ADVANCING THE APPLICABILITY AND ROBUSTNESS OF HIGH-ORDER DISCONTINUOUS SPECTRAL ELEMENT METHODS

A Dissertation

by

TARIK DZANIC

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Freddie D. Witherden
Committee Members,	Antony Jameson
	Sharath Girimaji
	Hamn-Ching Chen
Head of Department,	Sharath Girimaji

May 2023

Major Subject: Ocean Engineering

ABSTRACT

High-order discontinuous spectral element methods provide a potential direction for drastically reducing the computational cost of simulating complex unsteady fluid flows due to their high accuracy, geometric flexibility, and computationally efficient data structure. However, for certain application areas such as high Reynolds number flows, supersonic and hypersonic aeronautics, and rarefied gases, these methods can either become computationally intractable, suffer from numerical robustness problems, or are not well established as a numerical tool. This work presents various algorithmic developments for the use of discontinuous spectral element methods in the simulation of problems in continuum fluid mechanics and molecular gas dynamics.

First, as a mechanism for reducing the computational cost of simulating high Reynolds number flows, hybrid turbulence modeling within a high-order discontinuous spectral element framework was explored in the context of the Partially-averaged Navier–Stokes equations. It was observed that larger improvements were generally seen when increasing the discretization accuracy of the Partially-averaged Navier–Stokes method in comparison to methods without models. Furthermore, less sensitivity to the resolution-control parameter was observed with high-order discretizations.

Then, to allow for the use of discontinuous spectral element methods for flows in the supersonic and hypersonic regimes, a novel adaptive filtering approach was introduced as a shock capturing method by formulating convex invariants such as positivity of density and pressure and a local minimum entropy principle as constraints on the solution. The result of this approach is a provably robust, parameter-free method for resolving strong discontinuities that can be applied on general unstructured meshes with relatively low computational cost.

Finally, extensions to non-equilibrium and rarefied flow regimes were then performed through approximations of the polyatomic Boltzmann equation for molecular gas dynamics augmented with the Bhatnagar–Gross–Krook collision model. Through the combination of a positivity-preserving limiter and a discrete velocity model, the method guarantees discrete conservation and positivity of the macroscopic density and pressure. The approach was validated on experiments ranging from shock-dominated flows to direct numerical simulation of three-dimensional compressible turbulent flows, the latter of which is the first instance of such a flow computed by directly solving the Boltzmann equation.

DEDICATION

To my family and friends.

ACKNOWLEDGMENTS

First, I'd like to thank Freddie, I couldn't have asked for a better advisor. Thanks for giving me the support, mentorship, and freedom to do all of this research and for making this Ph.D. as painless as it could be. To Gigi, I would have never been able to get this far if you didn't teach me everything you did, thank you for all your guidance. To Will, half of this research would have never been possible without those cheeky pints. To my family, thank you for your unconditional support, you can read this dissertation and finally stop wondering what I did my Ph.D. on. To Linda, thank you for listening to me talk about shock capturing more than anyone should ever have to. Finally, to my dissertation committee, thank you for making this Ph.D. possible.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supported by a dissertation committee consisting of Professor Freddie D. Witherden of the Department of Ocean Engineering, Professor Antony Jameson of the Department of Aerospace Engineering, Professor Sharath Girimaji of the Department of Ocean Engineering, and Professor Hamn-Ching Chen of the Department of Ocean Engineering.

The conceptualization and methodology for the work presented in Chapter IV was performed in collaboration with Professor Luigi Martinelli in the Department of Mechanical and Aerospace Engineering at Princeton University. All other work conducted for the dissertation was completed by the student independently.

Funding Sources

This work was supported in part by the U.S. Air Force Office of Scientific Research via grant FA9550-21-1-0190 ("Enabling next-generation heterogeneous computing for massively parallel high-order compressible CFD") of the Defense University Research Instrumentation Program (DURIP) under the direction of Dr. Fariba Fahroo.

TABLE OF CONTENTS

LIST OF FIGURES

FIGURE Page

LIST OF TABLES

TABLE Page

1. INTRODUCTION

The efficient and accurate computation of complex unsteady fluid flows remains a driving force in the development of computational fluid dynamics (CFD) techniques. In more recent years, there has been an increasing interest in numerical methods which can provide high-fidelity approximations of the underlying physical phenomena of these fluid flows. These approaches can provide critical insight into problems for which experimental studies may be too costly or simply not possible, and, as a result, can be a useful tool for aiding engineering design and scientific discovery.

The approximation of fluid flows requires the spatio-temporal discretization of some form of governing equations, ranging from the Euler and Navier–Stokes equations for continuum fluid mechanics to the Boltzmann equation for molecular gas dynamics. While the numerical approximation of the temporal component of these equations is an ongoing field of research, a much more substantial amount of work has been performed on advancing the techniques for approximating the spatial component. These spatial discretization methods typically fall into one of three categories: finite difference methods, finite volume methods, and finite element methods. Each of these approaches has advantages and disadvantages with respect to their accuracy, geometric flexibility, computational efficiency, and algorithmic complexity.

In the context of simulating complex unsteady flows on arbitrary domains, the key factors for the choice of spatial discretization method are the accuracy and geometric flexibility of the scheme. For the former, the primary metric for evaluating a scheme's accuracy is its rate of convergence, i.e., the rate at which the discretization error decreases with respect to the spatial resolution. For many engineering and industry applications, it is commonplace to rely on methods with first- or second-order accuracy, usually in the context of finite difference and finite volume methods. However, these "low-order" accurate schemes typically introduce a significant amount of numerical dissipation, such that they can be exceedingly inefficient for simulating unsteady flow phenomena. For these types of flows, schemes that can provide "high-order" accuracy, which generally refers to higher than second-order, can significantly decrease the necessary computational cost as they typically have lower numerical dissipation and lower error for a given number of spatial degrees of freedom [7]. For the latter, the simulation of complex geometries requires the use of non-uniform meshes which, in many cases, are required to be unstructured and may consist of mixed element types. As such, the numerical scheme must be able to straightforwardly extend to general unstructured meshes to make it usable for practical engineering applications.

While finite difference methods can readily achieve arbitrarily high-order accuracy, they typically cannot be extended to general unstructured grids in an efficient manner. In contrast, finite volume methods are well-suited for unstructured grids, but modifications to increase their order of accuracy are much more involved, resulting in a significantly more algorithmically complex and less computationally efficient approach. With regards to these two desired properties, the use of finite element methods is particularly attractive as they can both achieve the arbitrarily high-order accuracy of finite difference methods while retaining the geometric flexibility of finite volume methods. In this approach, the solution within each element of an arbitrary mesh is represented by a set of basis functions and the governing equations are typically solved in their variational formulation. Various forms of finite element methods exist, the most ubiquitous method being the continuous Galerkin formulation [8] where the discretization is formed such as to enforce the continuity of the basis functions across element interfaces. However, this continuity in continuous Galerkin schemes, as well as in variants such as Petrov–Galerkin schemes [9], typically results in a centered-type scheme for which the lack of numerical dissipation may cause stability issues. Several approaches exist to ameliorate this issue, such as the use of artificial diffusion [10] or alternate formulations such as the streamline upwind Petrov–Galerkin scheme [11]. In addition to the centered nature of continuous Galerkin schemes, another drawback of these schemes is that their computational efficiency can suffer as a result of the lack of compact support in the element-wise solution approximation.

These two disadvantages motivate the use of discontinuous formulations of finite element methods such as discontinuous Galerkin [12, 13], spectral difference [14, 15], and flux reconstruction [16] schemes. These schemes not only retain the desirable properties of standard finite element methods, but also include a natural numerical stabilization mechanism through judicious formulations of the inter-element interactions and have a compact data structure extremely well-suited for modern massively-parallel computing architectures. While the nomenclature is not definitively agreed upon, these schemes are typical referred to as discontinuous spectral element methods (DSEM) when the basis functions are chosen to be polynomials of some high enough degree, typically to where the scheme is high-order accurate. The most commonly used form of DSEM are nodal DSEM, where the basis functions are constructed from nodal interpolating polynomials over a set of discrete solution nodes within each element [13]. This approach is typically the most algorithmically straightforward and computationally efficient class of DSEM, and their use in CFD is the underlying topic of this work. Due to their accuracy and efficiency, these schemes have the potential to allow for the simulation of problems that were hitherto intractable, but these benefits also come with disadvantages that must first be addressed.

1.1 Motivation

The overarching motivation of this work is to develop algorithmic advances for tackling certain deficiencies of DSEM that hinder their widespread use in industrial and scientific contexts. Various discretization techniques for CFD have been applied to a wide variety of flow regimes, but DSEM approaches in their standard form have generally been restricted to simulating subsonic flows at moderate Reynolds numbers [7, 17], shown by the red region in Fig. 1.1. This work will focus primarily on three distinct inadequacies of DSEM which drive this lack of generalizability. First, while DSEM have been shown to be highly accurate and efficient approaches for simulating unsteady flow phenomena in these flow regimes, their ability to simulate high Reynolds number flows is not nearly as well established. Due to their high accuracy and low numerical dissipation, the primary use case for DSEM is typically on directly resolving the flow physics, either through direct numerical simulation (DNS), or more commonly, large eddy simulation (LES). While these approaches can provide a thorough representation of the underlying flow, the computational cost stemming from the spatio-temporal resolution requirements of simulating high Reynolds number flows quickly becomes prohibitive, such that their use is typically limited to lower Reynolds num-

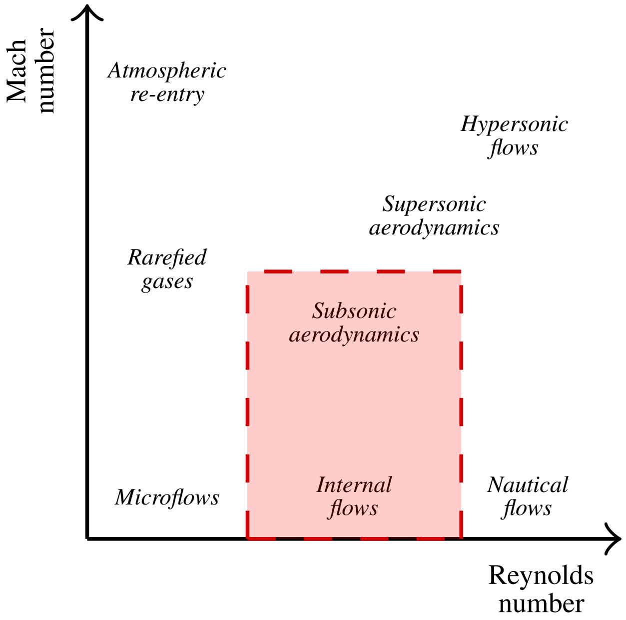


Figure 1.1: Diagram of various flow regimes characterized by the Mach and Reynolds numbers. Red region highlights the regimes standard DSEM approaches can typically be used for.

bers. To extend to higher Reynolds numbers in a tractable manner, it is commonplace to augment the governing equations with some form of a turbulence model to account for the subgrid-scale physics. While the use of DSEM for more fundamental subgrid-scale models such as methods deriving from the steady Reynolds-averaged Navier–Stokes (RANS) equations has been previously explored [18–21], their use for more complex approaches, such as hybrid turbulence models, is quite rare. The pairing of these approaches with DSEM has the potential to significantly reduce the computational cost of simulating high Reynolds number flows, but these models have been primarily applied to low-order FVM schemes [22–24], such that the analysis on the effects of using DSEM for hybrid turbulence modeling is extremely limited. Furthermore, this analysis can be particularly nuanced as these models can be more sensitive to the dissipative properties of the numerical scheme and typically have some dependency on the resolution of the scheme which can be affected by the accuracy of the discretization.

Another deficiency of DSEM stems from a fundamental shortfall of high-order nodal interpolating polynomials. It is well known that this interpolation is ill-behaved when approximating discontinuous features, resulting in spurious oscillations akin to the Gibbs phenomenon. For nonlinear hyperbolic conservation laws, this presents a problem as these systems are known to produce discontinuities in finite time even with smooth initial conditions [25]. In particular, for compressible gas dynamics, it is frequently the case that the system exhibits these features, typically in the form of shock waves and contact discontinuities [26]. In their native form, DSEM are severely ill-suited for the approximation of these types of flows, such that the use of these schemes may result in physical incorrect predictions or, more typically, the failure of the scheme altogether. Around discontinuities, it is usually necessary to apply additional numerical stabilization to these schemes via some form of a shock capturing algorithm. While various shock-capturing approaches exist [10, 27–31], they share a common goal of suppressing numerical instabilities in the vicinity of discontinuities while retaining high-order accuracy in smooth regions of the flow. It is still an open problem as to how to formulate robust and accurate shock capturing methods for DSEM as many existing approaches are either excessively dissipative, may degrade the accuracy of the scheme in smooth regions, do not guarantee the robustness of the scheme around strong discontinuities, and/or require problem-dependent tunable parameters.

Finally, as its application is not as widespread as finite difference and finite volume methods, the use of DSEM for more complex flow regimes is not nearly as common. One such example is the use of DSEM for non-equilibrium flows, such as rarefied gases, hypersonic flows, and microfluids, where the assumption of continuum or thermodynamic equilibrium is not well posed. In these regimes, governing equations stemming from the continuum assumption may produce erroneous results, and it is typically necessary to revert to the governing equations of molecular gas dynamics derived from the kinetic theory of gases [32]. One such statistical description of molecular gas dynamics can be given by the Boltzmann equation which characterizes the transport and interaction of particles. Although the Boltzmann equation is relatively simple from a mathematical perspective, its high-dimensional nature makes it exceedingly costly to simulate. As a result, direct simulation of the Boltzmann equation for non-trivial fluid flows has been considered out of reach due to this complexity. The use of DSEM has the potential to reduce this computational cost and allow for direct simulation of complex flows via the Boltzmann equation that were previously intractable. However, DSEM approximations of the Boltzmann equation have only recently been attempted [33, 34] and only for simple two-dimensional flows, such that there is significant room for improvement in the methodology and analysis as well as potential application to more complex three-dimensional flows.

