$$

$$g^*_\varepsilon(T,x,v) = 0, \tag{5.14b}$$

where

$$Q(g^*_\varepsilon) := \langle g^*_\varepsilon\rangle_M + \boldsymbol{v}\cdot\langle \boldsymbol{v} g^*_\varepsilon\rangle_M + 2\varepsilon\left[(\boldsymbol{u}_\varepsilon)^T\cdot\langle(\boldsymbol{v}\otimes\boldsymbol{v})g^*_\varepsilon\rangle_M\cdot\boldsymbol{v} - (\boldsymbol{v}\cdot\boldsymbol{u}_\varepsilon)\langle g^*_\varepsilon\rangle_M\right].$$

Similar to (5.4), system (5.14) admits local solutions $g^*_\varepsilon \in L^\infty(0,T : H^4(\Omega)\cap L^2_M(\Theta))$ (Zhongyang 2022). Now integrating (5.12a) against the adjoint field g^*_ε over the time-space-velocity domain $[0,T]\times\Omega\times\Theta$ with respect to the measure $M\,d\boldsymbol{v}$ and performing

integration by parts with respect to both time and space, we obtain

$$\begin{aligned}
&\int_0^T \int_\Omega \int_\Theta \left[\varepsilon \partial_t g'_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} g'_\varepsilon - \frac{1}{\nu\varepsilon} P(g'_\varepsilon) + \frac{1}{\nu\varepsilon} g'_\varepsilon\right] g^*_\varepsilon \, M \, d\boldsymbol{v}\, d\mathbf{x}\, dt \;=\; 0, \\
&\int_0^T \int_\Omega \int_\Theta \left[\varepsilon \partial_t g^*_\varepsilon + \boldsymbol{v}\cdot\boldsymbol{\nabla} g^*_\varepsilon + \frac{1}{\nu\varepsilon} Q(g^*_\varepsilon) - \frac{1}{\nu\varepsilon} g^*_\varepsilon\right] g'_\varepsilon \, M \, d\boldsymbol{v}\, d\mathbf{x}\, dt \;=\; \int_\Omega \int_\Theta g^*(0,\mathbf{x},\boldsymbol{v})\, \eta' \, M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x}, \\
&\underbrace{\int_0^T \int_\Omega \int_\Theta \left[-\frac{2\nu}{T} S\left(\boldsymbol{\Delta} \int_\Theta \grave{S} g_\varepsilon \, M(\grave{\boldsymbol{v}})\, d\grave{\boldsymbol{v}}\right)\right] g'_\varepsilon \, M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x}\, dt}_{\mathcal{J}_\varepsilon{}'(\eta;\eta')} \;=\; \int_\Omega \int_\Theta g^*_\varepsilon(0,\mathbf{x},\boldsymbol{v})\, \eta' \, M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x},
\end{aligned} \tag{5.15}$$

where spatial boundary terms vanish when performing integration by parts due to periodicity and the terminal condition from integration by parts is zero due to (5.14b). From (5.15), we obtain that $\mathcal{J}_\varepsilon{}'(\eta;\eta') = \int_\Omega \int_\Theta g^*(0,\mathbf{x},\boldsymbol{v})\, \eta' \, M(\boldsymbol{v})\, d\boldsymbol{v}\, d\mathbf{x}$, and can extract the L^2 gradient via (5.13), to obtain

$$\nabla^{L^2} \mathcal{J}_\varepsilon(\mathbf{x},\boldsymbol{v}) = g^*_\varepsilon(0,\mathbf{x},\boldsymbol{v}). \tag{5.16}$$

The creation of a numerical scheme and implementation of this optimization procedure is currently underway and results will be reported in the future.

5.2 Energy Transfer

5.2.1 Introduction

The energy cascade between the hierarchy of scales was described first by Richardson 1922, who provided a statistical description of this cascade for turbulent flows. This idea further leads to the concept of self-similar flow structures, predicted by Kolmogorov (1941c). Although this is described from a statistical point of view, it is still yet to be understood from the mathematical perspective. Of particular interest are the mechanisms and flow structures which sustain the self-similar hierarchy for the energy cascade dynamics. There has been recent work studying this problem using direct numerical

simulations (Goto et al. 2017; Yoneda et al. 2022), and in this project we hope to further explore this work by constructing optimal initial data that realizes this energy transfer over one eddy-turnover time. This will allow us to understand the flow structures of the energy cascade.

5.2.2 Burgers Equation and Optimization Approach

As a simple "toy" problem, we are interested in the energy transfer in flows described by the 1D Burgers equation on a domain $\Omega = [0, 2\pi]$ with periodic boundary conditions

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} - \nu\frac{\partial^2 u}{\partial x^2} = 0 \qquad \text{in } (0, T] \times \Omega, \tag{5.17a}$$

$$u(t = 0) = \phi \qquad \text{in } \Omega, \tag{5.17b}$$

where $\nu > 0$ is the viscosity and ϕ is the initial condition. Similar to the 3D Navier-Stokes system, (5.17) exhibits a forward energy cascade of kinetic energy where energy is transferred from large scales to small scales, which is the main focus of this project.

We are interested in studying this energy transfer over a single eddy-turnover time, T, between scales, which can be viewed as a difference between Fourier modes. To study the structures driving this energy transfer over T, we can observe the differences over Fourier modes by defining the functional

$$\begin{aligned}\mathcal{J}_\lambda(\phi) :=& \frac{1}{2}\int_0^\infty \left[|\widehat{\phi}(k)|^2 - \lambda|\widehat{u}(T, \lambda k; \phi)|^2\right]^2 dk, \\ =& \frac{1}{2}\int_0^\infty \left[|\widehat{\phi}(k)|^2 - \lambda\int_0^T \delta(t - T)\,|\widehat{u}(t, \lambda k; \phi)|^2 dt\right]^2 dk,\end{aligned} \tag{5.18}$$

where $\widehat{\phi}(k)$ is the Fourier transform of $\phi(x)$, $\delta(\cdot)$ is the Dirac delta distribution, and $\lambda \in \mathbb{N}^+$ is a parameter to define the energy transfer. As was done in Section 3.2.2, we will slightly abuse notation and consider $k \in \mathbb{R}$ for simplicity, however $k \in \mathbb{Z}$ will need to be used in actual implementation. Here, we seek to construct optimal initial conditions

that will sustain a self-similar hierarchy between scales of motion, which can be done by minimizing the functional (5.18). This leads to the following optimization problem.

Problem 5. *For fixed T and $\lambda \in \mathbb{N}^+$ for (5.18), find*

$$\breve{\phi}=\underset{\phi\in H^1(\Omega)}{\arg\min}\, \mathcal{J}_\lambda(\phi), \qquad \textit{subject to} \qquad \|\phi\|_{\dot{H}^1(\Omega)} = 1. \tag{5.19}$$

Similar to the procedure described in Section 5.1.4, we utilize the Riesz representation theorem (Berger 1977), and compute the Gâteaux (directional) differential defined as $\mathcal{J}_\lambda{}'(\phi;\phi') := \lim_{\epsilon\to 0}\epsilon^{-1}\left[\mathcal{J}_\lambda(\phi+\epsilon\phi') - \mathcal{J}_\lambda(\phi)\right]$ for some perturbation of the initial $\phi' \in H^1(\Omega)$, which is a bounded linear functional when viewed as a function of the second argument. Before computing the Gâteaux differential, we use Parseval's identity and note that for real-valued functions $\upsilon_1(x), \upsilon_2(x) \in L^2(\Omega)$,

$$\begin{aligned}
\int_0^\infty \widehat{\upsilon}_1(k)\overline{\widehat{\upsilon}}_2(k)\,dk &= \langle \widehat{\upsilon}_1(k), \widehat{\upsilon}_2(k)\rangle_{L^2(\mathbb{R})}\,, \\
&= \langle \upsilon_1(x), \upsilon_2(x)\rangle_{L^2(\Omega)}\,, \\
&= \overline{\langle \upsilon_1(x), \upsilon_2(x)\rangle}_{L^2(\Omega)}, \\
&= \overline{\langle \widehat{\upsilon}_1(k), \widehat{\upsilon}_2(k)\rangle}_{L^2(\mathbb{R})}, \\
&= \int_0^\infty \widehat{\upsilon}_2(k)\overline{\widehat{\upsilon}}_1(k)\,dk.
\end{aligned} \tag{5.20}$$

Since $|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2$ are real-valued for all $k \in \mathbb{R}$, we use (5.20) to compute the Gâteaux differential of $\mathcal{J}_\lambda(\phi)$ with respect to ϕ as

$$\begin{aligned}
\mathcal{J}_\lambda{}'(\phi;\phi') &= \int_0^\infty \left[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2\right]\,\left[\widehat{\phi}'(k)\overline{\widehat{\phi}}(k) + \widehat{\phi}(k)\overline{\widehat{\phi}}'(k)\right]\,dk \\
&\quad - \int_0^\infty \lambda\left[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2\right]\,\left[\widehat{u}'(T,\lambda k;\phi,\phi')\overline{\widehat{u}}(T,\lambda k;\phi)\right]\,dk \\
&\quad - \int_0^\infty \lambda\left[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2\right]\,\left[\widehat{u}(T,\lambda k;\phi)\overline{\widehat{u}}'(T,\lambda k;\phi,\phi')\right]\,dk, \\
&= 2\int_0^\infty \left[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2\right]\widehat{\phi}(k)\,\overline{\widehat{\phi}}'(k)\,dk \\
&\quad - 2\int_0^\infty \lambda\left[|\widehat{\phi}(k)|^2 - \lambda\,|\widehat{u}(T,\lambda k;\phi)|^2\right]\int_0^T \delta(t-T)\widehat{u}(t,\lambda k;\phi)\,\overline{\widehat{u}}'(t,\lambda k;\phi,\phi')dt\,dk,
\end{aligned} \tag{5.21}$$

where $u'(t,x;\phi,\phi')$ satisfies

$$\frac{\partial u'}{\partial t} + u'\frac{\partial u}{\partial x} + u\frac{\partial u'}{\partial x} - \nu\frac{\partial^2 u'}{\partial x^2} = 0 \qquad \text{in } (0,T]\times\Omega, \tag{5.22a}$$

$$u'(t=0) = \phi' \qquad \text{in } \Omega. \tag{5.22b}$$

In order to extract the gradient $\nabla_\phi \mathcal{J}_\lambda(\phi)$ from (5.21), we use the use the Riesz representation theorem (Berger 1977) to note

$$\mathcal{J}_\lambda{}'(\phi;\phi') = \left\langle \nabla_\phi^{H^1}\mathcal{J}_\lambda, \phi'\right\rangle_{H^1(\Omega)} = \left\langle \nabla_\phi^{L^2}\mathcal{J}_\lambda, \phi'\right\rangle_{L^2(\Omega)}. \tag{5.23}$$

The first term in (5.21) is in the appropriate Riesz form (5.23), however the perturbation ϕ' must be uncovered in the second term. To do this, we introduce the adjoint variable u^*, which satisfies

$$-\frac{\partial u^*}{\partial t} + u^*\frac{\partial u}{\partial x} - \frac{\partial (u\,u^*)}{\partial x} - \nu\frac{\partial^2 u^*}{\partial x^2} = 0, \tag{5.24a}$$

$$u^*(t=T,x) = W_0(x), \tag{5.24b}$$

where $W_0(x)$ is the terminal condition defined in Fourier space as

$$\widehat{W}_0(k) = -2\lambda \left[|\widehat{\phi}(k)|^2 - \lambda \, |\widehat{u}(T, \lambda k; \phi)|^2 \right] \widehat{u}(T, \lambda k; \phi). \tag{5.25}$$

We integrate (5.22) against the adjoint variable u^* over the time-space domain $[0, T] \times \Omega$ and perform integration by parts with respect to both time and space, to obtain

$$\begin{aligned}
&\int_0^T \int_\Omega \left[\frac{\partial u'}{\partial t} + u' \frac{\partial u}{\partial x} + u \frac{\partial u'}{\partial x} - \nu \frac{\partial^2 u'}{\partial x^2} \right] u^* \, dx \, dt = 0, \\
&\int_\Omega u^*(T, x) u'(T, x) + \int_0^T \left[-\frac{\partial u^*}{\partial t} + u^* \frac{\partial u}{\partial x} - \frac{\partial (u \, u^*)}{\partial x} - \nu \frac{\partial^2 u^*}{\partial x^2} \right] \phi' \, dt \, dx = \int_\Omega u^*(0, x) \, \phi' \, dx, \\
&\int_\Omega u^*(T, x) u'(T, x) \, dx = \int_\Omega u^*(0, x) \, \phi' \, dx,
\end{aligned} \tag{5.26}$$

where the adjoint equation is equal to zero from (5.24a) and spatial boundary terms resulting from integration by parts vanish due to periodicity. The expression given in (5.26) is precisely the second term in (5.21). Thus, invoking the Riesz representation theorem, we can deduce the L^2-gradient of (5.18), $\nabla_\phi^{L^2} \mathcal{J}_\lambda$, defined in Fourier space to be

$$\widehat{\nabla_\phi^{L^2} \mathcal{J}_\lambda}(k) = 2 \, \left[|\widehat{\phi}(k)|^2 - \lambda \, |\widehat{u}(T, \lambda k; \phi)|^2 \right] \widehat{\phi}(k) + \widehat{u}^*(0, k). \tag{5.27}$$

Results obtained by solving Problem 5 as described above will be reported in the near future. It is expected that the optimal initial conditions constructed will give us greater insight to the self-similar structure that leads to Kolmogorov–Richardson energy cascade. As this is simply a toy problem, the ultimate goal is to extend this work to the 3D Navier-Stokes system, where the initial conditions are expected to be antiparallel tubular vortices.