1.2 Overview

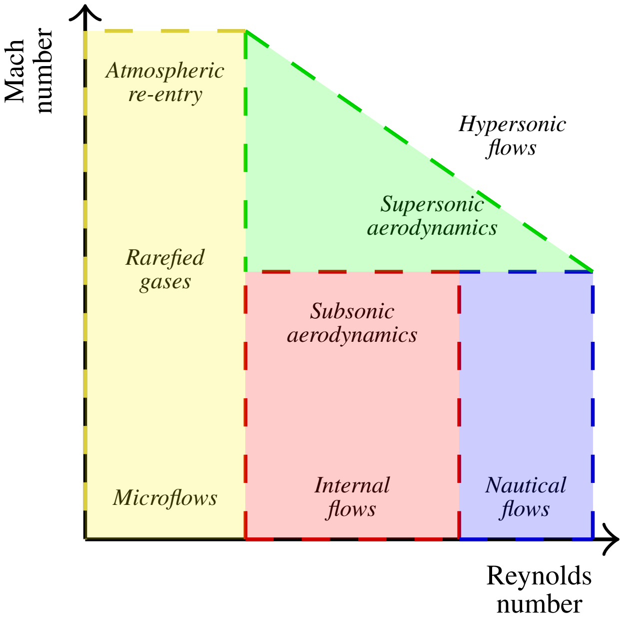


Figure 1.2: Diagram showing the flow regimes made accessible by this work through hybrid turbulence modeling (blue), advances in shock capturing (green), and discretizations of non-equilibrium systems (yellow).

This work consists of three chapters individually addressing the three previously presented deficiencies of DSEM, performed in the context of the high-order flux reconstruction scheme of Huynh

[16] implemented within the PyFR solver [35], a framework that can target massively-parallel CPU and GPU computing architectures. The work in these chapters will expand the applicability of DSEM to a wider variety of flow regimes, shown by the additional highlighted regions in Fig. 1.2. First, the applicability of hybrid turbulence modeling within a high-order DSEM framework is explored through scale-resolving simulations of turbulence via the Partially-averaged Navier–Stokes (PANS) approach of Girimaji [22]. Modifications for the robust implementation of this model within a high-order framework are presented, and the effects of discretization order on this model are analyzed by observing the relative accuracy benefits of low-order versus high-order methods in comparison to approaches without any explicit sub-grid scale modeling. These comparisons are performed over the wall-bounded, separated flow around a periodic hill and the wake flow around a circular cylinder at a Reynolds number of 3900.

Next, an adaptive filtering approach is introduced as a shock-capturing approach for DSEM approximations of hyperbolic and mixed hyperbolic-parabolic conservation laws. By adapting the filter strength to enforce positivity of density and pressure as well as a local discrete minimum entropy principle, the resulting approach can robustly resolve strong discontinuities with sub-element resolution, does not require problem-dependent parameter tuning, and can be easily implemented on general unstructured meshes with relatively low computational cost. The efficacy of the approach is shown in numerical experiments on the Euler and Navier–Stokes equations for problems including extreme shocks, shock-vortex interactions, and complex compressible turbulent flows.

Finally, a positivity-preserving and discretely conservative method for solving the Boltzmann equation closed with the Bhatnagar-Gross-Krook [36] collision model is presented. Through modeling the internal degrees of freedom, the approach is further extended to polyatomic molecules and can encompass arbitrary constitutive laws. The approach is validated on a series of large-scale complex numerical experiments, ranging from shock-dominated flows computed on unstructured grids to direct numerical simulation of three-dimensional compressible turbulent flows, the latter of which is the first instance of such a flow computed by directly solving the Boltzmann equation.

The organization of this work is as follows: the remainder of this chapter presents some preliminaries including an overview of hyperbolic and mixed hyperbolic-parabolic conservation laws and a description of the flux reconstruction approach. The implementation and analysis of the PANS approach within a high-order DSEM framework are then presented in Chapter 2. Next, the positivity-preserving entropy-based adaptive filtering approach is introduced in Chapter 3. Finally, a robust and accurate numerical approach for solving the polyatomic Boltzmann equation is presented in Chapter 4, followed by conclusions drawn in Chapter 5.

1.3 Preliminaries

1.3.1 Conservation Laws

This work primarily pertains to approximations of hyperbolic conservation laws or mixed hyperbolic-parabolic conservation laws of the form

$$\begin{cases} \partial_t \mathbf{u} + \boldsymbol{\nabla}\cdot\mathbf{F}(\mathbf{u}) = 0, \\ \mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_0(\mathbf{x}), \qquad \text{for } \mathbf{x} \in \Omega, \end{cases} \tag{1.1}$$

where $\mathbf{u} \in \mathbb{R}^m$ is a vector-valued solution, m is an arbitrary number of field variables, $\mathbf{F}(\mathbf{u}) \in \mathbb{R}^{m \times d}$ is the flux, $\Omega \in \mathbb{R}^d$ is the domain, and d is an arbitrary number of spatial dimensions. For mixed hyperbolic-parabolic systems, it is convenient to split the flux into two components as

$$\mathbf{F}(\mathbf{u}) = \mathbf{F}_I(\mathbf{u}) + \mathbf{F}_V(\mathbf{u}), \tag{1.2}$$

where the subscript I denotes the hyperbolic component and the subscript V denotes the parabolic component. This notation stems from the governing equations of compressible fluid flow which consists of an inviscid hyperbolic component and a viscous parabolic component. Note that the parabolic component of the conservation law is second-order with respect to the solution, i.e., $\mathbf{F}_V = \mathbf{F}_V(\mathbf{u}, \nabla\mathbf{u})$.

For the hyperbolic system, the focus of this work is primarily on the compressible Euler equations, given in the form of Eq. (1.1) as

$$\mathbf{u} = \begin{bmatrix} \rho \\ \boldsymbol{\rho}\mathbf{V} \\ E \end{bmatrix} \quad \text{and} \quad \mathbf{F} = \begin{bmatrix} \boldsymbol{\rho}\mathbf{V} \\ \boldsymbol{\rho}\mathbf{V} \otimes \mathbf{V} + P\mathbf{I} \\ (E+P)\mathbf{v} \end{bmatrix}, \tag{1.3}$$

where ρ is the density, $\boldsymbol{\rho}\mathbf{V}$ is the momentum, E is the total energy, $P = (\gamma - 1)\left(E - {}^{1}\!/_{2}\rho\mathbf{V}\cdot\mathbf{V}\right)$ is the pressure, and γ is the ratio of specific heat capacities. The symbol $\mathbf{I}$ denotes the identity matrix in $\mathbb{R}^{d\times d}$ and $\mathbf{V} = \boldsymbol{\rho V}/\rho$ denotes the velocity.

For the mixed hyperbolic-parabolic system, the focus of this work is primarily on the compressible Navier–Stokes equations. For this system, the inviscid (hyperbolic) and viscous (parabolic) components of the flux, denoted by the subscripts I and V, respectively, are given as

$$\mathbf{u} = \begin{bmatrix} \rho \\ \boldsymbol{\rho v} \\ E \end{bmatrix}, \quad \mathbf{F}_I = \begin{bmatrix} \boldsymbol{\rho V} \\ \boldsymbol{\rho V} \otimes \mathbf{V} + P\mathbf{I} \\ (E+P)\mathbf{v} \end{bmatrix}, \quad \text{and} \quad \mathbf{F}_V = \begin{bmatrix} 0 \\ -\mu\left(\nabla\mathbf{V} + \nabla\mathbf{V}^T\right) + \frac{2}{3}\mu\nabla\cdot\mathbf{V} \\ -\mu\left(\nabla\mathbf{V} + \nabla\mathbf{V}^T\right)\mathbf{v} - \mu\frac{\gamma}{Pr}\nabla e \end{bmatrix}, \tag{1.4}$$

where μ is the dynamic viscosity, Pr is the Prandtl number, and $e = \rho^{-1}(E - {}^{1}\!/_{2}\rho\mathbf{V}\cdot\mathbf{V})$ is the specific internal energy.

1.3.2 Flux Reconstruction Scheme

The flux reconstruction (FR) scheme, introduced by Huynh [16], is a form of nodal DSEM that can be considered as a generalization of the nodal discontinuous Galerkin method [13], encompassing both its standard form and other filtered variants. In this approach, the spatial domain Ω is partitioned into N_e elements Ω_k such that $\Omega = \bigcup_{N_e}\Omega_k$ and $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$. With a slight abuse of notation, the solution $\mathbf{u}(\mathbf{x})$ within each element Ω_k is approximated by a polynomial of the form

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{N_s} \mathbf{u}(\mathbf{x}_i^s)\phi_i(\mathbf{x}), \tag{1.5}$$

where $\mathbf{x}_i^s \in \Omega_k \; \forall \; i \in \{1, ..., N_s\}$ is a set of N_s solution nodes and $\phi_i(\mathbf{x})$ is a set of nodal basis functions with the property $\phi_i(\mathbf{x}_j^s) = \delta_{ij}$. The order of the approximation, represented by $\mathbb{P}_p$ for some order p, is defined as the maximal order of $\mathbf{u}(\mathbf{x})$.

Starting with the hyperbolic component, a discontinuous approximation of the flux is formed via a collocation projection of the flux onto the space spanned by the solution, i.e.,

$$\mathbf{f}^D(\mathbf{x}) = \sum_{i=1}^{N_s} \mathbf{F}_I \left(\mathbf{u}(\mathbf{x}_i^s)\right) \phi_i(\mathbf{x}), \tag{1.6}$$

To allow communication between elements and to ensure consistency between the polynomial spaces of the solution and the divergence of the flux, the corrected flux is formed by amending the discontinuous flux with additional correction terms which enforce C^0 continuity across $\partial\Omega_k$.

$$\mathbf{f}^C(\mathbf{x}) = \mathbf{f}^D(\mathbf{x}) + \sum_{i=1}^{N_f} \left[\overline{\mathbf{F}}_i - \mathbf{f}^D(\mathbf{x}_i^f) \cdot \mathbf{n}_i\right] \mathbf{h}_i(\mathbf{x}). \tag{1.7}$$

Here, $\mathbf{x}_i^f \in \partial\Omega_k \; \forall \; i \in \{1, ..., N_f\}$ is a set of N_f interface flux nodes, $\mathbf{n}_i$ is their associated outward-facing normal vector, $\overline{\mathbf{F}}_i$ is a common hyperbolic interface flux yet to be defined, and $\mathbf{h}_i$ is the correction function associated with the given flux node. These correction functions have the properties that

$$\mathbf{n}_i \cdot \mathbf{h}_j(\mathbf{x}_i^f) = \delta_{ij} \quad \text{and} \quad \sum_{i=1}^{N_f} \mathbf{h}_i(\mathbf{x}) \in RT_p, \tag{1.8}$$

where RT_p is the Raviart–Thomas space of order p [37]. In this work, these correction functions are chosen such as to recover the nodal discontinuous Galerkin approach [13, 16, 38], e.g., for tensor-product elements, the correction functions are the Radau polynomials of degree $p+1$.

At the interfaces, it is necessary to define a common hyperbolic interface flux. Due to the discontinuous nature of the solution approximation in DSEM, at each interface flux node, there exist two values of the solution, the solution value from the element in question, denoted by $\mathbf{u}^-$, and the solution value from the face-neighboring element, denoted by $\mathbf{u}^+$. To ensure conservation and the correct direction of information propagation, the common flux is computed from these two values by treating the interface as a Riemann problem. Various methods for solving the Riemann problem exist, ranging from exact Riemann solvers such as Godunov-type methods [39] to approximate Riemann solvers such as the methods of Rusanov [40] and Roe [41]. For example, the common interface flux can be computed via the Rusanov approximate Riemann solver as

$$\overline{\mathbf{F}} = \frac{1}{2}\left(\mathbf{f}^- + \mathbf{f}^+\right) \cdot \mathbf{n} - \frac{1}{2}\lambda_{\text{max}}\left(\mathbf{u}^+ - \mathbf{u}^-\right), \tag{1.9}$$

where λ_{max} is a suitable estimate of the maximum propagation speed in the system. One such estimate is given by Davis [42] as

$$\lambda_{\text{max}} = \max\left(\left|\mathbf{v}^- \cdot \mathbf{n}\right| + c_s^-, \left|\mathbf{v}^+ \cdot \mathbf{n}\right| + c_s^+\right), \tag{1.10}$$

where $c_s = \sqrt{\gamma P/\rho}$ is the local speed of sound.

For the parabolic component, it is necessary first to form an appropriate approximation of the gradient of the solution, represented as

$$\mathbf{w} \approx \nabla \mathbf{u}. \tag{1.11}$$

Similarly to the calculation of the hyperbolic component of the flux, a discontinuous approximation of the gradient of the solution can be computed as

$$\mathbf{w}^D(\mathbf{x}) = \sum_{i=1}^{N_s} \mathbf{u}(\mathbf{x}_i^s)\nabla\phi_i(\mathbf{x}), \tag{1.12}$$

after which a C^0 continuous approximation of the solution gradient can be computed as

$$\mathbf{w}^C(\mathbf{x}) = \mathbf{w}^D(\mathbf{x}) + \sum_{i=1}^{N_f}\left[\overline{\mathbf{u}}_i - \mathbf{u}(\mathbf{x}_i^f)\right]\nabla\mathbf{h}_i(\mathbf{x}). \tag{1.13}$$

The common interface solution $\overline{\mathbf{u}}_i$ can be computed using a simple centered average or penalty-like methods such as the BR2 approach of Bassi and Rebay [43]. The discontinuous approximation of the parabolic component of the flux is then computed as

$$\mathbf{g}^D(\mathbf{x}) = \sum_{i=1}^{N_s} \mathbf{F}_V\left(\mathbf{u}(\mathbf{x}_i^s), \mathbf{w}^C(\mathbf{x}_i^s)\right)\phi_i(\mathbf{x}), \tag{1.14}$$

after which adding the correction terms yields

$$\mathbf{g}^C(\mathbf{x}) = \mathbf{g}^D(\mathbf{x}) + \sum_{i=1}^{N_f}\left[\overline{\mathbf{G}}_i - \mathbf{g}^D(\mathbf{x}_i^f)\cdot\mathbf{n}_i\right]\mathbf{h}_i(\mathbf{x}). \tag{1.15}$$

Similarly to the common interface solution, the common parabolic interface flux can be computed using a centered average or through penalty-like methods. The temporal derivative of the solution can then be approximated as

$$\partial_t\mathbf{u} = -\nabla\cdot\mathbf{f}^C(\mathbf{x}) \tag{1.16}$$

for hyperbolic systems and

$$\partial_t\mathbf{u} = -\nabla\cdot\mathbf{f}^C(\mathbf{x}) - \nabla\cdot\mathbf{g}^C(\mathbf{x}) \tag{1.17}$$

for mixed hyperbolic-parabolic systems.

2. PARTIALLY-AVERAGED NAVIER–STOKES SIMULATIONS OF TURBULENCE

For a variety of practical engineering flows, the spatio-temporal resolution requirements of the underlying physical phenomena of complex turbulent flows make DNS and LES prohibitively expensive. Progress in this regard has generally followed two distinct paths: algorithmic advances and the development of higher fidelity subgrid-scale models. While the former approach attempts to reduce the computational cost, the goal of the latter is to reduce the resolution requirements without a significant detriment in accuracy.

In the context of algorithmic design, a class of methods that have seen increased usage for high-fidelity simulations of turbulence over the past several decades are the aforementioned DSEM as they can significantly decrease the computational cost requirements of turbulent flows. However, their application has generally focused on LES and DNS [44], and as such, they suffer from the spatio-temporal resolution requirements of these techniques which prohibits their use for higher Reynolds number flows. The computationally unfeasible resolution requirements of complex flows have driven the development of higher fidelity subgrid-scale models. There exist a large variety of options in terms of techniques for simulating fluid flows. On one end of the spectrum, approaches such as RANS attempt to model all of the spatio-temporal scales of the flow, doing so with relatively little computational effort. However, the effectiveness of the approach is highly dependent on the problem and the model in question. On the other end, DNS attempts to resolve all of the spatio-temporal scales of the flow at the expense of tremendous computational cost. This cost is slightly alleviated through approaches such as LES in which only the statistically significant scales are resolved, either via an explicit approach using a filter with a subgrid-scale model or via an implicit approach using the numerical dissipation of the scheme, the latter of which is typically denoted as implicit LES (ILES) or under-resolved DNS (URDNS). However, the cost of these approaches generally makes them impractical for engineering applications, and as such, there is a necessity for methods that can offer higher fidelity than RANS at a lower cost than LES. These methods are typically denoted as scale-resolving simulation (SRS) techniques, which attempt to relax the resolution requirements of LES without sacrificing its ability to accurately resolve the predominant flow physics. One such class of techniques that have shown promise in this regard are hybrid turbulence models [22, 24, 45–47]. The intent of these SRS approaches is generally to resolve only the coherent flow structures while modeling the stochastic portion of the flow. As a result, their computational cost tends to scale much more reasonably with respect to flow complexity while potentially retaining many of the benefits of approaches such as LES in terms of accuracy.

The unification of high-order numerical methods and hybrid turbulence models provides a potential for significant improvements in the computational cost of SRS. Aside from the increased fidelity provided by high-order methods, the decrease in numerical dissipation can particularly benefit SRS methods as the excessive dissipation of low-order schemes has a much more detrimental impact on the modeled physics in comparison to RANS-type approaches. The application of high-order numerical methods to various turbulence models has been explored in the literature [18–21], with the majority of these works using high-order schemes for RANS or zonal methods such as detached eddy simulation (DES). However, the analysis on the actual effects of the high-order discretization on these turbulence models is extremely limited, particularly for SRS with hybrid turbulence models in which the numerical dissipation plays a more significant role and the accuracy of the discretization affects the resolution, making this analysis more complex.

The goal of this chapter is therefore to analyze the effects of high-order discretizations on a hybrid turbulence model through a single numerical framework that can recover an arbitrary order of accuracy. For this, we employ the PANS approach of Girimaji [22], a bridging turbulence model, paired with the FR scheme. We study the effects of discretization order on the optimal resolution-control parameters, the relative accuracy benefits compared to approaches without physical models, and how the discretization affects the ability of the model to predict the dominant flow physics and flow structures. This analysis is performed on a wall-bounded, separated flow as well as the wake flow around a circular cylinder at a Reynolds number of 3900, the latter of which we provide DNS results and novel statistical analysis. The remainder of this chapter is structured as follows. In Section 2.1, the PANS is presented, as well as implementation details and modifications to make the model more amenable to high-order discretizations. The numerical experiments are described in Section 2.2 followed by the results. Conclusions are then drawn in Section 2.3.

2.1 Methodology

2.1.1 Partially-averaged Navier–Stokes Formulation

The PANS approach is derived from the application of a filter of arbitrary width to the Navier–Stokes equations [22]. For the compressible form, this yields second moment and source terms in the resolved momentum and energy equations, which allow the governing PANS equations to be written in the form of Eq. (1.1) with an additional flux contribution $\mathbf{F}_T(\mathbf{u})$ as

$$\mathbf{F}(\mathbf{u}) = \mathbf{F}_I(\mathbf{u}) + \mathbf{F}_V(\mathbf{u}) + \mathbf{F}_T(\mathbf{u}). \tag{2.1}$$

By invoking the Boussinesq approximation and enforcing the conservation of unresolved turbulent kinetic energy, these terms can be written as

$$\mathbf{F}_T(\mathbf{u}) = \begin{bmatrix} \mathbf{0} \\ -\mu_u \left(\nabla \mathbf{V} + \nabla \mathbf{V}^T\right) + \frac{2}{3} k_u \mathbf{I} \\ -\mu_u \frac{\gamma}{Pr_t} \nabla e + k_u \end{bmatrix}, \tag{2.2}$$

where μ_u is the unresolved eddy viscosity, k_u is the unresolved turbulent kinetic energy, and Pr_t is the turbulent Prandtl number. These relations create two additional unknowns, μ_u and k_u, which are determined through a modification of an underlying RANS turbulence model. In this work, the k-ω Shear Stress Transport (SST) model of Menter [48] is used. This model is adapted to the PANS formulation by reformulating the turbulence variables into their unresolved components, defined through the relations

$$f_k = \frac{k_u}{k} \quad \text{and} \quad f_\omega = \frac{\omega_u}{\omega}, \tag{2.3}$$

where the subscript $(\cdot)_u$ denotes the unresolved (i.e., modeled) component of the variable $(\cdot)$ and $f_{(\cdot)}$ denotes the unresolved-to-total ratio of $(\cdot)$. This formulation provides a mechanism for the

PANS approach to seamlessly blend between DNS ($f_k = f_\omega = 0$) and URANS ($f_k = f_\omega = 1$). The modified transport equations for the unresolved turbulent kinetic energy k_u and unresolved specific dissipation ω_u are given as

$$\partial_t \rho k_u + \boldsymbol{\nabla}\cdot \left[\rho k_u \mathbf{V} - \left(\mu + \mu_u \sigma_k \frac{f_\omega}{f_k}\right) \nabla k_u \right] = P_k - \beta^* \rho k_u \omega_u, \tag{2.4}$$

and

$$\begin{aligned} \partial_t \rho \omega_u + \boldsymbol{\nabla}\cdot &\left[\rho \omega_u \mathbf{V} - \left(\mu + \mu_u \sigma_\omega \frac{f_\omega}{f_k}\right) \nabla \omega_u \right] \\ &= \frac{\alpha}{\mu_u}\rho P_k - \rho\left(P' - \frac{P'}{f_\omega} + \frac{\beta \omega_u}{f_\omega}\right) + 2\rho \frac{\sigma_{\omega 2}}{\omega_u}\frac{f_\omega}{f_k}(1 - F_1)\nabla k_u \cdot \nabla \omega_u, \end{aligned} \tag{2.5}$$

where

$$P_k = \left[\mu\left(\nabla \mathbf{V} + \nabla \mathbf{V}^T\right) + \frac{2}{3}\mu(\boldsymbol{\nabla}\cdot\mathbf{V})\right]\cdot \nabla \mathbf{V} \tag{2.6}$$

and $P' = \rho\alpha\beta^* k_u/\mu_u$. The eddy viscosity is calculated as

$$\mu_u = \frac{\rho \alpha_1 k_u}{\max(\alpha_1 \omega_u, \Omega F_2)}, \tag{2.7}$$

with the auxiliary functions defined as

$$F_1 = \tanh\left(\min\left(\max\left(\frac{\sqrt{k_u}}{\beta^* \omega_u d}, \frac{500\mu}{\rho \omega_u d^2}\right), \frac{4\rho\sigma_{\omega 2} k_u}{CD_{k\omega} d^2}\right)^4 \right), \tag{2.8}$$

$$F_2 = \tanh\left(\max\left(\frac{2\sqrt{k_u}}{\beta^* \omega_u d}, \frac{500\mu}{\rho \omega_u d^2}\right)^2 \right), \tag{2.9}$$

$$CD_{k\omega} = \max\left(\frac{2\rho\sigma_{\omega 2}}{\omega_u} \nabla k_u \cdot \nabla \omega_u, 10^{-10}\right), \tag{2.10}$$

$$\Omega = \|\nabla \mathbf{V} - \nabla \mathbf{V}^T\|_2, \tag{2.11}$$

for some wall distance d. The values of the free parameters are tabulated in Chapter A, and description of the modifications applied to the PANS formulation to make it more amenable to high-order discretizations is presented in Section 2.1.2.

The effectiveness of the PANS approach is ultimately reliant on the values of the resolution-control parameters f_k and f_ω. In this work, these parameters are kept constant in space and time as the presence of spatio-temporal variation in the parameters requires additional modifications in the governing equations to account for commutation errors. Furthermore, the two-parameter model (f_k, f_ω) is reduced to a one-parameter model (f_k) through the assumption that the length scales associated with turbulent dissipation are entirely unresolved (i.e., $f_\epsilon = 1$), an assumption proposed in Girimaji and Wallin [49] and explored in Pereira et al. [23]. Subsequently, the one-parameter model is closed through the relation $f_\omega = f_\epsilon / f_k = 1/f_k$.

2.1.2 Numerical Implementation

Due to the inherently low numerical dissipation in high-order discretizations, several modifications were applied to the PANS equations to ensure robustness. In the context of the transport equations for the turbulence variables, ensuring that these variables remain positive requires adapting the formulation. For the ω_u transport equation, where the variable is *strictly positive*, we take the approach of Ilinca and Pelletier [18] as described by Bassi et al. [19], in which the transport equation for $\tilde{\omega}_u = \log(\omega_u)$ is solved instead.

$$\begin{aligned}\partial_t \rho\tilde{\omega}_u + \boldsymbol{\nabla}\cdot\left[\rho\tilde{\omega}_u \mathbf{V} - \left(\mu + \mu_u \sigma_\omega \frac{f_\omega}{f_k}\right)\nabla\tilde{\omega}_u\right] \\ = \frac{\alpha}{\mu_u}\rho P_k - \rho\left(P' - \frac{P'}{f_\omega} + \frac{\beta e^{\tilde{\omega}_u}}{f_\omega}\right) + 2\rho\frac{\sigma_{\omega 2}}{e^{\tilde{\omega}_u}}\frac{f_\omega}{f_k}(1 - F_1)\nabla k_u \cdot \nabla\tilde{\omega}_u, \quad (2.12)\end{aligned}$$

As a result, the specific dissipation term only appears in exponential form in the transport and auxiliary equations, guaranteeing positivity. Furthermore, the distribution of the logarithm form of ω_u is smoother than that of ω_u itself [19]. However, for k_u, where the variable is only *non-negative*, the logarithm form is not necessarily well-defined. In contrast to the approach of Bassi et al. [19] where negative values of k were allowed in the solution but limited to zero in the transport and auxiliary equations, we introduce a source term S_k in the k_u transport equation which activates if

k_u falls below some small constant k_{min}, such that k_u is effectively limited to k_{min}. This source term was formed via a forward Euler approximation of k_u, given as

$$S_k^n = \max\left[0,\ \frac{k_{min} - k_u^n}{\Delta t} + \boldsymbol{\nabla} \cdot \mathbf{F}_k^n\right], \tag{2.13}$$

where $\mathbf{F}_k$ is the k_u component of the flux and the superscript n denotes the time step. The value of k_{min} was set to 10^{-8}. The equations were discretized using a Rusanov-type [40] Riemann solver for the inviscid fluxes, the BR2 method of Bassi and Rebay [43] for the viscous fluxes, and an explicit fourth-order Runge–Kutta scheme for temporal integration. In the numerical experiments, a comparison between a low-order and high-order FR approach was conducted. The low-order approach was performed using $\mathbb{P}_1$ solution polynomials, resulting in a second-order accurate scheme comparable to a finite volume formulation. For the high-order approach, $\mathbb{P}_3$ solution polynomials were used, resulting in a fourth-order accurate scheme. Due to the increase in resolution afforded through higher-order representations of the solution, the meshes used for the low-order and high-order approaches were coarsened/refined appropriately such that the total degrees of freedom remained the same (i.e., the $\mathbb{P}_1$ meshes used approximately 6-8 times as many elements as the $\mathbb{P}_3$ meshes).

2.2 Numerical Experiments

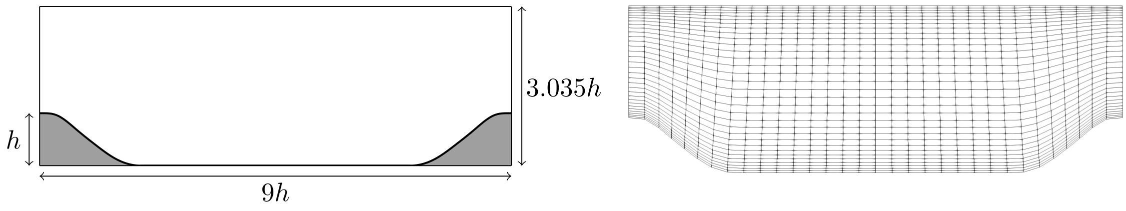
The effects of a high-order discretization of the PANS equations were evaluated on two distinct numerical experiments. The first test case, the periodic hill problem of Fröhlich et al. [50], consists of a recirculating, wall-bounded flow through a channel with periodic constrictions, a canonical benchmark for computing separated flows. The second test case, the flow around a circular cylinder at a Reynolds number of 3900, serves as an assessment of the methods for problems with significantly more complex flow physics, such as laminar separation, free-shear transition, and turbulent wakes. These assessments were performed at under-resolved to moderately-resolved levels of spatio-temporal resolution, where first-order statistics can be predicted reasonably well but second-order statistics and/or dominant flow physics may be poorly predicted. This range of resolution is where the application of hybrid turbulence models is most practical, as accurate predictions of the flow physics are achievable while still relaxing the resolution requirements for LES. As a result, the values of the free parameter f_k were investigated over a range of 0.1 to 0.3. Furthermore, the PANS approach was compared to a URDNS approach for which the same mesh and numerical setup were used without the addition of the PANS model to give a comparable assessment of the effects of the model. The focus of these comparisons is towards metrics that are more difficult to resolve, such as second-order statistics, temporal properties of the flow physics, and the characteristics of the coherent structures in the flow. We reiterate that the goal of this investigation is not to evaluate the efficacy of the PANS method, but to show the effect of high-order discretizations for SRS using PANS and how that effect differs from approaches without any physical models.