Chapter 6

Discussion & Conclusions

In this work, we investigate some fundamental theoretical and computational problems that arise in 2D turbulence. Formulated in the continuous setting as optimization problems, these problems are solved using an *optimize-then-discretize* approach, which was applied numerically via computational methods. Using pseudospectral methods to efficiently compute solutions to PDEs, solutions of the optimization problems were obtained via iterative gradient descent/ascent methods with gradients conveniently computed by solving adjoint systems. Furthermore, we introduce the technique of utilizing adjoint calculus to impose complicated nonhomogeneous constraints on the control variable of an optimization problem. In this chapter, we briefly discuss research contribution made in different chapters.

In Chapter 2, we probe the sharpness of the upper bound bounds on enstrophy dissipation in 2D Navier-Stokes flows in the limit of zero viscosity. Initial conditions that maximize enstrophy dissipation locally for a given time window and starting with a fixed palinstrophy are determined numerically as solutions to an optimization problem via an adjoint-based gradient ascent method. Multiple branches were found, which were noted to possess distinct features to increase palinstrophy over the prescribed time window. Ultimately, the sharpness of an estimate due to Ciampa et al. (2021) was shown to be

in agreement with the behavior of the extremal flows found in this chapter. Estimates of this bound have historically been approached using mathematical analysis, however here we have demonstrated the sharpness of this bound, through the use of optimization methods.

Chapter 3 focuses on the problem of fundamental performance limitations of closure models of turbulence flows. We construct optimal eddy viscosities as a closure model, using a generalized Leith model for 2D LES. Using minimum assumptions on eddy viscosities, LES solutions were constructed to best match filtered DNS data. This resulted in a PDE-optimization problem with a nonstandard structure, which was solved via adapting the adjoint-based gradient method. The main findings of this work indicate that the performance of LES improves as the regularization in the eddy viscosities is reduced. In turn, the eddy viscosities exhibit increasingly rapid oscillations and we conclude that the problem of finding optimal eddy viscosities is ill-posed as the optimal eddy viscosities do not converge to a well-defined limit. To achieve useful eddy viscosities in practice, which would not reach the theoretical performance limitations, stronger regularization is required, as demonstrated in this chapter as well.

Extending the work of Chapter 3, Chapter 4 investigates determining optimal eddy viscosities subject to an additional nonlinear and nonhomogeneous constraint. Modifying our objective functional from Chapter 3 to consider the time-averaged, least-squared error between palinstrophies in a filtered DNS and LES solutions, we add the constraint that the time-averaged enstrophy in each solution also be equal. This constraint defines a complex manifold in the space of eddy viscosities serving as the control variable, which is extremely difficult to enforce in a computationally efficient manner. To address this problem, we developed a method for approximately enforcing general nonlinear constraints defining complex constraint manifolds. Using adjoint calculus, we were able to define a subspace tangent to the constraint manifold, on which we projected the gradient

flow to approximately enforce the constraint. Although the constraint is only approximately enforced, the results demonstrated that this approach did result in a better match between the enstrophy in the target and optimized flows, when compared to the case without any constraint. As expected, in the constrained problem a worse decrease in the objective functional was obtained. However a comparison of the two optimal eddy viscosities, when the enstrophy constraint was and was not imposed, provided insight into qualitative features of the optimal eddy viscosities. The main contribution of this work is a general framework of using adjoint calculus in a "nonstandard" manner, namely, to approximately enforce general nonlinear constraints, in a computationally efficient method.

Two open problems are explored in Chapter 5, where they are motivated and we also present pathways for solutions for these problems. First, we outlined a research problem which extends the work conducted in Chapter 2, from the perspective of kinetic theory using the Boltzmann equation. The problem was formulated to study enstrophy dissipation in the context of the Boltzmann equation, which will provide a complementary viewpoint onto to this problem. In addition, we outline a problem to study Kolmogorov-Richardson energy cascade, by constructing self-similar initial conditions to understand the energy transfer between scales in 1D Burgers flows on a torus. Both these problems are likely to produce intriguing results which will be reported upon in the future.

This thesis provides new perspective and insight to the 2D Navier-Stokes system, using optimization methods. The study of fluid mechanics has been traditionally dominated by determining solutions of a governing equations, to understand flow properties. In this work, we showcase the use of optimization methods and computational mathematics to construct solutions that possess specific properties. This provides a new avenue for studying problems in fluid mechanics, which assist in solving some of the open problems that exist in the field. Moreover, this work acts as a stepping-stone, providing a

framework to approach these problems in the context of 3D Navier-Stokes flow.

Appendix A

Numerical Methods for the 2D Navier-Stokes System

We numerically solve the 2D Navier-Stokes system to obtain solutions for the optimization problems considered in Chapters 2, 3, and 4. Here we describe the numerical methods employed to solve the 2D Navier-Stokes system (1.3), which are suitable for both the unforced setting, cf. (2.1), and with the forced setting, cf. (3.1). These systems are solved forward in time, whereas the corresponding adjoint systems, (2.19) and (3.22), must be solved backwards in time. However, via the convenient change of the time variable, $t' = T - t$ in (2.19) and (3.22), these numerical methods are also used to solve the systems backwards in time. We implement a pseudospectral method for spatial discretization of the equations, where differentiation and other linear operators are computed in Fourier space using the FFT algorithm (based on the FFTW implementation (Frigo and Johnson 2005)) to evaluate the Fourier transform. Nonlinear terms are computed in the physical space, using an appropriate dealiasing method based on the standard 3/2 rule (Hou 2009). Time integration is performed using an implicit/explicit Runge-Kutta scheme (Alimo et al. 2021) for the problems considered in Chapters 2 and 4, whereas a Crank-Nicolson combined with third-order Runge–Kutta–Wray method was

used for problems considered in Chapter 3.

A1 Spatial Discretization — Pseudospectral Method

Consider a scalar field $w : (0,T] \times \Omega \to \mathbb{R}$, and the set of wavevectors $\mathscr{K} = \{\boldsymbol{k} \in \mathbb{Z}^2 : |\boldsymbol{k}| \leq K\}$, where K is the maximum resolved wavenumber. Without loss of generality, let $\Omega = [0, 2\pi]^2$. Then, denoting the Fourier coefficient $\widehat{w}(t, \boldsymbol{k})$, we can define the Galerkin approximation of $w(t, \mathbf{x})$ as

$$\begin{aligned} w_K(t, \mathbf{x}) &= \sum_{\boldsymbol{k} \in \mathscr{K}} \widehat{w}(t, \boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\mathbf{x}}, \\ &= \sum_{\boldsymbol{k} \in \mathscr{K}} \widehat{w}_{k_1,k_2}(t)\, e^{i(k_1 x_1 + k_2 x_2)}, \end{aligned}$$

where $\boldsymbol{k} = (k_1, k_2)$ and the sum is over these components such that $-K_1 \leq k_1 \leq K_1$, $-K_2 \leq k_2 \leq K_2$, with $K = \sqrt{K_1^2 + K_2^2}$. Similarly, we define the Galerkin approximation of the streamfunction $\psi(t, \mathbf{x})$ by denoting the Fourier coefficient $\widehat{\psi}(t, \boldsymbol{k})$, which we represent as

$$\psi_K(t, \mathbf{x}) = \sum_{\boldsymbol{k} \in \mathscr{K}} \widehat{\psi}(t, \boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\mathbf{x}}.$$

With this, we can approximate our 2D vorticity equation (1.3) as a finite-dimensional system

$$\frac{\partial \widehat{w}}{\partial t} + \widehat{\mathcal{N}}(\widehat{\psi}, \widehat{w})(t, \boldsymbol{k}) - \nu\, |\boldsymbol{k}|^2 \widehat{w} = \widehat{f}_\omega, \qquad \forall\ t \in (0,T],\ \forall\ \boldsymbol{k} \in \mathscr{K}, \tag{A.1a}$$

$$-|\boldsymbol{k}|^2 \widehat{\psi} = -\widehat{w}, \qquad \forall\ t \in (0,T],\ \forall\ \boldsymbol{k} \in \mathscr{K}, \tag{A.1b}$$

$$\widehat{w}(t=0) = \widehat{w}_0, \qquad \forall\ \boldsymbol{k} \in \mathscr{K}, \tag{A.1c}$$

where $\widehat{w}_0$ is the Fourier transform of the initial condition w_0, and $\widehat{\mathcal{N}}$ is Fourier transform of the nonlinear term, denoted $\mathcal{N}$. The Fourier-truncated system (A.1) can also be

written in the canonical form

$$\frac{d\boldsymbol{v}}{dt} = \mathcal{L}\boldsymbol{v} + \mathcal{N}(\boldsymbol{v}), \tag{A.2}$$

where $\boldsymbol{v}$ denotes the vector of the Fourier coefficients $\widehat{w}$, $\widehat{\psi}$ and $\mathcal{L}$ contains the linear operators on the RHS. The Fourier modes in (A.1) are coupled due to the nonlinear term $\widehat{\mathcal{N}}$, which can be represented as a convolution in Fourier space

$$\begin{aligned}\widehat{\mathcal{N}}(\widehat{\psi},\widehat{w})(t,\boldsymbol{k}) = \int_{\Omega} \Bigg[&\left(\sum_{q=-K_2}^{K_2-1} iq\widehat{\psi}_{k_1,q}(t)\right)\left(\sum_{p=-K_1}^{K_1-1} ip\widehat{w}_{p,k_2}(t)\right) - \\ &\left(\sum_{p=-K_1}^{K_1-1} ip\widehat{\psi}_{p,k_2}(t)\right)\left(\sum_{q=-K_2}^{K_2-1} iq\widehat{w}_{k_1,q}(t)\right)\Bigg] e^{-i\boldsymbol{k}\cdot\mathbf{x}}\, d\mathbf{x}, \\ = \int_{\Omega} &\left[\sum_{\substack{p=-K_1 \\ p^2+q^2=|\boldsymbol{k}|}}^{K_1-1}\sum_{q=-K_2}^{K_2-1} - pq\widehat{\psi}_{k_1,q}(t)\widehat{w}_{p,k_2}(t) + pq\widehat{\psi}_{p,k_2}(t)\widehat{w}_{k_1,q}(t)\right] e^{-i\boldsymbol{k}\cdot\mathbf{x}}\, d\mathbf{x}.\end{aligned}$$

Computing a convolution is computationally expensive, with cost proportional to $\mathcal{O}(N^2)$, and is impractical to be used in this setting. Instead, nonlinear products are computed in physical space, by taking the inverse Fourier transform of each term in $\mathcal{N}$ after performing differentiation in Fourier space, i.e.,

$$\frac{\partial \psi_K(t,\mathbf{x})}{\partial x_1} = \sum_{\boldsymbol{k}\in\mathscr{K}} (ik_1)\,\widehat{\psi}(t,\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\mathbf{x}}, \tag{A.3a}$$

$$\frac{\partial \psi_K(t,\mathbf{x})}{\partial x_2} = \sum_{\boldsymbol{k}\in\mathscr{K}} (ik_2)\,\widehat{\psi}(t,\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\mathbf{x}}, \tag{A.3b}$$

$$\frac{\partial w_K(t,\mathbf{x})}{\partial x_1} = \sum_{\boldsymbol{k}\in\mathscr{K}} (ik_1)\,\widehat{w}(t,\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\mathbf{x}}, \tag{A.3c}$$

$$\frac{\partial w_K(t,\mathbf{x})}{\partial x_2} = \sum_{\boldsymbol{k}\in\mathscr{K}} (ik_2)\,\widehat{w}(t,\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\mathbf{x}}, \tag{A.3d}$$

and then computing $\mathcal{N}$ in physical space

$$\mathcal{N}(t,\mathbf{x}) = \frac{\partial \psi_K(t,\mathbf{x})}{\partial x_2}\frac{\partial w_K(t,\mathbf{x})}{\partial x_1} - \frac{\partial \psi_K(t,\mathbf{x})}{\partial x_1}\frac{\partial w_K(t,\mathbf{x})}{\partial x_2}.$$

This nonlinear product is significantly less computationally expensive (with cost of order $\mathcal{O}(N)$), and simply requires a Fourier transform from physical space back to Fourier space. When transforming the nonlinear products back to Fourier space, dealiasing is performed using a Gaussian filtering based on the 3/2 rule (Hou 2009)

$$\rho(\boldsymbol{k}) = e^{-36\left(\frac{|\boldsymbol{k}|}{K}\right)^{36}},$$

where $K = \frac{2N}{3}$.

Fourier transforms are efficiently computed using the FFT algorithm, more specifically the FFTW implementation (Frigo and Johnson 2005), which only has a computational complexity of $\mathcal{O}(N\log(N))$. When performing these computation, these methods are combined with a parallel implementation using the Message Passing Interface (MPI).

A2 Four-Step, Third Order IMEX Method

We describe the four-step, globally third-order accurate mixed implicit/explicit (IMEX) Runge-Kutta scheme with low truncation error and low-storage requirements, introduced by Alimo et al. (2021). This time stepping method is preferred over the classic Crank-Nicolson combined with third-order low-storage Runge–Kutta–Wray (CN/RKW3) scheme that is typically used for fluids simulations (Le and Moin 1991; Bewley 2009), which is only A-stable and second-order accurate overall. Discretizing (A.2), the IMEX scheme incrementally advances the solution $\boldsymbol{u}_n$ at time t_n, using four substeps, to the solution $\boldsymbol{u}_{n+1}$ at time t_{n+1}, where we denote the step-size as Δt (hence $\Delta t = t_{n+1} - t_n$). This

scheme can be represented as follows

$$\begin{aligned}
\boldsymbol{u}^{(1)} &= \boldsymbol{u}_n + \Delta t\, \left[\alpha_1^I\, \mathcal{L}\, \boldsymbol{u}^{(1)} + \beta_1^I\, \mathcal{L}\, \boldsymbol{u}_n + \beta_1^E\, \mathcal{N}\Big(\boldsymbol{u}_n\Big)\right], \\
\boldsymbol{u}^{(2)} &= \boldsymbol{u}^{(1)} + \Delta t\, \left[\alpha_2^I\, \mathcal{L}\, \boldsymbol{u}^{(2)} + \beta_2^I\, \mathcal{L}\, \boldsymbol{u}^{(1)} + \beta_2^E\, \mathcal{N}\left(\boldsymbol{u}^{(1)}\right) + \gamma_2^E\, \mathcal{N}\Big(\boldsymbol{u}_n\Big)\right], \\
\boldsymbol{u}^{(3)} &= \boldsymbol{u}^{(2)} + \Delta t\, \left[\alpha_3^I\, \mathcal{L}\, \boldsymbol{u}^{(3)} + \beta_3^I\, \mathcal{L}\, \boldsymbol{u}^{(2)} + \beta_3^E\, \mathcal{N}\left(\boldsymbol{u}^{(2)}\right) + \gamma_3^E\, \mathcal{N}\left(\boldsymbol{u}^{(1)}\right)\right], \\
\boldsymbol{u}_{n+1} &= \boldsymbol{u}^{(3)} + \Delta t\, \left[\alpha_4^I\, \mathcal{L}\, \boldsymbol{u}_{n+1} + \beta_4^I\, \mathcal{L}\, \boldsymbol{u}^{(3)} + \beta_4^E\, \mathcal{N}\left(\boldsymbol{u}^{(3)}\right) + \gamma_4^E\, \mathcal{N}\left(\boldsymbol{u}^{(2)}\right)\right].
\end{aligned} \tag{A.4}$$

Determined by Alimo et al. (2021), the coefficients α_m^I, β_m^I, β_m^E, and γ_m^E in are given in Table A1.1, where m denotes the substep. To numerically implement the scheme given in (A.4), we rewrite the substeps as

$$\begin{aligned}
\boldsymbol{u}^{(1)} &= \Big(I - \Delta t\, \alpha_1^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_1^I\, \mathcal{L})\, \boldsymbol{u}_n + \Delta t\, \beta_1^E\, \mathcal{N}\Big(\boldsymbol{u}_n\Big) + \gamma_1^E\Big], \\
\boldsymbol{u}^{(2)} &= \Big(I - \Delta t\, \alpha_2^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_2^I\, \mathcal{L})\, \boldsymbol{u}^{(1)} + \Delta t\, \beta_2^E\, \mathcal{N}\left(\boldsymbol{u}^{(1)}\right) + \Delta t\, \gamma_2^E\, \mathcal{N}\Big(\boldsymbol{u}_n\Big)\Big], \\
\boldsymbol{u}^{(3)} &= \Big(I - \Delta t\, \alpha_3^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_3^I\, \mathcal{L})\, \boldsymbol{u}^{(2)} + \Delta t\, \beta_3^E\, \mathcal{N}\left(\boldsymbol{u}^{(2)}\right) + \Delta t\, \gamma_3^E\, \mathcal{N}\left(\boldsymbol{u}^{(1)}\right)\Big], \\
\boldsymbol{u}_{n+1} &= \Big(I - \Delta t\, \alpha_4^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_4^I\, \mathcal{L})\, \boldsymbol{u}^{(3)} + \Delta t\, \beta_4^E\, \mathcal{N}\left(\boldsymbol{u}^{(3)}\right) + \Delta t\, \gamma_4^E\, \mathcal{N}\left(\boldsymbol{u}^{(2)}\right)\Big],
\end{aligned}$$

where I is the identity matrix of suitable dimension. As noted in Alimo et al. (2021), this is a third-order method that possesses low-storage requirements and provides remarkable stability. In comparison, the classic CN/RKW3 scheme (Le and Moin 1991) is three-step, and only second-order accurate overall. Combining a Runge-Kutta and Crank-Nicolson methods, for the nonlinear (treated explicitly) and linear (treated implicitly) parts respectively, we can implement this scheme as

$$\begin{aligned}
\boldsymbol{u}^{(1)} &= \Big(I - \Delta t\, \alpha_1^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_1^I\, \mathcal{L})\, \boldsymbol{u}_n + \Delta t\, \beta_1^E\, \mathcal{N}\Big(\boldsymbol{u}_n\Big) + \gamma_1^E\Big], \\
\boldsymbol{u}^{(2)} &= \Big(I - \Delta t\, \alpha_2^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_2^I\, \mathcal{L})\, \boldsymbol{u}^{(1)} + \Delta t\, \beta_2^E\, \mathcal{N}\left(\boldsymbol{u}^{(1)}\right) + \Delta t\, \gamma_2^E\, \mathcal{N}\Big(\boldsymbol{u}_n\Big)\Big], \\
\boldsymbol{u}_{n+1} &= \Big(I - \Delta t\, \alpha_3^I\, \mathcal{L}\Big)^{-1}\, \Big[(I + \Delta t\, \beta_3^I\, \mathcal{L})\, \boldsymbol{u}^{(2)} + \Delta t\, \beta_3^E\, \mathcal{N}\left(\boldsymbol{u}^{(2)}\right) + \Delta t\, \gamma_3^E\, \mathcal{N}\left(\boldsymbol{u}^{(1)}\right)\Big].
\end{aligned} \tag{A.5}$$

Substep, m	α_m^I	β_m^I	β_m^E	γ_m^E
1	$\frac{343038331393}{1130875731271}$	$\frac{35965327958}{140127563663}$	$\frac{14}{25}$	0
2	$\frac{288176579239}{1140253497719}$	$\frac{19632212512}{2700543775099}$	$\frac{777974228744}{1346157007247}$	$\frac{-251352885992}{790610919619}$
3	$\frac{253330171251}{677500478386}$	$\frac{-173747147147}{351772688865}$	$\frac{251277807242}{1103637129625}$	$\frac{-383714262797}{1103637129625}$
4	$\frac{189462239225}{1091147436423}$	$\frac{91958533623}{727726057489}$	$\frac{113091689455}{220187950967}$	$\frac{-403360439203}{1888264787188}$

TABLE A1.1: Coefficients defining the IMEX scheme (A.4), from Alimo et al. (2021).

Substep, m	α_m^I	β_m^I	β_m^E	γ_m^E
1	$\frac{4}{15}$	$\frac{4}{15}$	$\frac{8}{15}$	0
2	$\frac{1}{15}$	$\frac{1}{15}$	$\frac{5}{12}$	$-\frac{17}{60}$
3	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{3}{4}$	$-\frac{5}{12}$

TABLE A1.2: Coefficients defining the CN/RKW3 scheme (A.5), from Le and Moin (1991).

Using the spatial discretization described in the previous section, Appendix A1, we compare the four-step IMEX and the three-step methods, given in (A.4) and (A.5), respectively. Solving (1.3) as a test problem, we use $N = N_x = N_y = 16$ spatial discretization points in the x and y directions with a periodic domain $\Omega := \mathbb{T}^2 = [0, 1]^2$ and a terminal time value $T = 100$. For simplicity of validating the numerical methods, we use remove forcing from the system, hence $f_\omega = 0$ in (1.3), and let $\nu = 2 \times 10^{-4}$. In order to test the methods, use an initial condition w_0 representing the Taylor-Green vortex

$$w_0(x, y) = 2 \sin(ax) \sin(by), \tag{A.6}$$

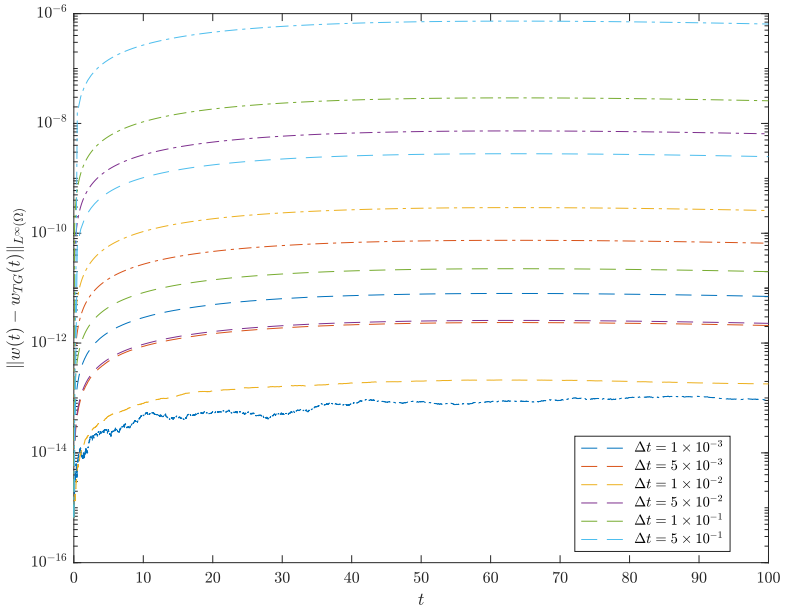
which produces an analytic solution

$$w_{TG}(t,x,y) = 2e^{-2(a^2+b^2)\nu t}\sin(ax)\sin(ay), \tag{A.7}$$

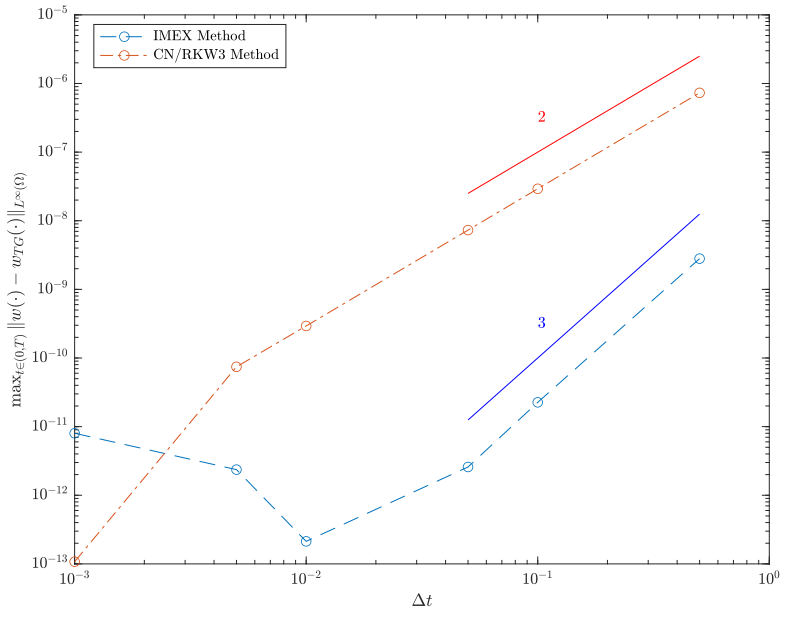
where we set $a = b = 2\pi$.

To demonstrate the accuracy of the numerical methods and the convergence to (A.7) as the time discretization is refined, we show the $L^\infty(\Omega)$ norm of the error between the numerical solution and the exact solution (A.7) in Figure A1.1. To understand the overall global accuracy of the time stepping methods, we also show the maximum error over the entire time window $[0,T]$, as a function the the time step Δt, for each of the methods. The numerical solutions converge to the analytic solution (A.7) as Δt is refined, down to machine precision. The four-step method converges to the analytic solution more rapidly, in which we eventually see round-off errors begin to arise.

A key advantage the IMEX method (A.4) provides over the CN/RKW3 method (A.5) is its remarkable stability. This allows coarser time steps Δt to be used for more efficient computational time while maintaining good accuracy, which was a crucial component for solving Problem 1. To test this, we considered (1.3) using the initial condition constructed in Jeong and Yoneda (2021), $f_\omega = 0$, $T = 1$, $N = 4096$, and $\nu = 5 \times 10^{-7}$, and noted the differences in stability and timing between (A.4) and (A.5). Using the a time step $\Delta t = 1 \times 10^{-4}$ both methods remained stable, however the CN/RKW3 method (A.5) became unstable when using $\Delta t = 1 \times 10^{-3}$, whereas the IMEX method (A.4) was stable. As a result, using the IMEX method (A.4) with a coarser time step was over 10 times computationally faster than the CN/RKW3 method (A.5) with a time step $\Delta t = 1 \times 10^{-4}$.



(a)



(b)

FIGURE A1.1: (a) Error as a function of time and (b) the maximum error over all time $[0, T]$ in the solution of system (1.3) with initial condition (A.6), comparing (dashed line) the IMEX method (A.4) and (dash-dotted line) the CN/RKW3 method (A.5). In (a), we show the step sizes Δt: (blue) 1×10^{-3}, (red) 5×10^{-3}, (yellow) 1×10^{-2}, (purple) 5×10^{-2}, (green) 1×10^{-1}, and (cyan) 5×10^{-1}.