2.2.1 Periodic Hill

The periodic hill problem of Fröhlich et al. [50] presents a general assessment for predicting flow separation arising from curved surfaces and the subsequent flow reattachment. The problem consists of a channel flow with an infinite series of smooth constrictions (hills) of height h separated by a crest-to-crest distance of $9h$. In numerical experiments, the infinite domain is approximated with a truncated region spanning one crest-to-crest distance with periodicity enforced between the inlet and outlet. A 2D cross-section of the model geometry in the streamwise-spanwise plane is shown in Fig. 2.1 for this truncated region. The 3D geometry is formed by extruding the cross-section along the spanwise direction over a length of $4.5h$.

Initially, a uniform flow with a Mach number of 0.1 is set throughout the domain, and the bulk Reynolds number, based on the bulk velocity U_b and height h, is set to 10,595. To enforce a constant mass flow rate across the domain, a body force in the form of an additional streamwise pressure gradient was imposed using the approach of Benocci and Pinelli [51] as described by Wang [52], given as

$$\left(\frac{dP}{dx}\right)^{n+1} = \left(\frac{dP}{dx}\right)^{n} - \frac{1}{A_c \Delta t}(\dot{m}^* - 2\dot{m}^n + \dot{m}^{n-1}), \tag{2.14}$$



(a) Geometry (b) Mesh

Figure 2.1: Cross-section of the periodic hill geometry in the streamwise-spanwise plane (left) and mesh for the $\mathbb{P}_1$ approach (right).

where the superscript denotes the time step iteration, A_c is the inlet/outlet area, and $\dot{m}$ is the mass flow rate at the inlet/outlet. The desired mass flow rate $\dot{m}^*$ was set based on the bulk Reynolds number, and the initial pressure gradient was set to zero.

Two structured, hexahedral meshes were generated for the periodic hill, one for the low-order ($\mathbb{P}_1$) approach and one for the high-order ($\mathbb{P}_3$) approach. The low-order mesh was generated by uniformly sub-dividing the high-order mesh along each direction. A description of the characteristics of these meshes is shown in Table 2.1, where N denotes the number of elements, DOF denotes the degrees of freedom, and x, y, and z denote the streamwise, normal, and spanwise directions, respectively. The Δy^+ value was estimated by normalizing the distance of the first solution point to the wall by the friction velocity of a flat plate at the bulk Reynolds number. The difference in Δy^+ values between the grids results from the nonuniformity of the solution point distribution at various orders. To highlight the differences in the predicted flow properties between the low-order and high-order approaches, significantly coarser grids were used in the present work, using 97% fewer degrees of freedom than the LES of Fröhlich et al. [50] and 85% fewer degrees of freedom than the low-order PANS approach of Razi et al. [53]. As such, the results are presented for f_k values of 0.2 and 0.3, as through *a posteriori* analysis, it was shown that this resolution level does not support lower f_k values.

Periodic boundary conditions were imposed between the inlet and outlet as well as along the spanwise direction. At both the top and bottom walls, no-slip, adiabatic boundary conditions were

Grid	Method	Order	$N_x \times N_y \times N_z$	$DOF_x \times DOF_y \times DOF_z$	Δy^+
A1	PANS/URDNS	$\mathbb{P}_1$	$32 \times 32 \times 16$	$64 \times 64 \times 32$	1.8
A2	PANS/URDNS	$\mathbb{P}_3$	$16 \times 16 \times 8$	$64 \times 64 \times 32$	1.0

Table 2.1: Description of mesh characteristics for the periodic hill case.

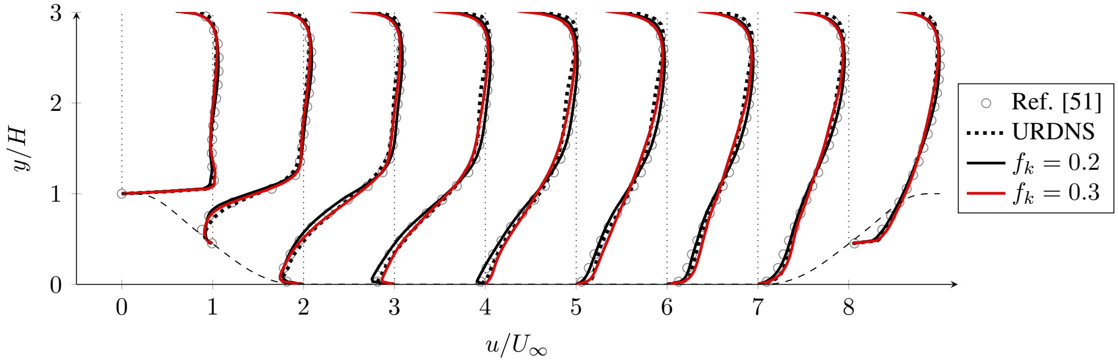
applied. Following the recommendation of Menter [48], the boundary conditions at the wall for the turbulence variables were set as

$$k_u = 0, \qquad \omega_u = f_\omega \frac{60\nu}{\beta_1 (\Delta d_1)^2}, \tag{2.15}$$

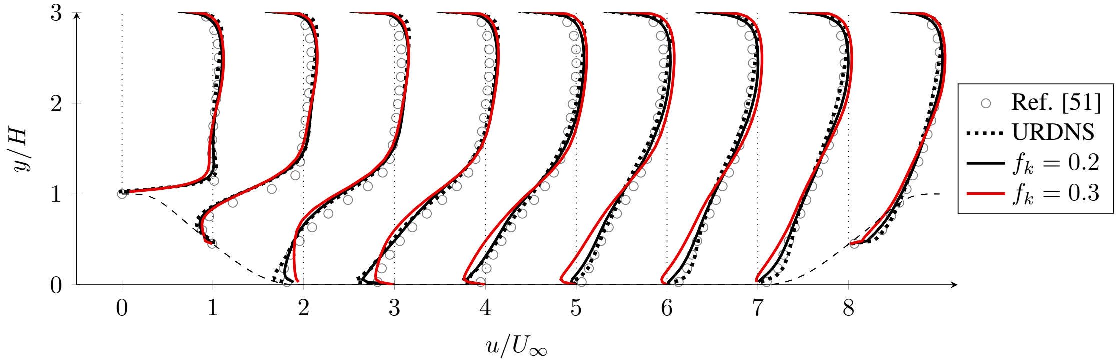
where β_1 is a constant defined in Chapter A and Δd_1 is the wall distance. To maintain consistency in the boundary conditions between low-order and high-order approaches in light of the nonuniformity of the solution point distribution, the wall distance was approximated as $\Delta d_1 = \Delta d_e/(p+1)$, where Δd_e is the height of the element and p is the order of the solution basis. As such, the values of the turbulence variables at the wall were identical between both approaches. The simulation was run for a period corresponding to 20 flow-through times of the domain. After 10 flow-through times, the flow was assumed to be fully-developed. Statistical quantities were then gathered over the final 10 flow-through times.

After averaging the flow across the time-averaging horizon and along the spanwise direction, the first-order and second-order statistics of the periodic hill flow were analyzed. These quantities were compared to the LES results of Fröhlich et al. [50]. A comparison of the profiles of the average streamwise velocity as predicted by the $\mathbb{P}_1$ and $\mathbb{P}_3$ methods is shown in Fig. 2.2. For both methods, the separation of the flow aft of the hill was evident, but discrepancies in the profiles in the separation region were observed. For the low-order method, both the PANS and URDNS approaches overpredicted the reversal of the flow in the separation region, which resulted in notable deviations from the reference data in the region $0 \leq x \leq 5$. The introduction of the PANS model did not significantly improve the low-order results. However, for the high-order scheme, both the

URDNS and PANS approaches performed significantly better, and the $f_k = 0.2$ results were in excellent agreement with the reference data. Less deviation in the mean velocity profiles between the two values of f_k were observed with the high-order scheme, indicating that the high-order approach is less sensitive to the application of the PANS model than the low-order approach.



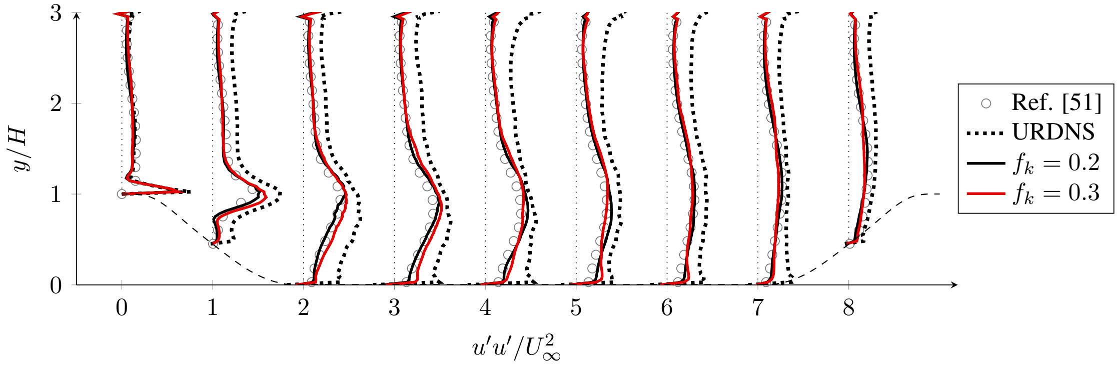
(a) $\mathbb{P}_3$



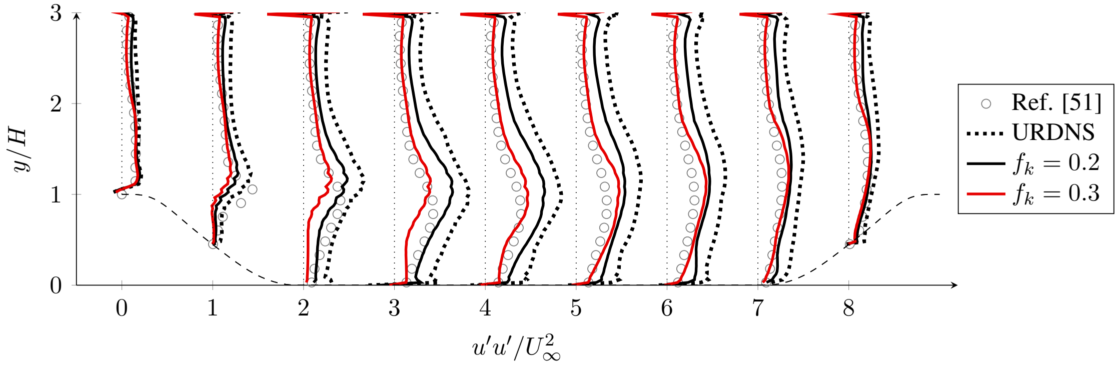
(b) $\mathbb{P}_1$

Figure 2.2: Time and span-averaged streamwise velocity profiles using a $\mathbb{P}_3$ (top) and $\mathbb{P}_1$ FR scheme (bottom). Profiles are shifted by 0, +1, +2, ..., +8, respectively, along the abscissa.

For the second-order statistics, the differences in the predictions by the various approaches were much more evident. In Fig. 2.3, the profiles of the streamwise velocity variance are shown. For both the low-order and high-order methods, the URDNS approach significantly overpredicted the variance across most of the domain, with the largest overprediction seen in the separation region. The change from the low-order to high-order method for the URDNS approach did not



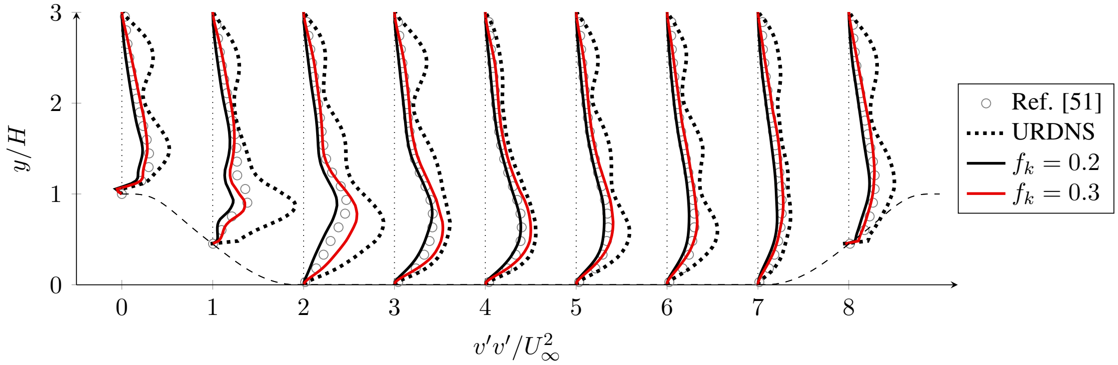
(a) $\mathbb{P}_3$



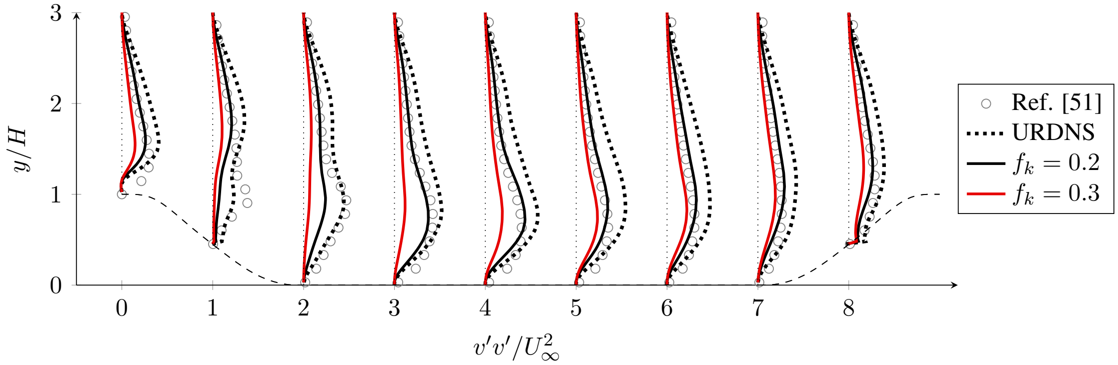
(b) $\mathbb{P}_1$

Figure 2.3: Time and span-averaged streamwise velocity variance profiles using a $\mathbb{P}_3$ (top) and $\mathbb{P}_1$ FR scheme (bottom). Profiles are scaled by a factor of 5 and shifted by 0, +1, +2, ..., +8, respectively, along the abscissa.

appreciably improve the results. However, unlike URDNS, notable improvement was seen with the PANS approach when switching from the low-order method to the high-order method. Excellent agreement was observed between the high-order PANS approach and the reference data for both values of f_k, whereas the accuracy of the low-order PANS approach was comparable to the URDNS approaches. As with the first-order statistics, significantly less deviation in the results was observed between the two f_k values when using the high-order scheme.