Appendix B

Validation of Gradients

A key element of the discrete gradient flow algorithms employed in this work are gradients of the objective functionals. Here we demonstrate the validity and accuracy of the gradients of the objective functionals considered in the optimization problems studied in Chapters 2, 3, and 4. As Sobolev H^s gradients can be viewed as low-pass filtering of the L^2 gradient (Protas et al. 2004), we only show verification tests for L^2 gradients here. In order to validate the form and discretization of the gradients, which involves solution of the adjoint system, we determine the accuracy of the cost functional gradient by comparing the Gâteaux differential by $\Phi'(\Gamma;\Gamma') := \lim_{\epsilon\to 0}\epsilon^{-1}\left[\Phi(\Gamma+\epsilon\Gamma')-\Phi(\Gamma)\right]$ approximated using a forward finite-difference formula and the corresponding Riesz form $\langle\nabla^{L^2}\Phi(\Gamma),\Gamma'\rangle_{L^2(\Omega)}$, where Φ and Γ are placeholders for the functionals and the control variables, respectively, appearing in different optimization problems. Thus, we compute the ratio

$$\kappa(\epsilon) := \frac{\epsilon^{-1}\left[\Phi(\Gamma+\epsilon\Gamma')-\Phi(\Gamma)\right]}{\left\langle\nabla^{L^2}\Phi,\Gamma'\right\rangle_{L^2(\Omega)}}, \tag{B.1}$$

where Γ is the reference element where the derivative is computed, Γ' is an arbitrary perturbation defining the direction in which the directional derivative is computed, and $\epsilon > 0$ is the magnitude of the perturbation, which serves as a parameter. For an exact gradient $\nabla^{L^2}\Phi$, we expect $\kappa(\epsilon) = 1$ since the expressions in the numerator and the

denominator in (B.1) are equal. However, evaluation of Φ and $\nabla^{L^2}\Phi$ involves numerical solutions of the governing system and the corresponding adjoint system, such that we can only have $\kappa(\epsilon) \approx 1$. In addition, there are expected errors to arise in (B.1) for both large and small values of ϵ. Large ϵ values provide a poor approximation due to truncation error in the finite-difference approximation of the Gâteaux differential, whereas small values of ϵ give rise to subtractive cancellation errors due to numerical round-off. Hence, for intermediate values of ϵ we expect the value of $\kappa(\epsilon)$ be close to unity and this should improve as the numerical parameters N and Δt used in the solution of the PDE systems are refined. The numerical methods described in Appendix A utilize a spectrally accurate discretization in space, thus the more important numerical parameter in these validations is Δt.

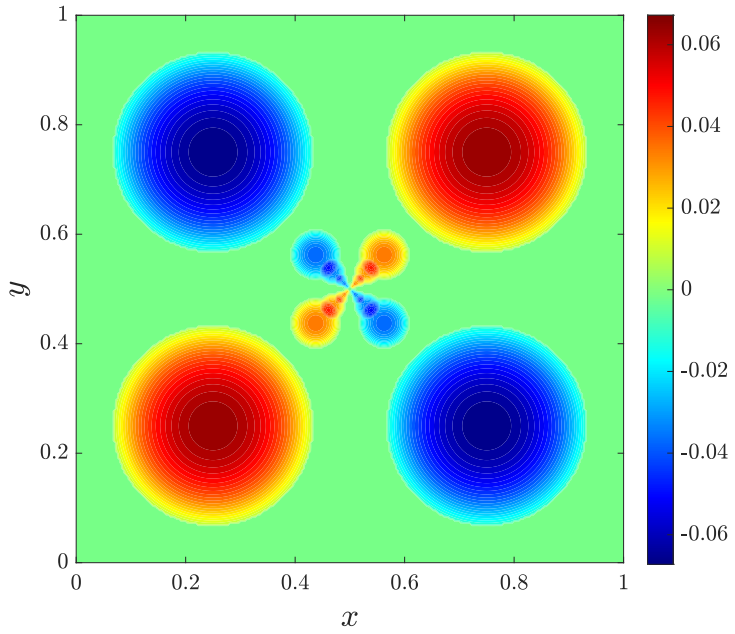
A1 Maximum Enstrophy Dissipation

We validate the gradient given in (2.21) by using the Riesz form (2.17), meaning that in (B.1) we set $\Phi = \chi_\nu$, $\Gamma = \phi \in H^1(\Omega)$, $\Gamma' = \phi' \in H^1(\Omega)$, and compute the ratio

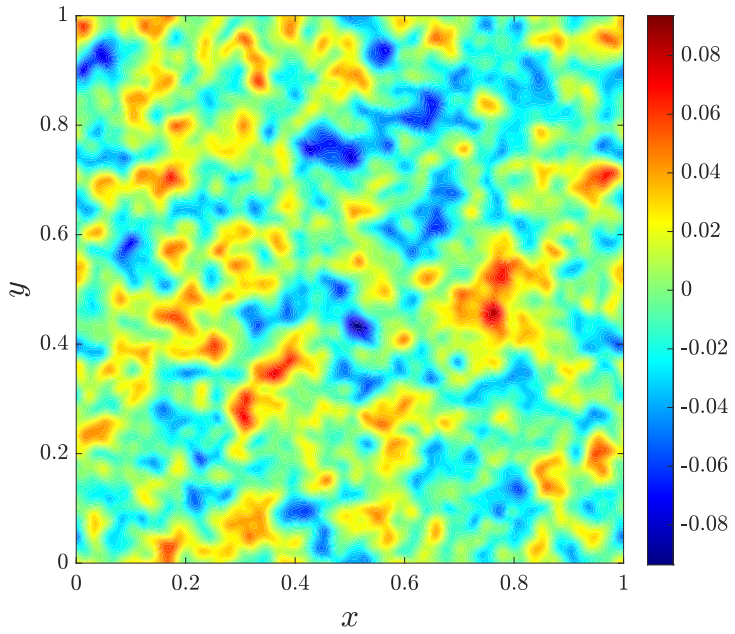
$$\kappa_1(\epsilon) := \frac{\epsilon^{-1}\left[\chi_\nu(\phi + \epsilon\phi') - \chi_\nu(\phi)\right]}{\left\langle \nabla^{L^2}\chi_\nu, \phi' \right\rangle_{L^2(\Omega)}}. \tag{B.2}$$

Since this Gâteaux differential is defined for arbitrary perturbations $\phi' \in H^1(\Omega)$, we ensure that the directional derivative is approximated accurately for multiple perturbations. We set $\phi(x)$ as the initial condition constructed in Jeong and Yoneda (2021), which is shown in Figure A2.1a, and test 2 perturbation functions. For the first perturbation we let $\phi'(\mathbf{x}) = \phi(\mathbf{x})$, whereas the second perturbation we use a vorticity field, with randomly determined Fourier modes; this random vorticity field is shown Figure A2.1b, in physical space. Computing (B.2) for $\epsilon \in [10^{-15}, 10^{-1}]$, with the parameters $\nu = 2.2361 \times 10^{-8}$, $N = 4096$, and $T = 0.0894$, we show $\kappa_1(\epsilon)$ in Figure A2.2a and $|\kappa_1(\epsilon) - 1|$ in Figure A2.2b, which demonstrates how close κ_1 is to unity. To demonstrate

the convergence to unity as the numerical parameters are refined, we consider two step sizes: $\Delta t = 1.1180 \times 10^{-5}$ and $\Delta t = 4.4721 \times 10^{-6}$. As shown in Figure A2.2b, as the time step size Δt is refined, the values of κ_1 tend closer to unity, which demonstrates that the gradients $\nabla^{L^2}\chi_\nu$ become more accurate.

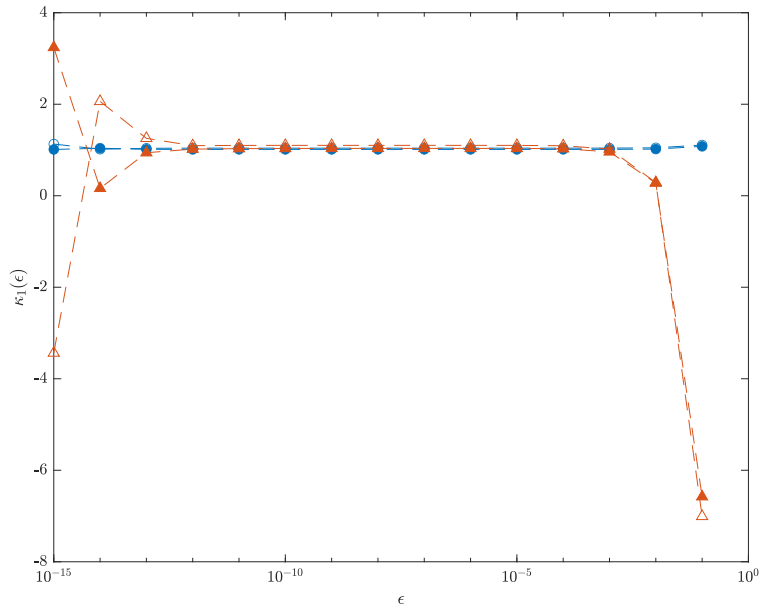


(a)

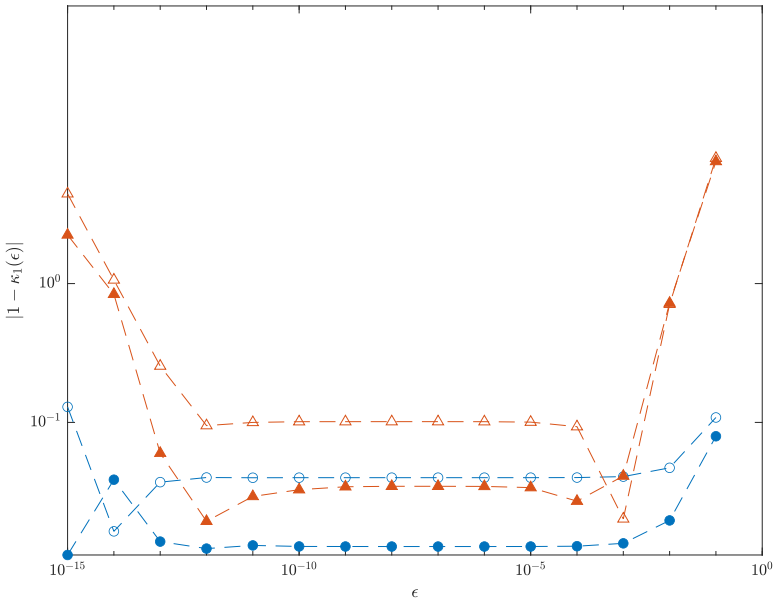


(b)

Figure A2.1: (a) The initial condition, which also acts as the first perturbation, used in (B.2) and (b) the second (random) perturbation, both shown in physical space.



(a)



(b)

FIGURE A2.2: Dependence of (a) $\kappa_1(\epsilon)$ and (b) $|1-\kappa_1(\epsilon)|$ on ϵ, cf. (B.2), for two different perturbations ϕ': perturbation 1 (blue circles) and perturbation 2 (red triangles), using two step sizes: $\Delta t = 1.1180 \times 10^{-5}$ (empty symbols) and $\Delta t = 4.4721 \times 10^{-6}$ (filled symbols).

A2 Optimal Eddy Viscosity

We validate the gradient given in (3.25) by letting $\Phi = \mathcal{J}_1$, $\Gamma = \varphi$, and $\Gamma' = \varphi'$ in (B.1) and computing the ratio

$$\kappa_2(\epsilon) := \frac{\epsilon^{-1}\left[\mathcal{J}_1(\varphi + \epsilon\varphi') - \mathcal{J}_1(\varphi)\right]}{\left\langle \nabla_{\varphi}^{L^2}\mathcal{J}_1, \varphi' \right\rangle_{L^2([0,1])}}, \tag{B.3}$$

for different values of ϵ. Although in the minimization problem (3.14) we use the function φ as our optimal control parameter, here we will present the results in terms of ν and ν', which are related to φ and φ', via (3.7). As the reference eddy viscosity, we use a standard Leith-type model

$$\nu_0(s) = (C_l\, k_c)^3 \sqrt{s+\gamma}, \tag{B.4}$$

where $C_l = 6.7 \times 10^{-4}$ is the Leith constant and $\gamma = 100$ is a regularization term which ensures that our reference eddy viscosity is bounded away from zero. For the perturbations ν' in (B.3), we test 4 different functions,

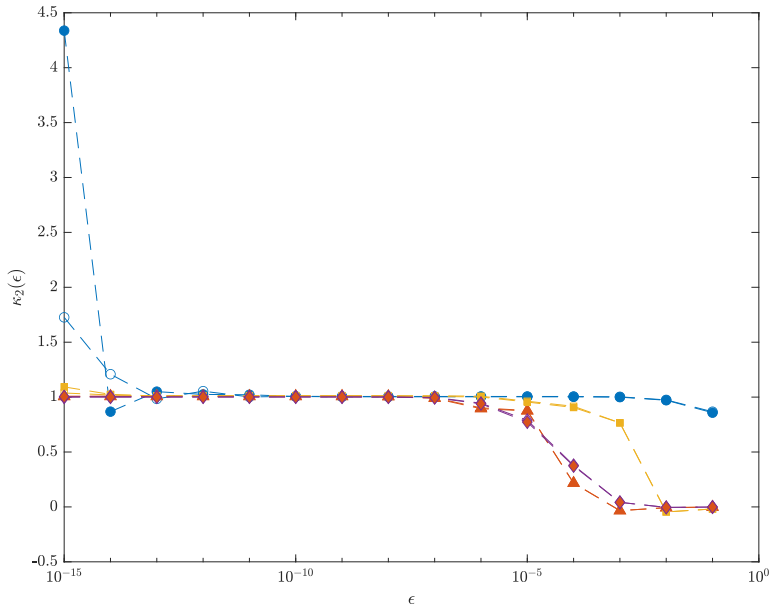
$$\nu'(s) = \nu_0, \tag{B.5a}$$

$$\nu'(s) = \left(4\cos\left(\frac{s}{750000}\right)\right)^2, \tag{B.5b}$$

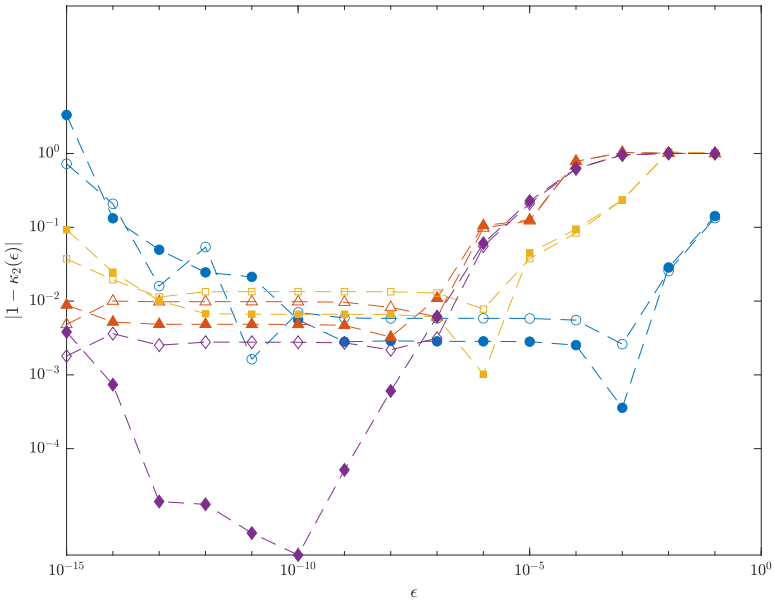
$$\nu'(s) = \left(\exp\left(\frac{-s}{300000}\right)\right)^2, \tag{B.5c}$$

$$\nu'(s) = \left(\frac{s}{100000}\right)^2. \tag{B.5d}$$

In Figure A2.3 we show the behaviour of (B.3) for $\epsilon \in [10^{-15}, 10^{-1}]$. We note as the time discretization Δt is refined, the values of (B.3) for all perturbations become closer to 1.



(a)



(b)

FIGURE A2.3: Dependence of (a) $\kappa_2(\epsilon)$ and (b) $|1-\kappa_2(\epsilon)|$ on ϵ, cf. (B.3), for four different perturbations ν' given in (B.5a) (blue circles), (B.5b) (red triangles), (B.5c) (yellow squares), and, (B.5d) (purple diamonds). Results for two different time discretizations Δt are shown: (empty symbols) $\Delta t = 1 \times 10^{-4}$ and (filled symbols) $\Delta t = 0.5 \times 10^{-4}$.

A3 Enforcing Constraints via Adjoint Analysis

Here we validate the standard gradient in the space $\mathcal{S}$, for the functional provided in (4.2). Thus, we check the ratio given in (B.1), with $\Phi = \mathcal{J}$, $\Gamma = \upsilon$, and $\Gamma' = \upsilon'$, hence,

$$\kappa_3(\epsilon) := \frac{\epsilon^{-1}\left[\mathcal{J}(\upsilon + \epsilon\upsilon') - \mathcal{J}(\upsilon)\right]}{\left\langle \nabla^{L^2}_{\upsilon}\mathcal{J}, \upsilon' \right\rangle_{L^2([0,1])}}, \tag{B.6}$$

for multiple values of ϵ, with 4 perturbations defined in terms of $\nu'(s)$,

$$\nu'(s) = \nu_0, \tag{B.7a}$$

$$\nu'(s) = \left(4\,\cos\left(\frac{s}{7.5}\right)\right)^2, \tag{B.7b}$$

$$\nu'(s) = \left(\exp\left(\frac{-s}{30}\right)\right)^2, \tag{B.7c}$$

$$\nu'(s) = \left(\frac{s}{150}\right)^2, \tag{B.7d}$$

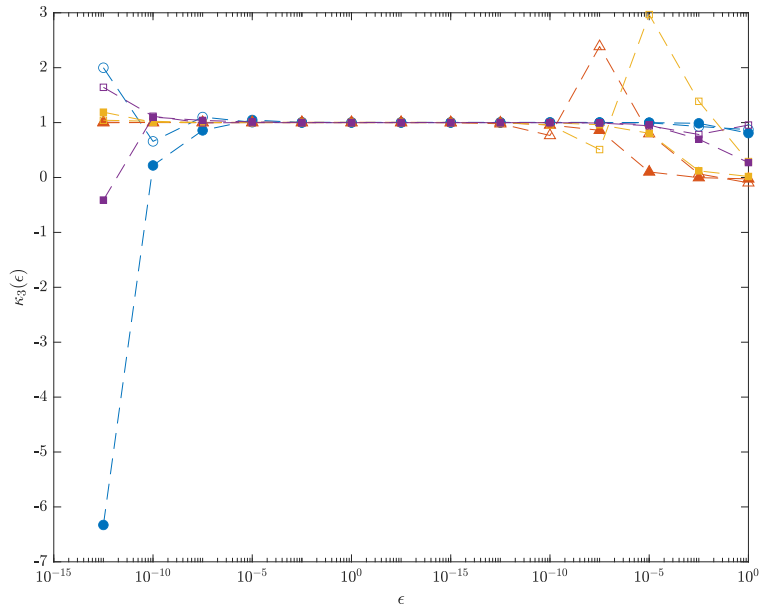
where ν_0 is given in (B.4) with Leith constant $C_l = 1.215 \times 10^{-2}$ and $\gamma = 100$. In Figure A2.4, we demonstrate that the value of (B.3) does indeed yield values close to unity, and they approach unity as the time discretization Δt is refined.

Furthermore, we must also validate computation of the "normal vector" $\boldsymbol{N}_\upsilon$, cf. Figure 4.1, by analyzing the ratio of the Gâteaux differential $\left[\widetilde{\mathcal{E}}'(\cdot;\upsilon,\upsilon')\right]_T := \lim_{\epsilon\to 0}\epsilon^{-1}\left[\widetilde{\mathcal{E}}(\cdot;\upsilon+\epsilon\upsilon') - \widetilde{\mathcal{E}}(\cdot;\upsilon)\right]_T$ approximated using a forward finite-difference formula and the corresponding Riesz form (4.13),

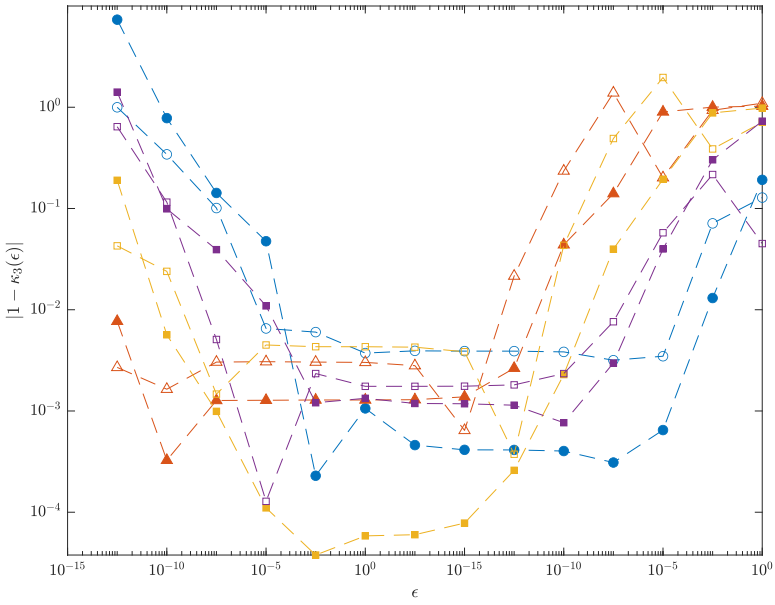
$$\kappa_4(\epsilon) := \frac{\epsilon^{-1}\left[\widetilde{\mathcal{E}}(\cdot;\upsilon + \epsilon\upsilon') - \widetilde{\mathcal{E}}(\cdot;\upsilon)\right]_T}{\left\langle \boldsymbol{N}_\upsilon, \upsilon' \right\rangle_{L^2([0,1])}}, \tag{B.8}$$

which is recovered in (B.1) by setting $\Phi = \left[\widetilde{\mathcal{E}}(\cdot;\upsilon)\right]_T$, $\Gamma = \upsilon$, and $\Gamma' = \upsilon'$, in (B.1). Again, we test with respect to $\breve{\nu}$ and $\breve{\nu}'$ which is defined in terms of υ and υ', cf. (3.7). We show $\kappa_4(\epsilon)$ and $|1 - \kappa_4(\epsilon)|$, testing the "normal vector" given in (4.14), in Figure A2.5. From

Figure A2.5a and Figure A2.5b, we can see the value that $\kappa_4(\epsilon) \approx 1$ for intermediate values ϵ. In general, we also notice that as we refine the time discretization, $\kappa_4(\epsilon) \to 1$.

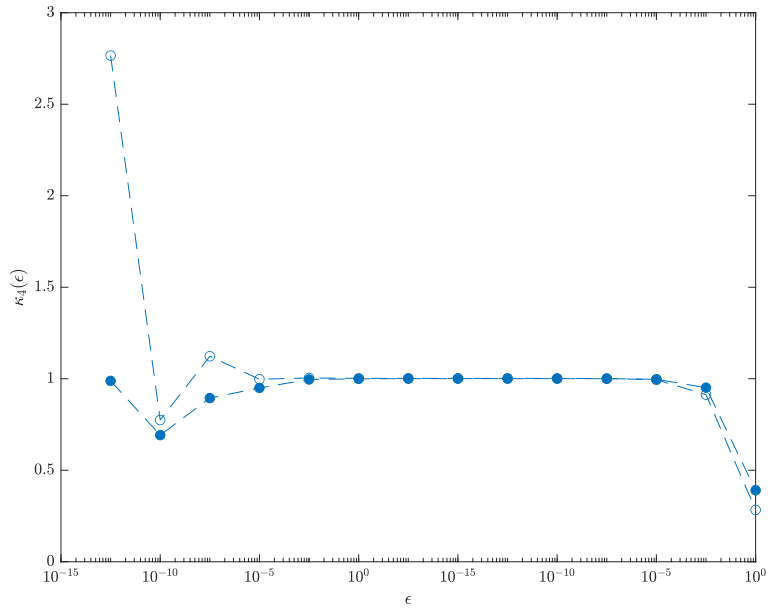


(a)

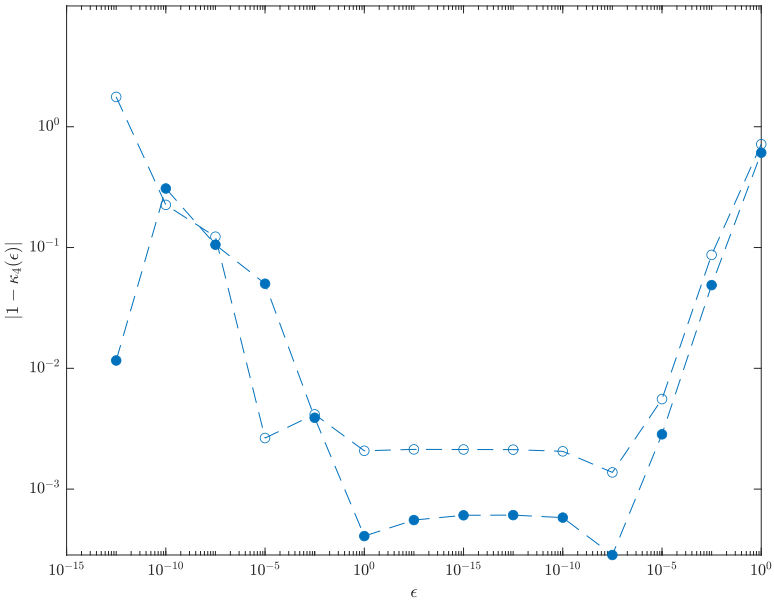


(b)

FIGURE A2.4: Dependence of (a) $\kappa_3(\epsilon)$ and (b) $|1-\kappa_3(\epsilon)|$ on ϵ, for the perturbation ν' given in (B.7a) (blue circles), (B.7b) (red triangles), (B.7c) (yellow squares), and (B.7d) (purple squares). Results for two different time discretizations Δt are shown: (empty symbols) $\Delta t = 1 \times 10^{-2}$ and (filled symbols) $\Delta t = 5 \times 10^{-3}$.



(a)



(b)

FIGURE A2.5: Dependence of (a) $\kappa_4(\epsilon)$ and (b) $|1-\kappa_4(\epsilon)|$ on ϵ, for the perturbation ν' given in (B.7a) (blue circles). Results for two different time discretizations Δt are shown: (empty symbols) $\Delta t = 1 \times 10^{-2}$ and (filled symbols) $\Delta t = 5 \times 10^{-3}$.