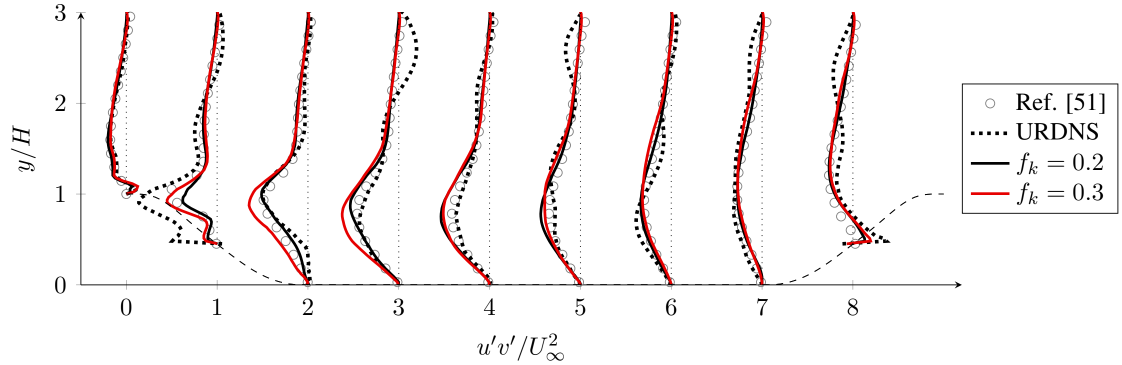
(a) $\mathbb{P}_3$



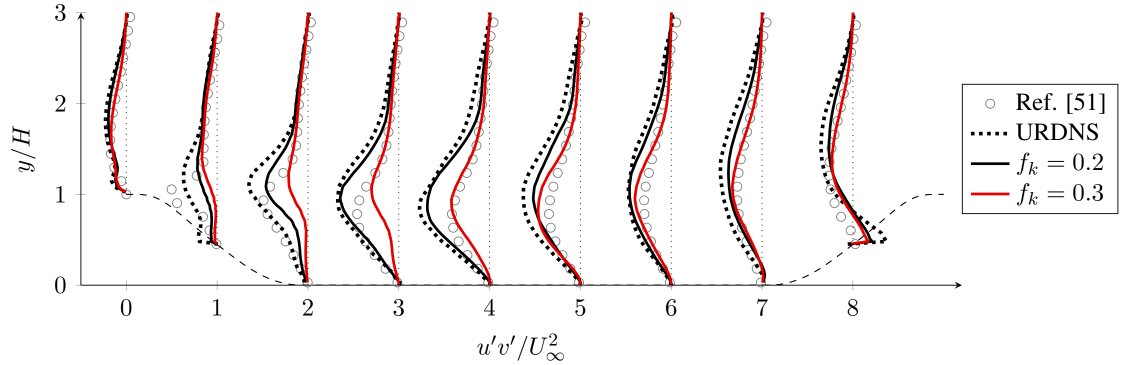
(b) $\mathbb{P}_1$

Figure 2.4: Time and span-averaged normal velocity variance profiles using a $\mathbb{P}_3$ (top) and $\mathbb{P}_1$ FR scheme (bottom). Profiles are scaled by a factor of 10 and shifted by 0, +1, +2, ..., +8, respectively, along the abscissa.

The profiles of the normal velocity variance are shown in Fig. 2.4. For the low-order method, poor agreement with the reference data was generally observed for both PANS and URDNS, with



(a) $\mathbb{P}_3$



(b) $\mathbb{P}_1$

Figure 2.5: Time and span-averaged streamwise-normal velocity covariance profiles using a $\mathbb{P}_3$ (top) and $\mathbb{P}_1$ FR scheme (bottom). Profiles are scaled by a factor of 10 and shifted by 0, +1, +2, ..., +8, respectively, along the abscissa.

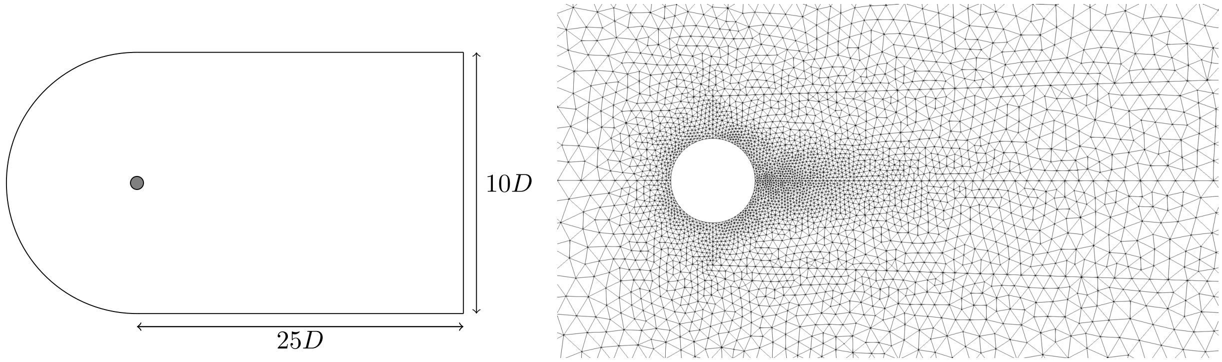
the PANS approach underpredicting the variance in the separation region and the URDNS approach overpredicting the variance outside of the separation region. However, outside of the separation region, the prediction by the low-order PANS approach significantly improved. When switching to the high-order method, the results of the URDNS approach degraded, with large oscillations observed in the variance profiles. This effect can likely be attributed to aliasing-driven instabilities evident in high-order approximations of under-resolved turbulent flows, which can introduce spurious high-frequency oscillations in the flow that are more evident in higher-order statistics [17]. When using the high-order method for the PANS approach, this effect was not seen, likely due to the suppression of aliasing errors as a result of the introduction of a physically appropriate eddy viscosity. Furthermore, significant improvements were observed in the variance profiles between the low-order and high-order PANS approaches, with the high-order PANS profiles showing excellent agreement with the reference data for both f_k values and notably less deviation in the results between different f_k values.

Similar observations were seen in the streamwise-normal velocity covariance profiles, shown in Fig. 2.5. No appreciable improvement in the covariance profiles was observed when switching from low-order to high-order URDNS, with the high-order approach showing better results in the separation region but at the expense of spurious oscillations outside of the separation region as in Fig. 2.4. The low-order PANS approach showed reasonable results outside of the separation region with $f_k = 0.3$, but underpredicted the magnitude of the covariance aft of the separation point for both values of f_k. The high-order PANS approach showed good agreement with the reference data for both f_k values, with significantly better prediction within the separation region and monotonic covariance profiles outside of the separation region. The decrease in sensitivity to the f_k parameter when switching to the high-order approach was not as evident for the streamwise-normal velocity covariance as it was with the streamwise and normal variance.

2.2.2 Cylinder

Due to the variety of physical phenomena that appear, the flow around a circular cylinder is of interest for many applications in fluid dynamics. At moderately-low Reynolds numbers (400-

5000), the flow lies in a subcritical regime where the transition of the laminar shear layer to a turbulent wake presents a challenge for numerical studies as the flow physics are very sensitive to the methods used which can lead to significant discrepancies in the results between various approaches [54]. The flow at a Reynolds number (Re_D) of 3900, based on the freestream velocity U_∞ and cylinder diameter D, lies in this regime, and as such, there exists vast amounts of numerical and experimental data for this configuration [54–59]. This problem was explored to conduct a comparison between low-order and high-order PANS and URDNS methods for resolving more complex flow physics. These results were compared to experimental and numerical data, and a DNS study was performed to provide additional data for comparison that was not available in the literature. The comparisons were carried out at levels of resolution which can be considered to be "moderately-resolved", where the prediction of first-order statistics can be done to a reasonable accuracy without turbulence modeling (i.e., implicit LES/under-resolved DNS). As such, the metric for comparison between the methods was their ability to resolve the more complex flow physics of the problem, such as the vortex shedding and Kelvin–Helmholtz frequencies and the presence of coherent turbulent structures in the wake.



(a) Geometry

(b) Mesh

Figure 2.6: Cross-section of the cylinder geometry in the streamwise-spanwise plane (left) and mesh for the $\mathbb{P}_3$ approach in the cylinder wake region (right).

The simulations were performed using a C-grid domain on $[x/D, y/D, z/D] \in [-10, 25] \times$

$[-10, 10] \times [0, \pi]$, where x, y, and z denote the streamwise, normal, and spanwise directions, respectively. The domain and grid are shown in Fig. 2.6. At the cylinder surface, a no-slip, adiabatic wall boundary condition was applied, with the turbulence variables taking on the values in Eq. (2.15). Periodic boundary conditions were applied along the spanwise direction. At the inlet, a uniform density and momentum were fixed, and the turbulence variables were set according to Menter [48] as

$$k_u = 10^{-3} f_k \frac{U_\infty^2}{Re_D}, \qquad \omega_u = f_\omega \frac{U_\infty}{D}. \tag{2.16}$$

At the outlet, the pressure was set to a fixed value corresponding to a Mach number of 0.1 while the remaining variables were free. For all simulations, the flow was assumed to be fully developed after $t = D/U_\infty = 100$, after which statistical quantities were gathered and analyzed until $t = D/U_\infty = 300$.

Three unstructured, prismatic grids were generated by extruding a 2D, triangular grid along the spanwise axis, one for the low-order ($\mathbb{P}_1$) PANS/URDNS approaches, one for the high-order ($\mathbb{P}_3$) PANS/URDNS approaches, and one for the DNS. A description of the characteristics of these meshes is shown in Table 2.2, where N denotes the number of elements and DOF denotes the degrees of freedom. The values corresponding to the unstructured, triangular elements in the streamwise-normal plane are denoted by the subscript xy while the values corresponding to the structured extrusion along the spanwise direction are denoted by the subscript z. The low-order mesh was generated by uniformly sub-dividing the high-order mesh along each direction. As a result of the subdivision of the triangular surfaces, the low-order mesh had approximately 20% more degrees of freedom than the high-order mesh. In comparison to the low-order PANS approach used for this configuration in Pereira et al. [24], the grids in the present work consist of approximately 20-30% fewer degrees of freedom.

The DNS study was performed using a $\mathbb{P}_3$ approximation with roughly 55.6 million degrees of freedom and an identical problem configuration (excluding the turbulence model). Through *a posteriori* analysis, the minimum Kolmogorov length scale in the domain was found to be $\eta/D = 0.011$. Therefore, this resolution was sufficient to achieve a $\Delta s_{xy}/\eta$ ratio between 0.25 and 0.83

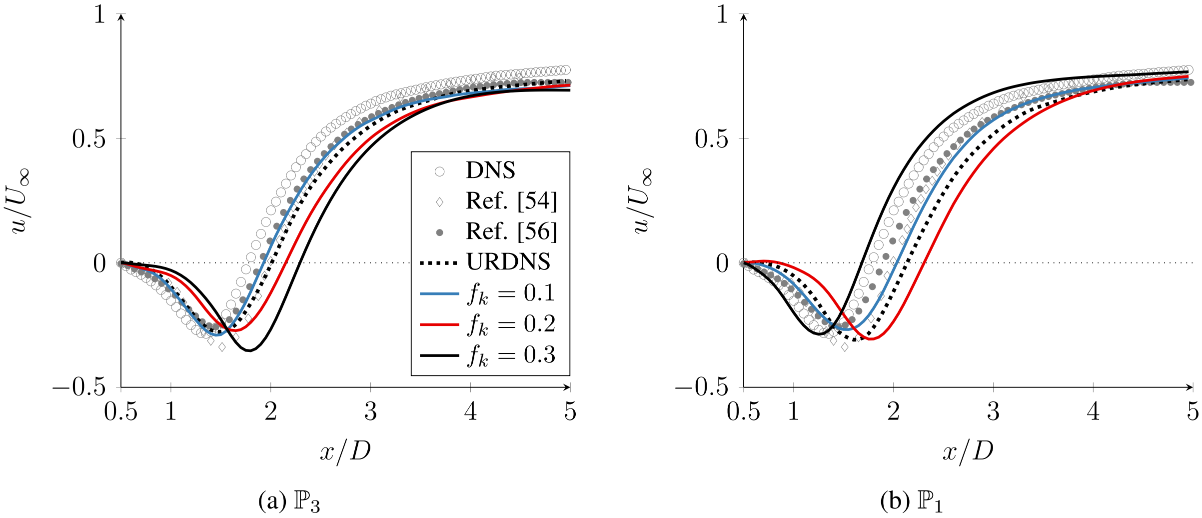
Grid	Method	Order	$N_{xy} \times N_z$	$DOF_{xy} \times DOF_z$
B1	PANS/URDNS	$\mathbb{P}_1$	$26,000 \times 20$	$78,000 \times 40$
B2	PANS/URDNS	$\mathbb{P}_3$	$6,500 \times 10$	$65,000 \times 40$
B3	DNS	$\mathbb{P}_3$	$58,000 \times 24$	$580,000 \times 96$

Table 2.2: Description of mesh characteristics for the cylinder case.

for $x/D < 10$, where Δs_{xy} is the average subcell size within an element in the streamwise-normal plane. Furthermore, the $\Delta s_z/\eta$ ratio was 0.85, where Δs_z is the average subcell size within an element in the spanwise direction. The results of the DNS study were used to perform a Proper Orthogonal Decomposition (POD) analysis to identify the presence of coherent turbulent structures which may not be evident in the first and second-order statistics. This method, described in detail in Chapter B, was used to characterize the time-dependent velocity field by a set of orthonormal spatial modes, such that the modes containing the most energy corresponded to the most dominant features in the flow field. To evaluate the ability of the various approaches in this work in predicting the dominant flow physics of the problem, the highest energy mode of the streamwise velocity fluctuations was compared to the DNS results.