Appendix C

Gradient of the Error Functional $\mathcal{J}_2$

Here we discuss computation of the gradients $\nabla_{\varphi}^{L^2}\mathcal{J}_2$ and $\nabla_{\varphi}^{H^2}\mathcal{J}_2$ of the error functional (3.12). The difference with respect to the formulation used in Section 3.3 is that functional (3.12) is defined in the Fourier space and we adopt with suitable modifications the approach developed in (Farazmand et al. 2011). Proceeding as in Section 3.3, we first compute the Gâteaux differential of the error functional (3.12) with respect to φ

$$\mathcal{J}_2'(\varphi;\varphi') = \frac{1}{2T}\int_{t=0}^{T}\int_{k=0}^{k_c}\Big(\left[E_{\widetilde{\omega}}(\cdot,k;\varphi)\right]_T - \left[E_w(\cdot,k)\right]_T\Big)\left(\int_{\mathscr{C}(k)}\widehat{\widetilde{\omega}}\,\overline{\widehat{\widetilde{\omega}}}' + \overline{\widehat{\widetilde{\omega}}}\,\widehat{\widetilde{\omega}}'\,dS(\boldsymbol{k})\right)dk\,dt, \tag{C.1}$$

where $\bar{\cdot}$ denotes the complex conjugate and $\widehat{\widetilde{\omega}}'$ is the Fourier transform of the solution $\widetilde{\omega}'$ to (3.17). We note that the gradients $\nabla_{\varphi}^{L^2}\mathcal{J}_2$ and $\nabla_{\varphi}^{H^2}\mathcal{J}_2$ satisfy Riesz identities analogous to (3.19). Next we introduce new adjoint fields $\widetilde{\omega}^*$ and $\widetilde{\psi}^*$ assumed to satisfy the same adjoint system (3.22), but with a different source term W whose form is to be determined. Utilizing Parseval's identity and the fact that all fields are real-valued in

physical space, we rewrite the duality relation (3.21) as

$$\begin{aligned}\left(\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}, \mathcal{K}^*\begin{bmatrix}\widetilde{\omega}^*\\ \widetilde{\psi}^*\end{bmatrix}\right) &= \frac{1}{2}\left(\widehat{\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}}, \widehat{\mathcal{K}^*\begin{bmatrix}\widetilde{\omega}^*\\ \widetilde{\psi}^*\end{bmatrix}}\right) + \frac{1}{2}\overline{\left(\widehat{\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}}, \widehat{\mathcal{K}^*\begin{bmatrix}\widetilde{\omega}^*\\ \widetilde{\psi}^*\end{bmatrix}}\right)},\\ &= \frac{1}{2T}\int_{t=0}^{T}\int_{k=0}^{k_c}\int_{\mathscr{C}(k)} \widehat{\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}}\cdot\overline{\widehat{\mathcal{K}^*\begin{bmatrix}\widetilde{\omega}^*\\ \widetilde{\psi}^*\end{bmatrix}}} + \overline{\widehat{\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}}}\cdot\widehat{\mathcal{K}^*\begin{bmatrix}\widetilde{\omega}^*\\ \widetilde{\psi}^*\end{bmatrix}}\, dS(\boldsymbol{k})\, dk\, dt.\end{aligned} \tag{C.2}$$

Combining (3.17), (3.21), (3.22), (C.1) and (C.2) results in

$$\left(\begin{bmatrix}\widetilde{\omega}'\\ \widetilde{\psi}'\end{bmatrix}, \mathcal{K}^*\begin{bmatrix}\widetilde{\omega}^*\\ \widetilde{\psi}^*\end{bmatrix}\right) = \overbrace{\frac{1}{2T}\int_{t=0}^{T}\int_{k=0}^{k_c}\Big([E_{\widetilde{\omega}}(\cdot,k;\varphi)]_T - [E_w(\cdot,k)]_T\Big)\left(\int_{\mathscr{C}(k)}\widehat{\widetilde{\omega}}\overline{\widehat{\widetilde{\omega}}}' + \overline{\widehat{\widetilde{\omega}}}\widehat{\widetilde{\omega}}'\, dS(\boldsymbol{k})\right)dk\, dt}^{\mathcal{J}_2'(\varphi;\varphi')},$$

from which we deduce the form of the source term in the adjoint system as

$$\widehat{W}(t,\boldsymbol{k}) = \Big([E_{\widetilde{\omega}}(\cdot,k;\varphi)]_T - [E_w(\cdot,k)]_T\Big)\,\widehat{\widetilde{\omega}}(t,\boldsymbol{k}). \tag{C.3}$$

Once the adjoint system (3.22) with the source term (C.3) is solved, the L^2 gradient $\nabla_\varphi^{L^2}\mathcal{J}_2$ can be computed using expression (3.25). The Sobolev gradient $\nabla_\varphi^{H^2}\mathcal{J}_2$ is then obtained as discussed in Section 3.3 by solving system (3.26). In summary, the difference in the computation of the gradients of the error functionals $\mathcal{J}_1$ and $\mathcal{J}_2$ is confined to the form of the source term W in the adjoint system (3.22).

Appendix D

Decomposition of the Non-linear Term

Here we shall justify why the decomposition of non-linear term in (3.4) with use of the sharp spectral filter (3.3) is preferred for this work; further details can be found in (Pope 2000). For the Navier-Stokes system, we typically define the analogous *residual-stress tensor* as

$$M^R := \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla} \omega} - \widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}, \tag{D.1}$$

which is the difference between the filtered product and product of the filtered quantities. Typically for the 3D incompressible Navier-Stokes system, when substituting in the nonlinear term to formulate the LES equations, M^R is written in terms of the residual kinetic energy and the anisotropic residual-stress tensor. In Leonard (1975), a three component decomposition of M^R was proposed, however as pointed out by Speziale (1985), two of these components are not Galilean invariant. Instead, Germano (1986) proposed a Galilean-invariant decomposition. The analogous vorticity form of this decomposition

in 2D is

$$\begin{aligned} M^R &= \widetilde{\boldsymbol{u}\cdot\boldsymbol{\nabla}\omega} - \widetilde{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\widetilde{\omega}, \\ &= \widetilde{(\widetilde{\boldsymbol{u}}+\boldsymbol{u}')\cdot\boldsymbol{\nabla}(\widetilde{\omega}+\omega')} - \widetilde{(\widetilde{\boldsymbol{u}}+\boldsymbol{u}')}\cdot\boldsymbol{\nabla}\widetilde{(\widetilde{\omega}+\omega')}, \\ &= \widetilde{\widetilde{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\widetilde{\omega}} + \widetilde{\widetilde{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega'} + \widetilde{\boldsymbol{u}'\cdot\boldsymbol{\nabla}\widetilde{\omega}} + \widetilde{\boldsymbol{u}'\cdot\boldsymbol{\nabla}\omega'} - \widetilde{\widetilde{\boldsymbol{u}}}\cdot\boldsymbol{\nabla}\widetilde{\widetilde{\omega}} - \widetilde{\widetilde{\boldsymbol{u}}}\cdot\boldsymbol{\nabla}\widetilde{\omega'} - \widetilde{\boldsymbol{u}'}\cdot\boldsymbol{\nabla}\widetilde{\widetilde{\omega}} - \widetilde{\boldsymbol{u}'}\cdot\boldsymbol{\nabla}\widetilde{\omega'}, \\ &= \mathcal{L} + \mathcal{C} + \mathcal{R}, \end{aligned} \tag{D.2}$$

where the *Leonard stresses* are

$$\mathcal{L} := \widetilde{\widetilde{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\widetilde{\omega}} - \widetilde{\widetilde{\boldsymbol{u}}}\cdot\boldsymbol{\nabla}\widetilde{\widetilde{\omega}}, \tag{D.3}$$

the *cross stresses* are

$$\mathcal{C} := \widetilde{\widetilde{\boldsymbol{u}}\cdot\boldsymbol{\nabla}\omega'} + \widetilde{\boldsymbol{u}'\cdot\boldsymbol{\nabla}\widetilde{\omega}} - \widetilde{\widetilde{\boldsymbol{u}}}\cdot\boldsymbol{\nabla}\widetilde{\omega'} - \widetilde{\boldsymbol{u}'}\cdot\boldsymbol{\nabla}\widetilde{\widetilde{\omega}}, \tag{D.4}$$

and the *subgrid-scale Reynolds stresses* are

$$\mathcal{R} := \widetilde{\boldsymbol{u}'\cdot\boldsymbol{\nabla}\omega'} - \widetilde{\boldsymbol{u}'}\cdot\boldsymbol{\nabla}\widetilde{\omega'}. \tag{D.5}$$

This decomposition is often used for a general filtering operation $\widetilde{\cdot}$, but in the present investigation we are using the filter kernel defined in (3.3), which is a sharp spectral filter. Using the fact that a sharp spectral filter is idempotent (see Appendix E), we

note

$$\begin{aligned}\mathcal{L} &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}} - \widetilde{\widetilde{\boldsymbol{u}}} \cdot \boldsymbol{\nabla}\widetilde{\widetilde{\omega}}, \quad \leftarrow \widetilde{\widetilde{\omega}} = \widetilde{\omega}, \\ &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}} - \widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega},\end{aligned}$$

$$\begin{aligned}\mathcal{C} &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega'} + \widetilde{\boldsymbol{u}' \cdot \boldsymbol{\nabla}\widetilde{\omega}} - \widetilde{\widetilde{\boldsymbol{u}}} \cdot \boldsymbol{\nabla}\widetilde{\omega'} - \widetilde{\boldsymbol{u}'} \cdot \boldsymbol{\nabla}\widetilde{\widetilde{\omega}}, \quad \leftarrow \widetilde{\omega'} = \widetilde{\boldsymbol{u}'} = 0, \\ &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega'} + \widetilde{\boldsymbol{u}' \cdot \boldsymbol{\nabla}\widetilde{\omega}},\end{aligned}$$

$$\begin{aligned}\mathcal{R} &= \widetilde{\boldsymbol{u}' \cdot \boldsymbol{\nabla}\omega'} - \widetilde{\boldsymbol{u}'} \cdot \boldsymbol{\nabla}\widetilde{\omega'}, \quad \leftarrow \widetilde{\omega'} = \widetilde{\boldsymbol{u}'} = 0, \\ &= \widetilde{\boldsymbol{u}' \cdot \boldsymbol{\nabla}\omega'}.\end{aligned}$$

Denoting $M^{\omega} = \mathcal{C} + \mathcal{R}$ and using eqs. (D.1) and (D.3) to (D.5), we can rewrite the nonlinear term as

$$\begin{aligned}M^R &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega} - \widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}, \\ \mathcal{L} + \mathcal{C} + \mathcal{R} &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega} - \widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}, \\ \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}} - \widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega} + \mathcal{C} + \mathcal{R} &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega} - \widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}, \\ \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla}\omega} &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}} + \mathcal{C} + \mathcal{R}, \\ &= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}} + M^{\omega}, \end{aligned} \tag{D.6}$$

showing that we obtain the same decomposition as shown in (E.6) (see Appendix E). Thus, the two decompositions, (D.2) and (D.6), are equivalent when using a sharp spectral filter defined in (3.3). As discussed in Pope (2000), each term in the decomposition given in (E.6) is a filtered quantity. Thus, they do not have contributions coming from wavenumbers greater than the maximum resolved wavenumber in Fourier space. Since using (D.1) and (D.2) for any general filtering operation requires contributions greater than the maximum resolved (up to twice the maximum resolved wavenumber, due to the

Leonard stresses), this would be difficult for a residual stress model to compensate for. Thus, we argue that using a closure model with a sharp spectral filter and decomposition (D.6), where M^{ω} is modelled, is a preferred approach since the decomposition contains contributions from wavenumbers up to the maximum resolved wavenumber.

Appendix E

Galilean Invariance of the 2D Vorticity Equation

A system of equations and closure model are said to be Galilean invariant if they remain unchanged in different inertial frames. This is a key aspect when considering fluid flows, as it ensures that models do not depend on the frame of reference.. In this section, we establish the Galilean invariance of the 2D vorticity equation (3.1), its filtered form, cf. (E.7), and the Leith model (3.5), in a general form.

A1 2D Vorticity Equation

We will prove that the 2D vorticity equation and the associated LES system are Galilean invariant. The transformations we consider between the two coordinate systems are

$$\mathbf{x}^* = \mathbf{x} - \boldsymbol{v}\, t, \tag{E.1a}$$

$$t^* = t, \tag{E.1b}$$

where $\boldsymbol{v} \in \mathbb{R}^2$ is a fixed velocity. An equation is said to be *Galilean invariant* if upon applying the transformation (E.1), the equation remains unchanged. We state our first

result as a theorem, and prove it below.

Theorem 2. *The 2D Navier-Stokes system (3.1a) is Galilean invariant.*

Proof. First, we write our operators under the coordinate transformations (E.1)

$$\frac{d}{dt^*} = \frac{d}{dt}, \tag{E.2a}$$

$$\frac{d\mathbf{x}^*}{dt} = \boldsymbol{u}^* = \frac{d\mathbf{x}}{dt} - \boldsymbol{v},$$

$$\boldsymbol{u}^*(\mathbf{x}^*, t) = \boldsymbol{u}(\mathbf{x}, t) - \boldsymbol{v}, \tag{E.2b}$$

$$\frac{\partial}{\partial t^*} = \frac{\partial t}{\partial t^*}\frac{\partial}{\partial t} + \frac{\partial x}{\partial t^*}\frac{\partial}{\partial x} + \frac{\partial y}{\partial t^*}\frac{\partial}{\partial y} + \frac{\partial z}{\partial t^*}\frac{\partial}{\partial z} = \frac{\partial}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}, \tag{E.2c}$$

$$\frac{\partial}{\partial x^*} = \frac{\partial t}{\partial x^*}\frac{\partial}{\partial t} + \frac{\partial x}{\partial x^*}\frac{\partial}{\partial x} + \frac{\partial y}{\partial x^*}\frac{\partial}{\partial y} + \frac{\partial z}{\partial x^*}\frac{\partial}{\partial z} = \frac{\partial}{\partial x}, \tag{E.2d}$$

$$\frac{\partial}{\partial y^*} = \frac{\partial t}{\partial y^*}\frac{\partial}{\partial t} + \frac{\partial x}{\partial y^*}\frac{\partial}{\partial x} + \frac{\partial y}{\partial y^*}\frac{\partial}{\partial y} + \frac{\partial z}{\partial y^*}\frac{\partial}{\partial z} = \frac{\partial}{\partial y}, \tag{E.2e}$$

$$\frac{\partial}{\partial z^*} = \frac{\partial t}{\partial z^*}\frac{\partial}{\partial t} + \frac{\partial x}{\partial z^*}\frac{\partial}{\partial x} + \frac{\partial y}{\partial z^*}\frac{\partial}{\partial y} + \frac{\partial z}{\partial z^*}\frac{\partial}{\partial z} = \frac{\partial}{\partial z}, \tag{E.2f}$$

$$\boldsymbol{\nabla}^* = \boldsymbol{\nabla}, \tag{E.2g}$$

$$\boldsymbol{\Delta}^* = \boldsymbol{\Delta}. \tag{E.2h}$$

We note that $\boldsymbol{v}$ is constant, so $\boldsymbol{\nabla}\boldsymbol{v} = 0$, $\boldsymbol{\Delta}\boldsymbol{v} = 0$, and $\frac{\partial \boldsymbol{v}}{\partial t} = 0$. We also note that the vorticity is Galilean invariant, since

$$\begin{aligned} w^* &= \boldsymbol{\nabla} \times \boldsymbol{u}^*, \\ &= -\boldsymbol{\nabla}^\perp \cdot \boldsymbol{u}^*, \\ &= \boldsymbol{\nabla}^\perp \cdot (\boldsymbol{u} - \boldsymbol{v}), \\ &= \boldsymbol{\nabla}^\perp \cdot \boldsymbol{u} - \boldsymbol{\nabla}^\perp \cdot \boldsymbol{v}, \\ &= w. \end{aligned} \tag{E.3}$$

Now, we consider the 2D vorticity equation (3.1a), written with the velocity field $\boldsymbol{u}(t,\mathbf{x}) = (u_1,u_2)^T = \boldsymbol{\nabla}^{\perp}\psi$,

$$\partial_t w + \boldsymbol{u}\cdot\boldsymbol{\nabla} w = \nu_N \boldsymbol{\Delta} w - \alpha w + f_\omega, \tag{E.4a}$$

$$w = \boldsymbol{\nabla}\times\boldsymbol{u}. \tag{E.4b}$$

Since we already showed above (E.4b) is Galilean invariant, we apply transformation (E.1) to (E.4a), to obtain

$$\begin{aligned}
\frac{\partial w^*}{\partial t^*} + \boldsymbol{u}^*\cdot\boldsymbol{\nabla}^* w^* &= \nu_N \boldsymbol{\Delta}^* w^* - \alpha w^* + f_\omega,\\
\frac{\partial w}{\partial t^*} + \boldsymbol{u}^*\cdot\boldsymbol{\nabla} w &= \nu_N \boldsymbol{\Delta} w - \alpha w + f_\omega,\\
\frac{\partial w}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} w + (\boldsymbol{u}-\boldsymbol{v})\cdot\boldsymbol{\nabla} w &= \nu_N \boldsymbol{\Delta} w - \alpha w + f_\omega,\\
\frac{\partial w}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla} w + \boldsymbol{u}\cdot\boldsymbol{\nabla} w - \boldsymbol{v}\cdot\boldsymbol{\nabla} w &= \nu_N \boldsymbol{\Delta} w - \alpha w + f_\omega,\\
\frac{\partial w}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla} w &= \nu_N \boldsymbol{\Delta} w - \alpha w + f_\omega.
\end{aligned}$$

□

Therefore, system (E.4) is Galilean invariant. Frisch (1995) notes that adding a driving force to Navier-Stokes system usually breaks Galilean invariance (except if the force is delta-correlated in time).

A2 Filtered 2D Vorticity Equation

Now, we consider the filtered version of equation (E.4). Using the sharp spectral filter defined in (3.3), we denote $\widetilde{\boldsymbol{u}} := G_\delta * \boldsymbol{u}$. Noting that the operation of filtering and differentiation commute, we consider the decomposition of filtered variables $\widetilde{\omega} := G_\delta * \omega = \boldsymbol{\nabla}\times\widetilde{\boldsymbol{u}}$, such that $\omega = \widetilde{\omega} + \omega'$, where $\widetilde{\omega}$ is the resolved field and ω' is the unresolved field. As stated in Section 3.2.1, the maximum resolved wavenumber $k_c < k_b$, such that $\widetilde{f_\omega} = f_\omega$.

In addition, we note that the sharp spectral filter is a projection operator, thus it is *idempotent*, hence

$$\widetilde{\omega'} = \widetilde{\omega - \widetilde{\omega}} = \widetilde{\omega} - \widetilde{\widetilde{\omega}} = \widetilde{\omega} - \widetilde{\omega} = 0.$$

It should also be noted that the operation of filtering and differentiation commute and since $\boldsymbol{v}$ is a constant vector, the filtering operation gives us

$$\widetilde{\boldsymbol{v} \cdot \boldsymbol{\nabla}\omega} = \boldsymbol{v} \cdot \widetilde{\boldsymbol{\nabla}\omega} = \boldsymbol{v} \cdot \boldsymbol{\nabla}\widetilde{\omega}.$$

Applying the filter defined in (3.3) to (E.4), we obtain

$$\partial_t \widetilde{\omega} + \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla}\omega} = \nu_N \boldsymbol{\Delta}\widetilde{\omega} - \alpha\widetilde{\omega} + f_\omega, \tag{E.5a}$$

$$\widetilde{\omega} = \boldsymbol{\nabla} \times \widetilde{\boldsymbol{u}}. \tag{E.5b}$$

Considering the nonlinear term in (E.5a)

$$\begin{aligned}
\widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla}\omega} &= \widetilde{(\widetilde{\boldsymbol{u}} + \boldsymbol{u}') \cdot \boldsymbol{\nabla}(\widetilde{\omega} + \omega')}, \\
&= \widetilde{(\widetilde{\boldsymbol{u}} + \boldsymbol{u}') \cdot \boldsymbol{\nabla}\widetilde{\omega}} \quad + \quad \widetilde{(\widetilde{\boldsymbol{u}} + \boldsymbol{u}') \cdot \boldsymbol{\nabla}\omega'} \quad , \\
&= \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}} + \underbrace{\widetilde{\boldsymbol{u}' \cdot \boldsymbol{\nabla}\widetilde{\omega}} + \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\omega'} + \widetilde{\boldsymbol{u}' \cdot \boldsymbol{\nabla}\omega'}}_{M^\omega}, \\
M^\omega &:= \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla}\omega} - \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla}\widetilde{\omega}}.
\end{aligned} \tag{E.6}$$

As stated by Pope (2000), the decomposition of the nonlinear term given in (E.6) is preferable for a sharp spectral filter and is used in (3.4). This is because each component can be represented exactly in terms of the resolved modes (see Appendix D for further discussion on the nonlinear decomposition). Substituting (E.6) into the filtered vorticity

equation gives

$$\partial_t \widetilde{\omega} + \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} = \nu_N \boldsymbol{\Delta} \widetilde{\omega} - \alpha \widetilde{\omega} + f_\omega - M^\omega, \tag{E.7a}$$

$$\widetilde{\omega} = \boldsymbol{\nabla} \times \widetilde{\boldsymbol{u}}. \tag{E.7b}$$

Now we consider the Galilean invariance of (E.7), as the following theorem.

Theorem 3. *The filtered 2D Navier-Stokes system (E.7) is Galilean invariant.*

Proof. It simply follows from (E.3), that $\widetilde{\omega}^* = \boldsymbol{\nabla} \times \widetilde{\boldsymbol{u}}^* = \widetilde{\omega}$, so we will focus on proving (E.7a) is Galilean invariant. Since the $\boldsymbol{v}$ is a constant vector, it is invariant with respect to the filtering operation, hence $\widetilde{\boldsymbol{v}} = \boldsymbol{v}$. First, we will prove M^ω is Galilean invariant,

$$\begin{aligned}
M^{\omega^*} &= \widetilde{\boldsymbol{u}^* \cdot \boldsymbol{\nabla}^* \omega^*} - \widetilde{\widetilde{\boldsymbol{u}^*} \cdot \boldsymbol{\nabla}^* \widetilde{\omega}^*}, \\
&= \overbrace{(\boldsymbol{u} - \boldsymbol{v}) \cdot \boldsymbol{\nabla} \omega}^{} - \overbrace{\widetilde{(\boldsymbol{u} - \boldsymbol{v}) \cdot \boldsymbol{\nabla} \widetilde{\omega}}}^{}, \\
&= \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla} \omega} - \widetilde{\boldsymbol{v} \cdot \boldsymbol{\nabla} \omega} - \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} + \widetilde{\widetilde{\boldsymbol{v}} \cdot \boldsymbol{\nabla} \widetilde{\omega}}, \quad \leftarrow \boldsymbol{v} \text{ is constant and } \widetilde{\boldsymbol{v}} = \boldsymbol{v} \\
&= \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla} \omega} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} - \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\widetilde{\omega}}, \quad \leftarrow \widetilde{\widetilde{\omega}} = \widetilde{\omega} \text{ since (3.3) is a idempotent} \\
&= \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla} \omega} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} - \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega}, \\
&= \widetilde{\boldsymbol{u} \cdot \boldsymbol{\nabla} \omega} - \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} = \tau^\omega.
\end{aligned}$$

Using this, we can prove (E.7) is Galilean invariant:

$$\begin{aligned}
&\frac{\partial \widetilde{\omega}^*}{\partial t^*} + \widetilde{\widetilde{\boldsymbol{u}^*} \cdot \boldsymbol{\nabla}^* \widetilde{\omega}^*} = \nu_N \boldsymbol{\Delta}^* \widetilde{\omega}^* - \alpha \widetilde{\omega}^* + f_\omega - \tau^{\omega^*}, \\
&\frac{\partial \widetilde{\omega}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} + \overbrace{\widetilde{\boldsymbol{u} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega}}}^{} = \nu_N \boldsymbol{\Delta} \widetilde{\omega} - \alpha \widetilde{\omega} + f_\omega - \tau^\omega, \\
&\frac{\partial \widetilde{\omega}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} - \widetilde{\boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega}} + \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} = \nu_N \boldsymbol{\Delta} \widetilde{\omega} - \alpha \widetilde{\omega} + f_\omega - \tau^\omega, \quad \leftarrow \widetilde{\boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega}} = \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} \\
&\frac{\partial \widetilde{\omega}}{\partial t} + \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} - \boldsymbol{v} \cdot \boldsymbol{\nabla} \widetilde{\omega} + \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} = \nu_N \boldsymbol{\Delta} \widetilde{\omega} - \alpha \widetilde{\omega} + f_\omega - \tau^\omega, \\
&\frac{\partial \widetilde{\omega}}{\partial t} + \widetilde{\widetilde{\boldsymbol{u}} \cdot \boldsymbol{\nabla} \widetilde{\omega}} = \nu_N \boldsymbol{\Delta} \widetilde{\omega} - \alpha \widetilde{\omega} + f_\omega - \tau^\omega.
\end{aligned}$$

□

A3 Leith Model

Lastly, for our closure problem, we must show that the considered closure model is also Galilean invariant. Letting $M = -M^{\omega}$ in (E.7), we wish to model this term using an eddy viscosity closure model given in (3.5),

$$M \approx \widetilde{M} = \boldsymbol{\nabla} \cdot (\widetilde{\nu \boldsymbol{\nabla} \widetilde{\omega}}).$$

Since $\omega^* = \omega$ and the Reynolds stresses in the Leith model depend on the vorticity (more specifically, on $\boldsymbol{\nabla}\widetilde{\omega}$), the Leith model is Galilean invariant. Moreover, the closure term $\widetilde{M}$ is also Galilean invariant.

Bibliography

Adams, R. A. and Fournier, J. F. (2003). *Sobolev Spaces*. Second. Elsevier.

Alexakis, A. and Doering, C. R. (2006). Energy and enstrophy dissipation in steady state 2d turbulence. *Physics Letters A* 359(6), 652–657. ISSN: 0375-9601.

Alimo, R., Cavaglieri, D., Beyhaghi, P., and Bewley, T. R. (2021). Design of IMEXRK time integration schemes via Delaunay-based derivative-free optimization with non-convex constraints and grid-based acceleration. *J. Glob. Optim.* 79(3), 567–591. ISSN: 1573-2916.

Ayala, D. and Protas, B. (2014a). Maximum Palinstrophy Growth in 2D Incompressible Flows. *Journal of Fluid Mechanics* 742, 340–367.

Ayala, D. and Protas, B. (2014b). Vortices, maximum growth and the problem of finite-time singularity formation. *Fluid Dynamics Research* 46(3), 031404.

Ayala, D. and Protas, B. (2017). Extreme Vortex States and the Growth of Enstrophy in 3D Incompressible Flows. *Journal of Fluid Mechanics* 818, 772–806.

Ayala, D., Doering, C. R., and Simon, T. M. (2018). Maximum palinstrophy amplification in the two-dimensional Navier-Stokes equations. *Journal of Fluid Mechanics* 837, 839–857.

Bardos, C., Golse, F., and Levermore, D. (1991). Fluid dynamic limits of kinetic equations. I. Formal derivations. *Journal of statistical physics* 63(1), 323–344.

Batchelor, G. K. (1969). Computation of the Energy Spectrum in Homogeneous Two-Dimensional Turbulence. *Phys. Fluids* 12(12), II-233–239.

Berger, M. S. (1977). *Nonlinearity and Functional Analysis*. Academic Press.

Bewley, T. R. (2009). *Numerical Renaissance*. Renaissance Press.

Boffetta, G. (2007). Energy and enstrophy fluxes in the double cascade of two-dimensional turbulence. *Journal of Fluid Mechanics* 589, 253–260.

Boffetta, G. and Ecke, R. E. (2012). Two-Dimensional Turbulence. *Annual Review of Fluid Mechanics* 44(1), 427–451.

Boffetta, G. and Musacchio, S. (2010). Evidence for the double cascade scenario in two-dimensional turbulence. *Phys. Rev. E* 82 (1), 016307.

Bracco, A. and McWilliams, J. C. (2010). Reynolds-number dependency in homogeneous, stationary two-dimensional turbulence. *Journal of Fluid Mechanics* 646, 517–526.

Bukshtynov, V. and Protas, B. (2013). Optimal reconstruction of material properties in complex multiphysics phenomena. *J. Comput. Phys.* 242, 889–914. ISSN: 0021-9991.

Bukshtynov, V., Volkov, O., and Protas, B. (2011). On optimal reconstruction of constitutive relations. *Physica D* 240(16), 1228–1244. ISSN: 0167-2789.

Ciampa, G., Crippa, G., and Spirito, S. (2021). Strong convergence of the vorticity for the 2D Euler equations in the inviscid limit. *Arch. Ration. Mech. Anal.* 240(1), 295–326. ISSN: 0003-9527.

Constantin, P. (2007). On the Euler equations of incompressible fluids. *Bulletin of the American Mathematical Society* 44(4), 603–621.

Constantin, P., Drivas, T. D., and Elgindi, T. M. (2022). Inviscid Limit of Vorticity Distributions in the Yudovich Class. *Communications on Pure and Applied Mathematics* 75(1), 60–82.