After averaging across the time-averaging horizon and along the spanwise direction, the PANS and URDNS approaches were compared to the DNS results of the present work, the LES results of Parnaudeau et al. [54], and the LES results of Witherden et al. [56]. For comparison with the work of Witherden et al. [56] where individual data is presented for the two distinct shedding modes across a large time horizon, we take the average of the two modes. These results were analyzed with respect to the first and second-order statistics, frequency spectra, and POD modes.

The profiles of the averaged centerline streamwise velocity are shown in Fig. 2.7 in comparison to the reference data. Between all of the approaches, reasonable approximations of the centerline streamwise velocity were obtained when taking into account the variation in the reference results. For the URDNS approaches, the low-order method overpredicted the edge of the recirculation region, but this was remedied with the high-order method. For the PANS approaches, the effects of the higher-order discretization were not immediately evident. With the low-order method, the efficacy of the PANS model was sensitive to the value of f_k, with $f_k = 0.1$ showing reasonable agreement with the LES of Parnaudeau et al. [54] and $f_k = 0.3$ showing reasonable agreement with the DNS results. With the high-order method, less variation of the results with respect to f_k was seen. Excellent agreement between the $f_k = 0.1$ results and the LES of Witherden et al. [56] was observed. The introduction of the PANS model generally prolonged the recirculation region with the high-order method, whereas for the low-order method, the effect was not clear.



(a) $\mathbb{P}_3$

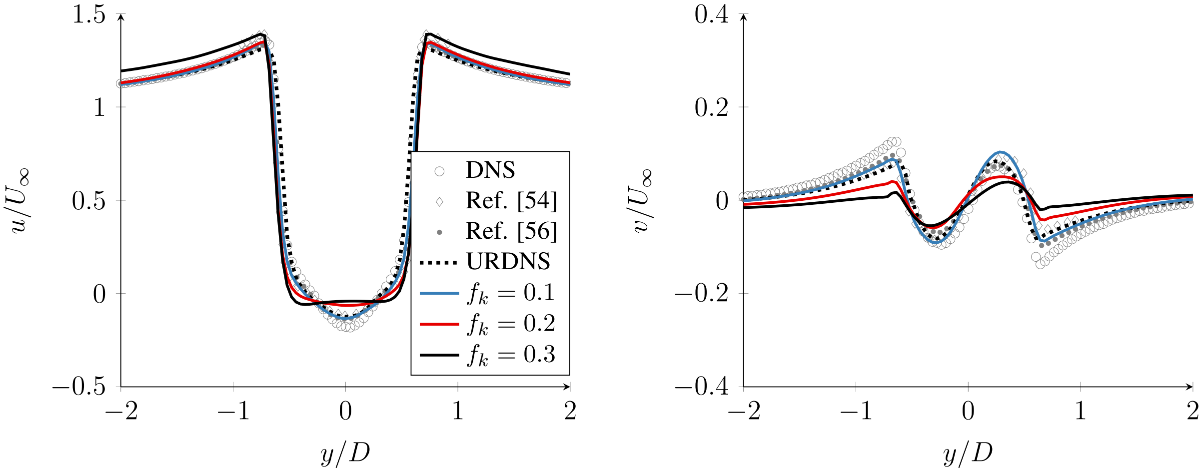
(b) $\mathbb{P}_1$

Figure 2.7: Time and span-averaged streamwise velocity profiles along the centerline ($y/D = 0$).

The averaged streamwise and normal velocity profiles at $x/D = 1.06$, 1.54, and 2.02 are shown in Fig. 2.8 and Fig. 2.9 for the $\mathbb{P}_3$ and $\mathbb{P}_1$ schemes, respectively. For the streamwise velocity profiles at $x/D = 1.06$ and 1.54, minimal differences were observed between the low-order and high-order PANS and URDNS approaches, with all methods showing reasonable agreement with the reference data. At $x/D = 2.02$, effects similar to the centerline velocity profiles were seen, where the sensitivity of the PANS approach to the f_k parameter decreased with the use of the high-order approach. For the normal velocity profiles, the differences between the various approaches were most evident at $x/D = 1.54$ but without clear distinction in the effects of higher-order discretizations on the PANS and URDNS approaches. Across the range of sampling locations, the best accuracy for the low-order approach was observed with $f_k = 0.3$, showing good agreement with the DNS results, and the best accuracy for the high-order approach was observed with $f_k = 0.1$, showing excellent agreement with the LES of Witherden et al. [56]. This observation is consistent with the PANS methodology, as the increase in fidelity afforded by high-order methods coincides with a decrease in the unresolved portion of the flow.

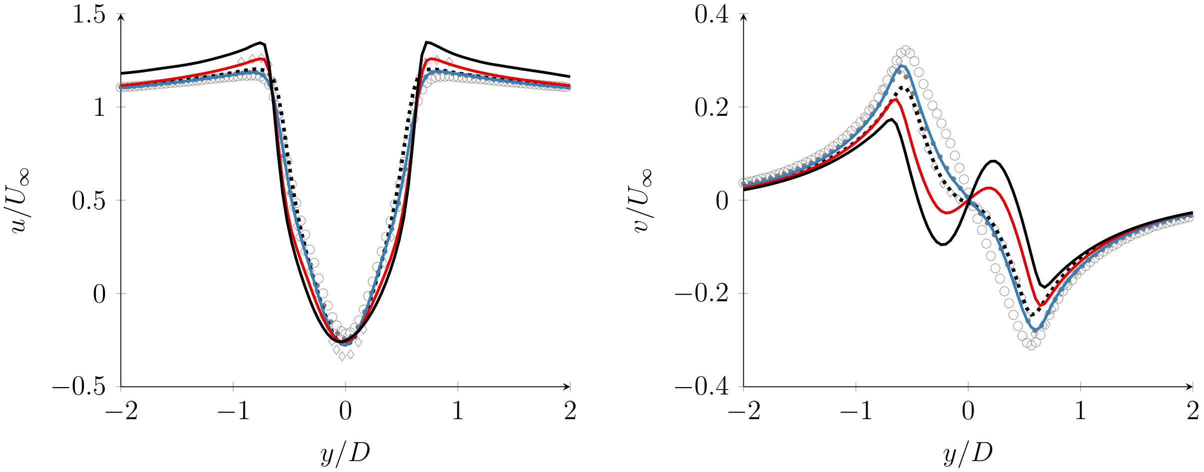
For the second-order statistics, the streamwise and normal velocity variance profiles as well as the streamwise-normal velocity covariance profiles at $x/D = 1.06$, 1.54, and 2.02 are shown in Fig. 2.10 and Fig. 2.11 for the $\mathbb{P}_3$ and $\mathbb{P}_1$ schemes, respectively. Between the various quantities and locations, the effects of the PANS model as well as the order of approximation were not consistent. Overall, although the low-order and high-order URDNS approach gave reasonable approximations of the second-order statistics at $x/D = 1.06$, the effects of the discretization as well as the introduction of the PANS model were more evident further away from the cylinder. Except for the case of $f_k = 0.3$, the switch from a low-order to a high-order approach generally tended to increase the magnitude of the second-order statistics, likely due to the decrease in numerical dissipation. The accuracy of the discretization did not necessarily affect the accuracy of the second-order statistics, but instead it primarily impacted the optimal value of f_k, with the low-order approach showing optimal results at $f_k = 0.3$ and the high-order approach showing optimal results at $f_k = 0.1$.

Due to the relative resolution of the grids, the first and second-order statistics were captured reasonably well across the various approaches used. For a more extensive comparison of these approaches, we instead focus on the flow physics and the characteristics of the coherent structures of the flow. The power spectra of the centerline normal velocity fluctuations at $x/D = 3$ and $x/D = 7$ were analyzed to evaluate the ability of the various methods to predict the frequencies of the vortex shedding and Kelvin–Helmholtz instabilities, as shown in Fig. 2.12 and Fig. 2.13. From the DNS results, the Strouhal number was determined to be $St = 0.209$, which is in agreement with the numerical and experimental results of Parnaudeau et al. [54], and the frequency



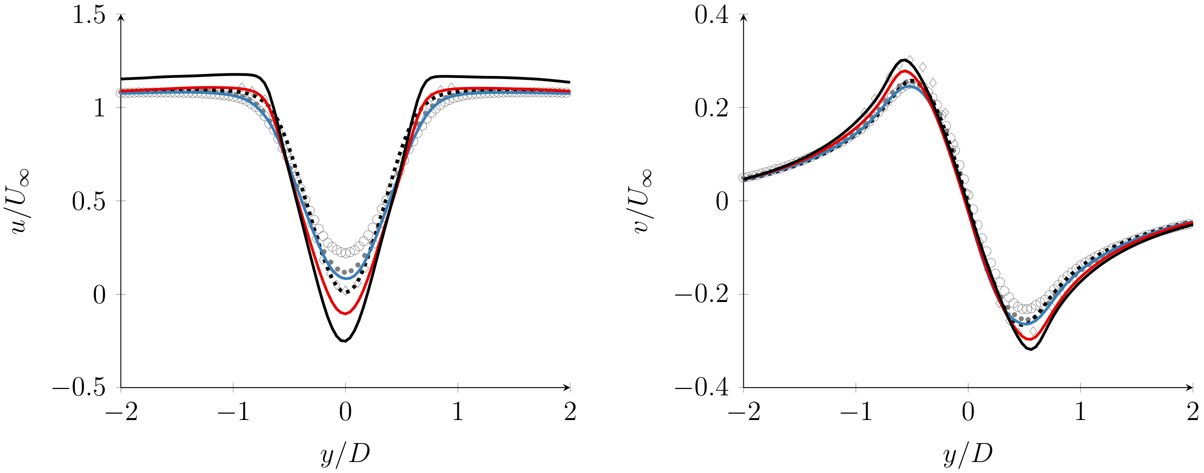
(a) $x/D = 1.06$

(b) $x/D = 1.06$



(c) $x/D = 1.54$

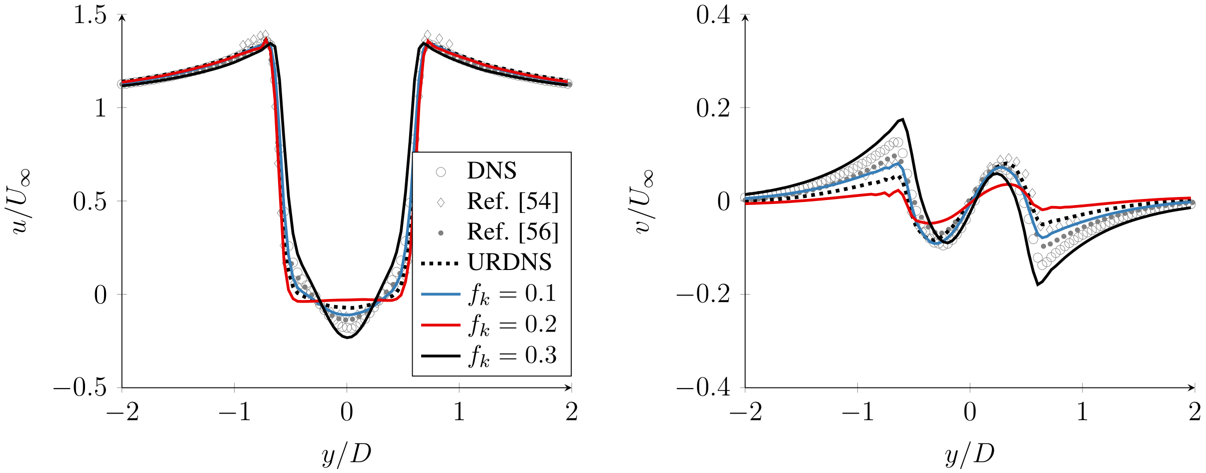
(d) $x/D = 1.54$



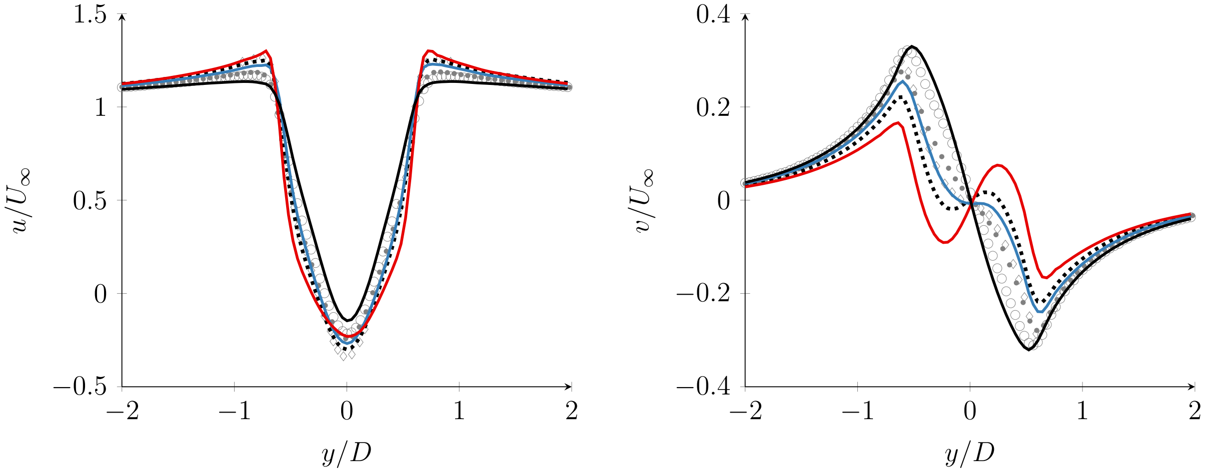
(e) $x/D = 2.02$

(f) $x/D = 2.02$

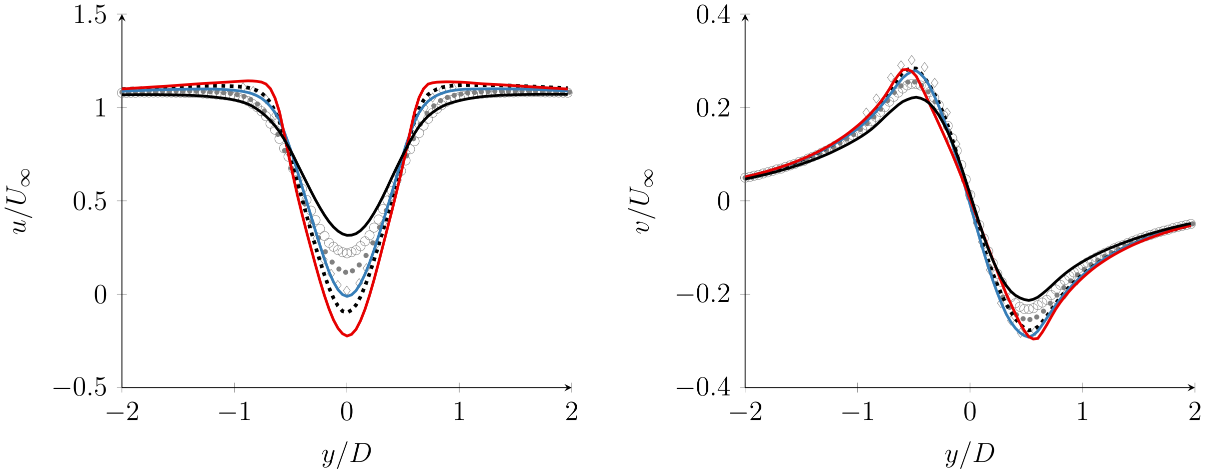
Figure 2.8: Time and span-averaged streamwise (left) and normal (right) velocity profiles at $x/D = 1.06$, 1.54, and 2.02 using a $\mathbb{P}_3$ FR scheme.