Darrigol, O. (2005). *Worlds of flow*. A history of hydrodynamics from the Bernoullis to Prandtl. Oxford University Press, New York, xiv+356. ISBN: 0-19-856843-6.

Dascaliuc, R. and Grujić, Z. (2012). Anomalous dissipation and energy cascade in 3D inviscid flows. *Communications in Mathematical Physics* 309(3), 757–770.

Davidson, P. A. (2015). *Turbulence: An introduction for scientists and engineers*. Second. Oxford University Press, Oxford, xvi+630. ISBN: 978-0-19-872259-5; 978-0-19-872258-8.

Driscoll, T. A., Hale, N., and Trefethen, L. N. (2014). *Chebfun Guide*. Oxford, UK: Pafnuty Publications.

Duraisamy, K., Iaccarino, G., and Xiao, H. (2019). Turbulence Modeling in the Age of Data. *Annu. Rev. Fluid Mech.* 51(1), 357–377.

Duraisamy, K. (2021). Perspectives on machine learning-augmented Reynolds-averaged and large eddy simulation models of turbulence. *Phys. Rev. Fluids* 6 (5), 050504.

Eyink, G. L. and Drivas, T. D. (2015). Spontaneous Stochasticity and Anomalous Dissipation for Burgers Equation. *Journal of Statistical Physics* 158(2), 386–432. ISSN: 1572-9613.

Farazmand, M. and Sapsis, T. P. (2017). A variational approach to probing extreme events in turbulent dynamical systems. *Science Advances* 3(9), e1701533.

Farazmand, M. M., Kevlahan, N. K.-R., and Protas, B. (2011). Controlling the dual cascade of two-dimensional turbulence. *Journal of Fluid Mechanics* 668, 202–222.

Fefferman, C. L. (2006). Existence and smoothness of the Navier-Stokes equation. In: *The millennium prize problems*. Clay Math. Inst., Cambridge, MA, 57–67.

Filho, M., Mazzucato, A., and Nussenzveig Lopes, H. (2006). Weak Solutions, Renormalized Solutions and Enstrophy Defects in 2D Turbulence. English (US). *Archive for Rational Mechanics and Analysis* 179(3), 353–387. ISSN: 0003-9527.

Frigo, M. and Johnson, S. G. (2005). The Design and Implementation of FFTW3. *Proceedings of the IEEE* 93(2). Special issue on "Program Generation, Optimization, and Platform Adaptation", 216–231.

Frisch, U. (1995). *Turbulence*. The legacy of A. N. Kolmogorov. Cambridge University Press, Cambridge, xiv+296. ISBN: 0-521-45103-5.

Gamahara, M. and Hattori, Y. (2017). Searching for turbulence models by artificial neural network. *Phys. Rev. Fluids* 2 (5), 054604.

Germano, M. (1986). A proposal for a redefinition of the turbulent stresses in the filtered Navier–Stokes equations. *The Physics of fluids* 29(7), 2323–2324.

Goto, S., Saito, Y., and Kawahara, G. (2017). Hierarchy of antiparallel vortex tubes in spatially periodic turbulence at high Reynolds numbers. *Phys. Rev. Fluids* 2 (6), 064603.

Graham, J. P. and Ringler, T. (2013). A framework for the evaluation of turbulence closures used in mesoscale ocean large-eddy simulations. *Ocean Modelling* 65, 25–39. ISSN: 1463-5003.

Gunzburger, M. D. (2003). *Perspectives in Flow Control and Optimization*. SIAM.

Hou, T. Y. (2009). Blow-up or no blow-up? A unified computational and analytic approach to 3D incompressible Euler and Navier–Stokes equations. *Acta Numer.* 18, 277–346. ISSN: 0962-4929.

Jeong, I.-J. and Yoneda, T. (2021). Enstrophy dissipation and vortex thinning for the incompressible 2D Navier–Stokes equations. *Nonlinearity* 34(4), 1837–1853.

Jimenez, J. (2018). Machine-aided turbulence theory. *J. Fluid Mech.* 854, R1.

Kármán, T. H. von and Howarth, L. (1938). On the Statistical Theory of Isotropic Turbulence. *Proc. Roy. Soc. London A* 164, 192–215.

Kevlahan, N. K.-R. and Farge, M. (1997). Vorticity filaments in two-dimensional turbulence: creation, stability and effect. *J. Fluid Mech.* 346, 49–76.

Kiya, M. and Arie, M. (1979). Helmholtz instability of a vortex sheet in uniform shear flow. *The Physics of Fluids* 22(2), 378–379.

Kolmogorov, A. N. (1962). A refinement of previous hypotheses concerning the local structure of turbulence in a viscous incompressible fluid at high Reynolds number. *Journal of Fluid Mechanics* 13(1), 82–85.

Kolmogorov, A. (1941a). Dissipation of energy in locally isotropic turbulence. *Dokl. Akad. Nauk SSSR* 32. (translated and reprinted 1991 in Proc. R. Soc. Lond. A **434**, 15–17), 16–18.

Kolmogorov, A. (1941b). On degeneration (decay) of isotropic turbulence in an incompressible viscous liquid. *Dokl. Akad. Nauk SSSR* 31, 538–540.

Kolmogorov, A. (1941c). The local structure of turbulence in incompressible viscous fluid for very large Reynolds number. *Dokl. Akad. Nauk SSSR* 30. (translated and reprinted 1991 in Proc. R. Soc. Lond. A **434**, 9–13), 9–13.

Kraichnan, R. H. (1967). Inertial Ranges in Two-Dimensional Turbulence. *The Physics of Fluids* 10(7), 1417–1423.

Kreiss, H. and Lorenz, J. (2004). *Initial-Boundary Value Problems and the Navier-Stokes Equations*. Vol. 47. Classics in Applied Mathematics. SIAM.

Krüger, T., Kusumaatmaja, H., Kuzmin, A., Shardt, O., Silva, G., and Viggen, E. M. (2017). *The Lattice Boltzmann Method*. 1st ed. Springer Cham. ISBN: 978-3-319-44649-3.

Kutz, J. N. (2017). Deep learning in fluid dynamics. *J. Fluid Mech.* 814, 1–4.

Langford, J. A. and Moser, R. D. (1999). Optimal LES formulations for isotropic turbulence. *J. Fluid Mech.* 398, 321–346. ISSN: 0022-1120.

Le, H. and Moin, P. (1991). An improvement of fractional step methods for the incompressible Navier-Stokes equations. *J. Comput. Phys.* 92(2), 369–379. ISSN: 0021-9991.

Leith, C. E. (1968). Diffusion Approximation for Two-Dimensional Turbulence. *The Physics of Fluids* 11(3), 671–672.

Leith, C. E. (1971). Atmospheric Predictability and Two-Dimensional Turbulence. *Journal of the Atmospheric Sciences* 28(2), 145–161.

Leith, C. (1996). Stochastic models of chaotic systems. *Physica D: Nonlinear Phenomena* 98(2). Nonlinear Phenomena in Ocean Dynamics, 481–491. ISSN: 0167-2789.

Leonard, A. (1975). Energy Cascade in Large-Eddy Simulations of Turbulent Fluid Flows. In: *Turbulent Diffusion in Environmental Pollution*. Ed. by F. N. Frenkiel and R. E. Munn. Vol. 18. Advances in Geophysics. Elsevier, 237–248.

Lesieur, M. (1993). *Turbulence in Fluids*. 2nd. Dordrecht, Boston, London: Kluwer Academic Publishers.

Lindborg, E. and Vallgren, A. (2010). Testing Batchelor's similarity hypotheses for decaying two-dimensional turbulence. *Physics of Fluids* 22(9), 091704.

Lions, J. (1968). *Contrôle Optimal de Systèmes Gouvernés par des Équations aux Dérivées Partielles*. (English translation, Springer-Verlag, New-York (1971)). Paris: Dunod.

Lopes, H. J. N., Seis, C., and Wiedemann, E. (2021). On the vanishing viscosity limit for 2D incompressible flows with unbounded vorticity. *Nonlinearity* 34(5), 3112–3121.

Luenberger, D. (1969). *Optimization by Vector Space Methods*. John Wiley and Sons.

Matharu, P. and Protas, B. (2020). Optimal Closures in a Simple Model for Turbulent Flows. *SIAM J. Sci. Comput.* 42(1), B250–B272.

Matharu, P. and Protas, B. (2022). Optimal eddy viscosity in closure models for two-dimensional turbulent flows. *Phys. Rev. Fluids* 7 (4), 044605.

Maulik, R. and San, O. (2017). A dynamic framework for functional parameterisations of the eddy viscosity coefficient in two-dimensional turbulence. *International Journal of Computational Fluid Dynamics* 31(2), 69–92.

Maulik, R., San, O., and Jacob, J. D. (2020). Spatiotemporally dynamic implicit large eddy simulation using machine learning classifiers. *Physica D* 406, 132409. ISSN: 0167-2789.

Maute, K., Nikbay, M., and Farhat, C. (2003). Sensitivity analysis and design optimization of three-dimensional non-linear aeroelastic systems by the adjoint method. *International Journal for Numerical Methods in Engineering* 56(6), 911–933.

Mazzucato, A. (2022). Remarks on anomalous dissipation for passive scalars. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 380(2218), 20210099.

Nocedal, J. and Wright, S. (2006). *Numerical Optimization*. Second. Springer Series in Operations Research and Financial Engineering. Springer, New York, xxii+664. ISBN: 978-0387-30303-1.

Pappalardo, C. M. and Guida, D. (2018). Use of the Adjoint Method for Controlling the Mechanical Vibrations of Nonlinear Systems. *Machines* 6(2). ISSN: 2075-1702.

Parisi, G. and Frisch, U. (Jan. 1985). On the singularity structure of fully developed turbulence in Turbulence and predictability in geophysical fluid dynamics and climate dynamics. *NTurbulence and Predictability of Geophysical Flows and Climate Dynamics* 88.

Pawar, S. and San, O. (2021). Data assimilation empowered neural network parametrizations for subgrid processes in geophysical flows. *Phys. Rev. Fluids* 6 (5), 050501.

Pope, S. B. (2000). *Turbulent flows*. Cambridge: Cambridge University Press, xxxiv+771. ISBN: 0-521-59886-9.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes 3rd Edition: The Art of Scientific Computations*. Cambridge University Press, xxii+1235. ISBN: 978-0-521-88068-8.

Protas, B., Bewley, T. R., and Hagen, G. (2004). A computational framework for the regularization of adjoint analysis in multiscale PDE systems. *J. Comp. Phys.* 195(1), 49–89. ISSN: 0021-9991.

Protas, B. (2022). Systematic search for extreme and singular behaviour in some fundamental models of fluid mechanics. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 380(2225), 20210035.

Richardson, L. F. (1922). *Weather prediction by numerical process*. Cambridge, UK: Cambridge University Press.

Rodi, W., Constantinescu, G., and Stoesser, T. (2013). *Large-Eddy Simulation in Hydraulics*. CRC Press.

Saint-Raymond, L. (2003). From the BGK model to the Navier–Stokes equations. *Annales Scientifiques de l'École Normale Supérieure* 36(2), 271–317. ISSN: 0012-9593.

Sakajo, T. and Okamoto, H. (1996). Numerical computation of vortex sheet roll-up in the background shear flow. *Fluid Dyn. Res.* 17, 195–212.

Seis, C. (2021). A note on the vanishing viscosity limit in the Yudovich class. *Canad. Math. Bull.* 64(1), 112–122. ISSN: 0008-4395.

Smagorinsky, J. (1963). General circulation experiments with the primitive equations: I. The basic experiment. *Mon. Weather Rev.* 91(3), 99–164.

Speziale, C. G. (1985). Galilean invariance of subgrid-scale stress models in the large-eddy simulation of turbulence. *Journal of fluid mechanics* 156, 55–62.

Sreenivasan, K. R. (1998). An update on the energy dissipation rate in isotropic turbulence. *Physics of Fluids* 10(2), 528–529.

Sreenivasan, K. R. (2019). Turbulent mixing: A perspective. *Proceedings of the National Academy of Sciences* 116(37), 18175–18183.

Taylor, G. I. (1935). Statistical theory of turbulenc. *Proc. Roy. Soc. London A* 151, 421–444.

Taylor, G. I. (1938). The spectrum of turbulence. *Proc. Roy. Soc. London A* 164(919), 476–490.

Tennekes, H. and Lumley, J. L. (1972). *A First Course in Turbulence*. MIT press.

Tran, C. V. and Dritschel, D. G. (2006). Vanishing enstrophy dissipation in two-dimensional Navier-Stokes turbulence in the inviscid limit. *Journal of Fluid Mechanics* 559, 107–116.

Trefethen, L. N. (2013). *Approximation Theory and Approximation Practice*. Philadelphia: SIAM, viii+305 pp.+back matter. ISBN: 978-1-611972-39-9.

Vallgren, A. and Lindborg, E. (2011). The enstrophy cascade in forced two-dimensional turbulence. *Journal of Fluid Mechanics* 671, 168–183.

Yang, Z., Liu, M., and Luo, X. (2021). First-Optimize-Then-Discretize Strategy for the Parabolic PDE Constrained Optimization Problem With Application to the Reheating Furnace. *IEEE Access* 9, 90283–90294.

Yeung, P. K., Zhai, X. M., and Sreenivasan, K. R. (2015). Extreme events in computational turbulence. *Proceedings of the National Academy of Sciences* 112(41), 12633–12638.

Yoneda, T., Goto, S., and Tsuruhashi, T. (2022). Mathematical reformulation of the Kolmogorov–Richardson energy cascade in terms of vortex stretching. *Nonlinearity* 35(3), 1380–1401.

Zhongyang, G. (2022). *Solvability of the modified BGK model*. personal communication.