(a) $x/D = 1.06$ (b) $x/D = 1.06$

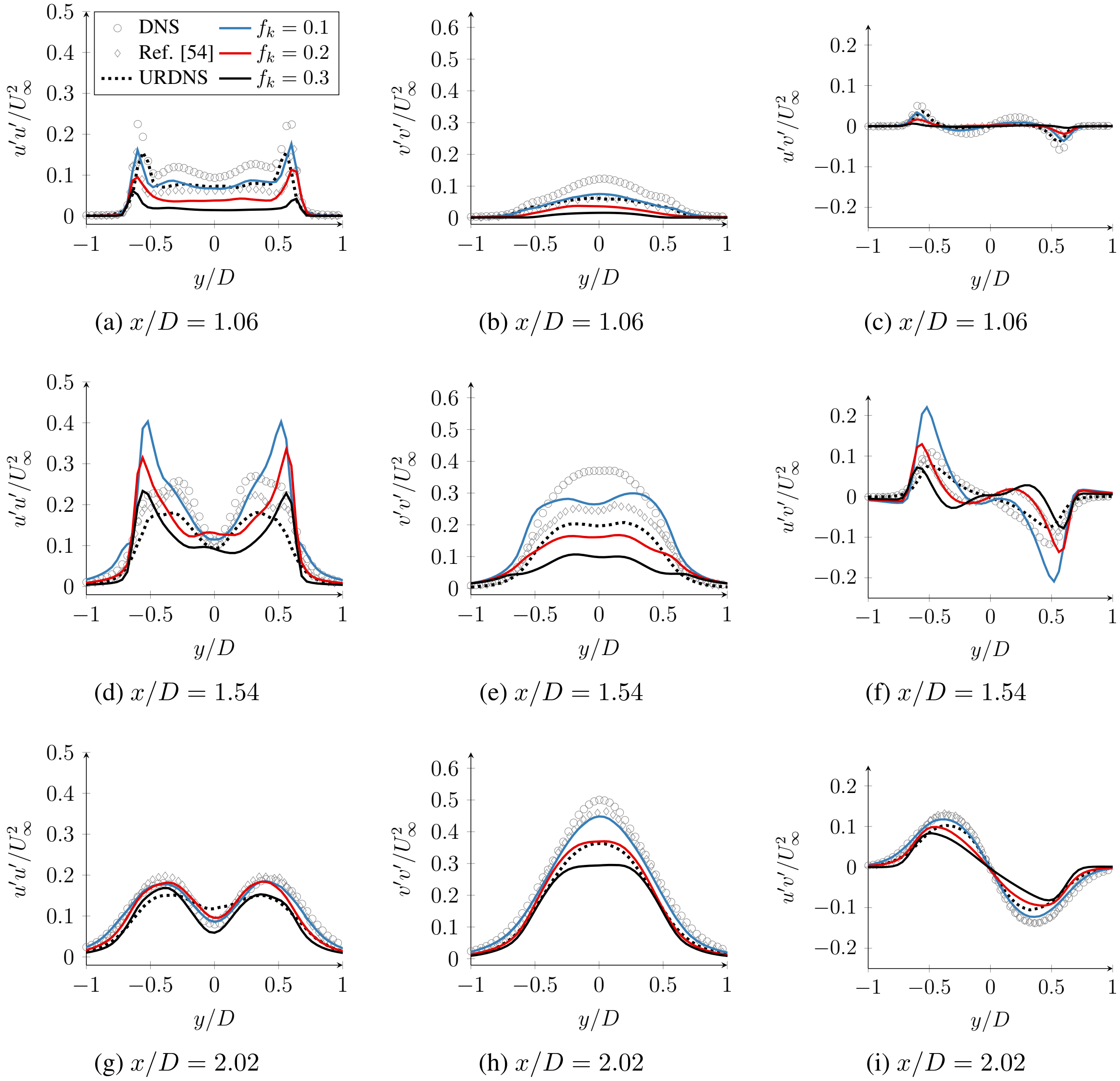


(c) $x/D = 1.54$ (d) $x/D = 1.54$



(e) $x/D = 2.02$ (f) $x/D = 2.02$

Figure 2.9: Time and span-averaged streamwise (left) and normal (right) velocity profiles at $x/D = 1.06$, 1.54, and 2.02 using a $\mathbb{P}_1$ FR scheme.

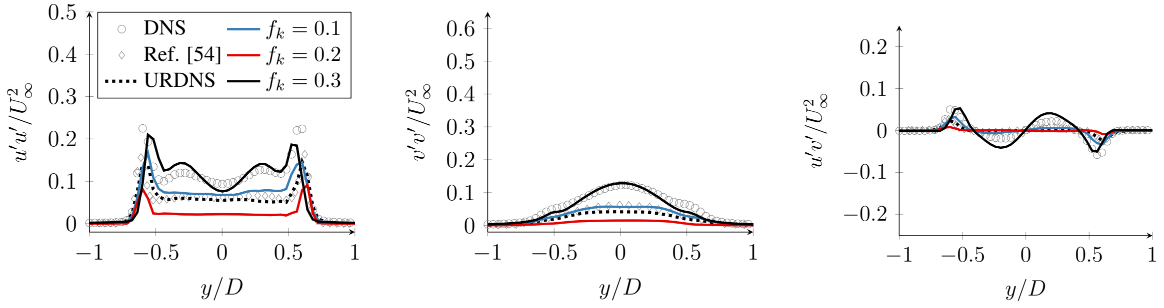


(a) $x/D = 1.06$ (b) $x/D = 1.06$ (c) $x/D = 1.06$

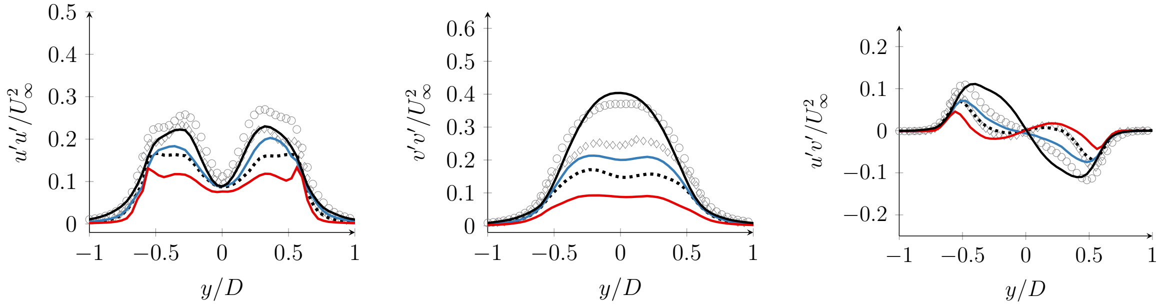
(d) $x/D = 1.54$ (e) $x/D = 1.54$ (f) $x/D = 1.54$

(g) $x/D = 2.02$ (h) $x/D = 2.02$ (i) $x/D = 2.02$

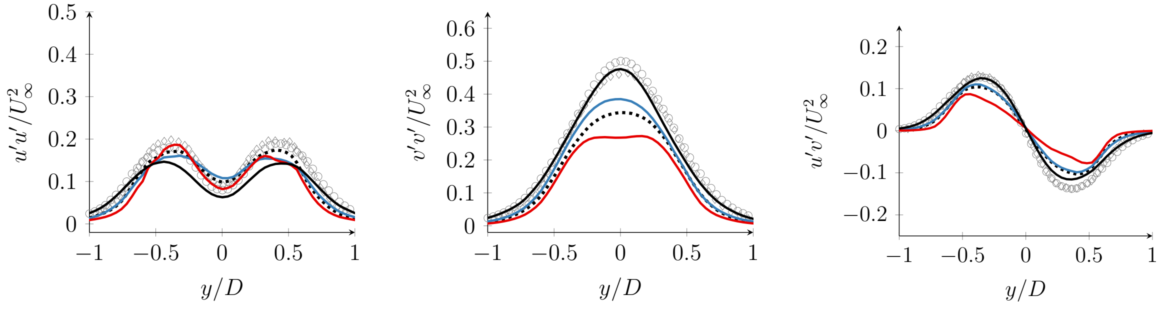
Figure 2.10: Time and span-averaged streamwise velocity variance (left), normal velocity variance (middle), and streamwise-normal velocity covariance (right) profiles at $x/D = 1.06$, 1.54, and 2.02 using a $\mathbb{P}_3$ FR scheme.



(a) $x/D = 1.06$ (b) $x/D = 1.06$ (c) $x/D = 1.06$



(d) $x/D = 1.54$ (e) $x/D = 1.54$ (f) $x/D = 1.54$



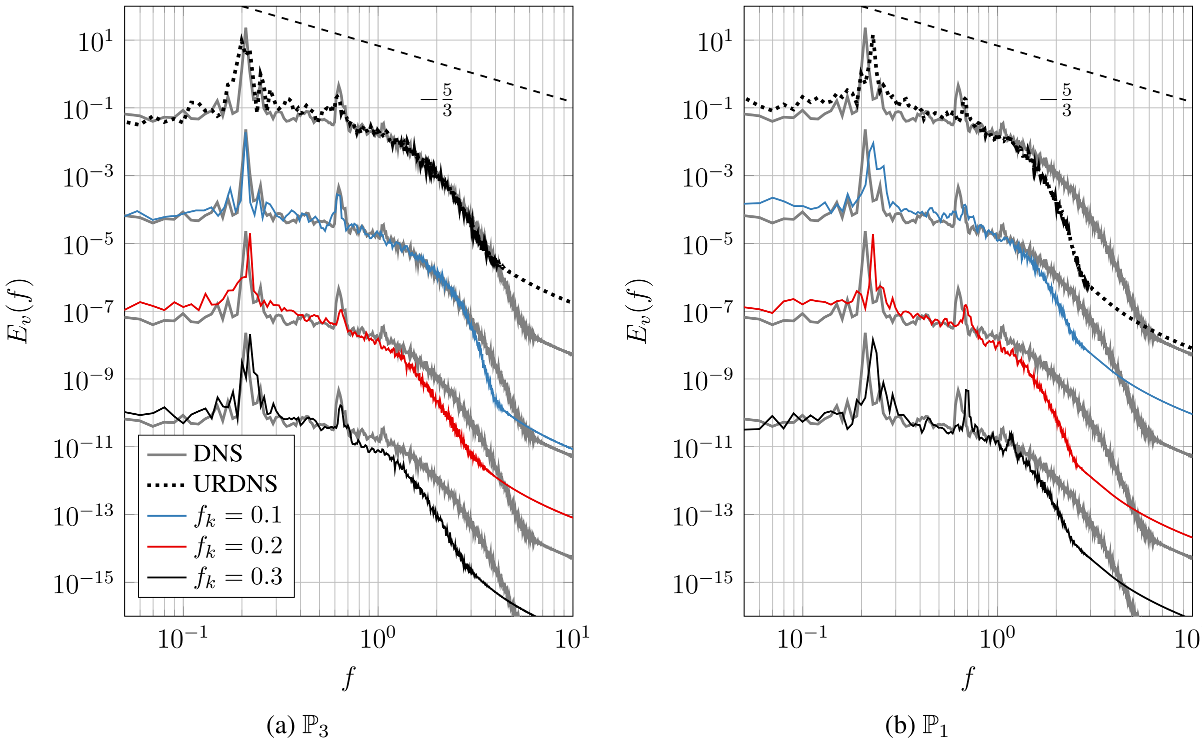
(g) $x/D = 2.02$ (h) $x/D = 2.02$ (i) $x/D = 2.02$

Figure 2.11: Time and span-averaged streamwise velocity variance (left), normal velocity variance (middle), and streamwise-normal velocity covariance (right) profiles at $x/D = 1.06$, 1.54, and 2.02 using a $\mathbb{P}_1$ FR scheme.

of the Kelvin–Helmholtz instability was determined to be $f_{KH} = 0.668$. At $x/D = 3$, the frequency spectra of the low-order approaches showed two distinct peaks corresponding to the vortex shedding and Kelvin–Helmholtz frequencies. However, the peak frequencies were overpredicted by roughly 10% and 6% for the vortex shedding and Kelvin–Helmholtz frequencies, respectively. Furthermore, except for the $f_k = 0.3$ case where the peak of the Kelvin–Helmholtz instability was prominent, the low-order methods generally did not adequately resolve the frequency of the Kelvin–Helmholtz instability. For the high-order methods, the frequencies of the vortex shedding and Kelvin–Helmholtz instability were inline with the DNS results. A more substantial improvement in the PANS results was observed in comparison to the URDNS results. For all values of f_k, both of the peaks were significantly more prominent than in the URDNS spectra where the peaks were distributed across a larger frequency range. Additionally, there was notably less sensitivity in the spectra to the f_k parameter with the high-order method than the low-order method, with $f_k = 0.1$ showing excellent agreement with the DNS results and $f_k = 0.2 - 0.3$ showing reasonable agreement.

At $x/D = 7$, where the propagation of the flow over a larger computational domain presents more challenges in resolving capability, these effects were amplified. The low-order approaches again overpredicted the vortex shedding and Kelvin–Helmholtz frequencies, but at this position further away from the cylinder, the Kelvin–Helmholtz frequency was not prominently resolved by neither URDNS nor the PANS approaches. With the high-order method, an even proportionally larger improvement was observed in the PANS approaches than the URDNS approach at this position. The URDNS approach did not adequately resolve the Kelvin–Helmholtz frequency and the peak of the vortex shedding frequency was distributed across a larger frequency range. However, the accuracy of the high-order PANS approaches did not deteriorate, with both the vortex shedding and Kelvin–Helmholtz frequencies prominently resolved and the spectra for all values of f_k showing good agreement with the DNS results.

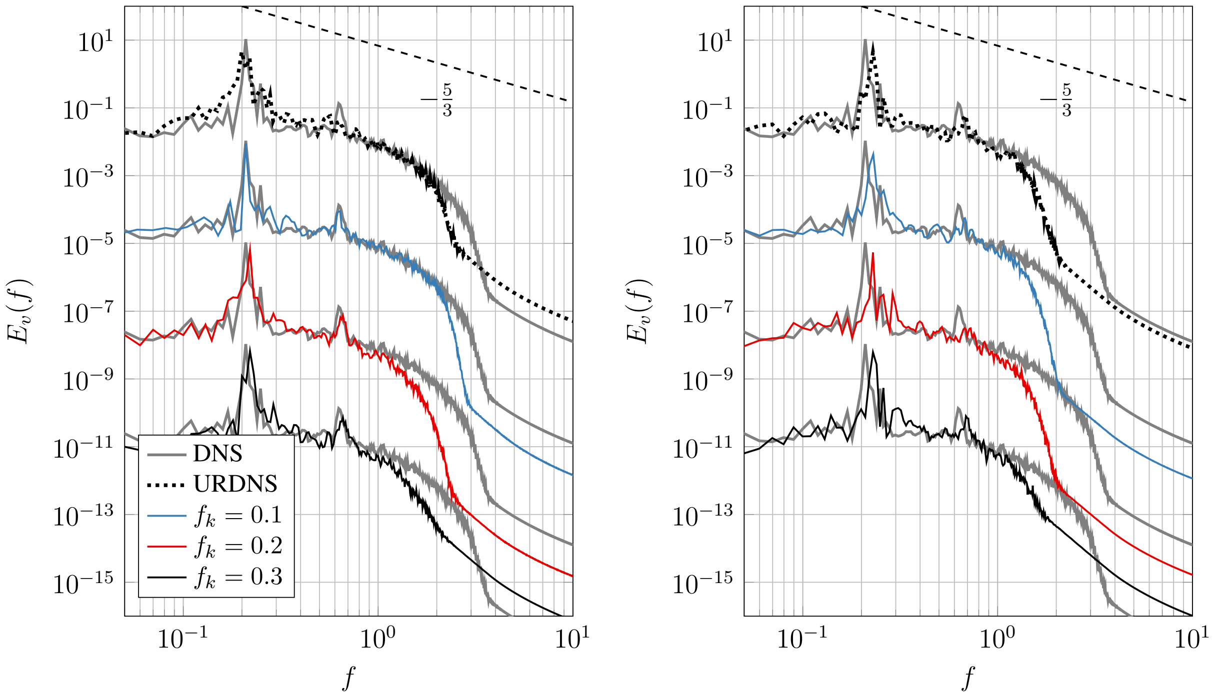
For an evaluation of the ability of the various methods in predicting the coherent structures in the flow, the primary POD mode of the streamwise velocity fluctuations was compared, shown



(a) $\mathbb{P}_3$

(b) $\mathbb{P}_1$

Figure 2.12: Power spectra of the centerline normal velocity fluctuations at $x/D = 3$. A scaling factor of 10^{-3} is applied between profiles. The frequency is nondimensionalized by D/U_∞.

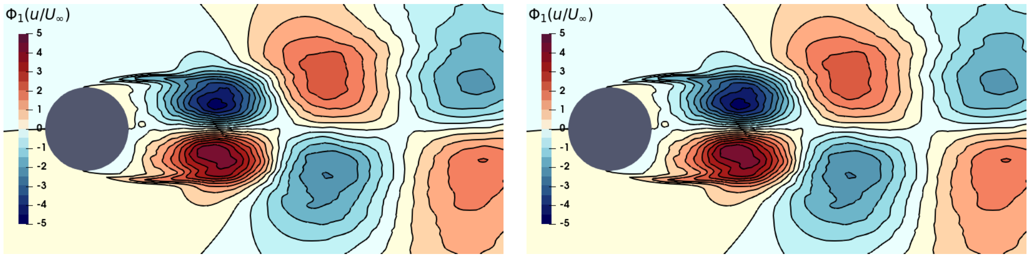


(a) $\mathbb{P}_3$

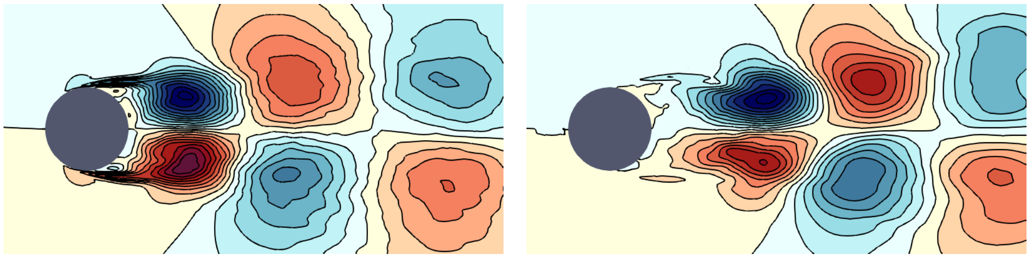
(b) $\mathbb{P}_1$

Figure 2.13: Power spectra of the centerline normal velocity fluctuations at $x/D = 7$. A scaling factor of 10^{-3} is applied between profiles. The frequency is nondimensionalized by D/U_∞.

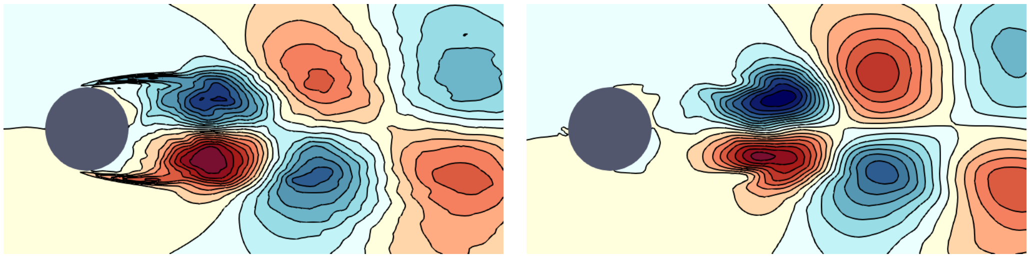
in Fig. 2.14. The primary POD mode from the DNS results depicts regions of strong correlation that are anti-symmetric across the centerline, indicative of anti-phased vortex shedding. Strong correlations in the separation region between $x/D = 2$ and $x/D = 3$ were observed as well as along the shear line emanating from the cylinder surface. Further along the wake, these correlated regions became less concentrated. Between the various approaches, the primary differences were in the positioning and shape of the strongly correlated region in the wake (i.e., the vortex shedding region) and the presence of the strongly correlated region in the shear layer. For the low-order method, the POD modes as predicted by the URDNS and PANS approaches only appreciably differed in the location of the vortex shedding region. Both approaches predicted this location farther aft of the DNS results regardless of the value of f_k, with the overprediction ranging monotonically from the minimum with the URDNS approach to the maximum with $f_k = 0.3$. Furthermore, neither the PANS nor the URDNS approaches replicated the circular shape of the vortex shedding region effectively, and the presence of a strongly correlated region in the shear layer was not seen. With the high-order method, these issues were generally rectified, with all approaches showing reasonable agreement with the DNS results in terms of the shape of the vortex shedding region and the presence of the strongly correlated region in the shear layer. As with the low-order method, the differences between the URDNS and PANS approaches were primarily with respect to the location of the vortex shedding region. The URDNS approach predicted this region much closer to the cylinder in comparison to the DNS results, whereas the PANS approaches generally showed good agreement with the DNS results. Due to the low variation in the results between various f_k values, no significant decrease in sensitivity to the f_k parameter was observed with the high-order approach.



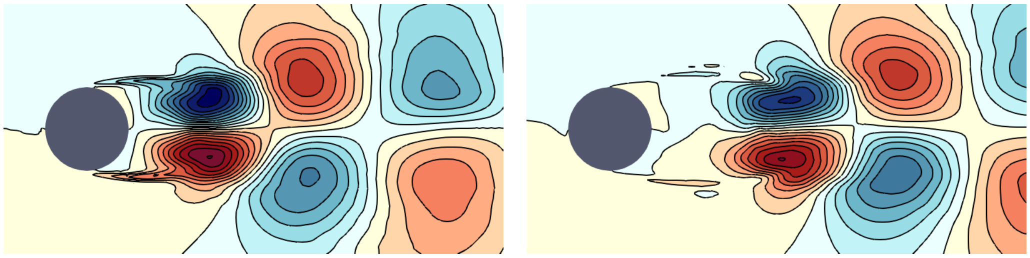
(a) DNS (b) DNS



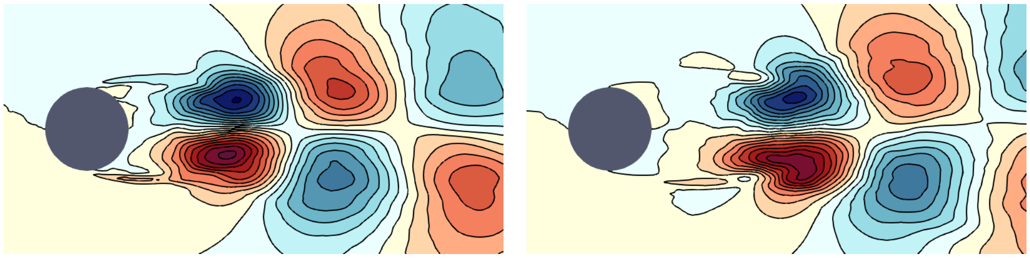
(c) $\mathbb{P}_3$ URDNS (d) $\mathbb{P}_1$ URDNS



(e) $\mathbb{P}_3$ $f_k = 0.1$ (f) $\mathbb{P}_1$ $f_k = 0.1$



(g) $\mathbb{P}_3$ $f_k = 0.2$ (h) $\mathbb{P}_1$ $f_k = 0.2$



(i) $\mathbb{P}_3$ $f_k = 0.3$ (j) $\mathbb{P}_1$ $f_k = 0.3$

Figure 2.14: Isocontour maps of the primary POD mode of the streamwise velocity. Contour lines represent 20 equispaced subdivisions across the range. DNS results are repeated across the top row.

2.3 Conclusion

The effects of high-order discretizations on a hybrid turbulence model were explored in the context of scale-resolving simulations of turbulent flows [60]. The PANS approach was discretized using the FR scheme and employed on two canonical benchmarks: the wall-bounded, separated flow around a periodic hill and the wake flow around a circular cylinder at $Re = 3900$. By varying the order of the approximation while fixing the total degrees of freedom, the effects of the discretization error on the PANS approach was independently investigated and compared to under-resolved DNS approaches.

In general, the switch from a low-order to a high-order approximation tended to proportionally improve both the URDNS and PANS approaches equally with respect to the first-order statistics. However, for highly under-resolved simulations such as the periodic hill in the present work, the high-order discretization improved the PANS prediction of the second-order statistics notably more than the URDNS prediction. For the more complex wake flow problem around the cylinder, the grid resolution was proportionally higher, and therefore the focus of the comparison was placed on the prediction of the flow physics since the first and second-order statistics could be reasonably approximated regardless of the methods used. The prediction of the flow physics improved significantly more through a high-order approximation using PANS than with URDNS, with much better prediction of the frequencies of the vortex shedding and Kelvin-Helmholtz instabilities, especially at distances farther along the wake. Furthermore, high-order approximations of the PANS approach showed a larger improvement in the prediction of the dominant POD mode of the streamwise velocity than high-order approximations of the URDNS approach.

Overall, high-order approximations tended to benefit the PANS approach proportionally more than the URDNS approach, as larger improvements in the prediction of the statistics and flow physics were generally seen with PANS. These benefits were attributed to the lower numerical dissipation of the high-order schemes, which allowed for better resolution of the small-scale features predicted by the model equations that can be dissipated by low-order schemes. Additionally, less sensitivity to the resolution-control parameter was observed with the high-order PANS approach, resulting in less variation in the predictions than with the low-order approximation. These findings indicate that high-order discretizations may be an effective approach for increasing the accuracy and reliability of hybrid turbulence models for scale-resolving simulations without a significant increase in computational effort.

3. POSITIVITY-PRESERVING ENTROPY-BASED ADAPTIVE FILTERING

Although DSEM offer many advantages for simulations of complex problems, their robustness is severely reduced for problems that exhibit discontinuities as the presence of spurious oscillations due to Gibbs phenomena can result in nonphysical solutions or the failure of the scheme altogether. Consequently, this lack of reliability is one of the limitations preventing the widespread adoption of these methods in the industry. To extend the use of DSEM to a wider variety of problems, various stabilization techniques have been proposed to increase the robustness of these schemes in the vicinity of discontinuities. A common goal of these shock capturing methods is to suppress numerical instabilities in the vicinity of a shock without degrading the accuracy of the underlying numerical scheme in regions where the solution is smooth.

These stabilization techniques can be broadly categorized as artificial viscosity, limiting, stencil modification, or filtering methods. The most ubiquitous approach is the addition of artificial viscosity to explicitly introduce numerical dissipation in the vicinity of a shock, a method pioneered by von Neumann and Richtmyer [61] with various advancements and alternative approaches over the decades [10, 31, 62–64]. Limiting methods have also shown promise [27, 28], with the premise of these schemes generally relying on a combination of a constraint-satisfying low-order scheme and a constraint-violating high-order scheme. Additionally, modifying the numerical stencil to alleviate the issues of Gibbs phenomena has shown to be accurate and robust in other numerical settings [29, 65–67], but the use of these methods is not very prevalent in the context of DSEM, primarily due to the incompatibility of the approach with the compact data structure, local compute, and geometric flexibility afforded by DSEM. In contrast, filtering [30, 68], where spurious oscillations are removed by reducing high-frequency modes in the solution, is particularly attractive as a shock capturing method for DSEM as the filter can sharply resolve discontinuities with minimal computational cost and without sacrificing the computational efficiency of the underlying numerical scheme.

However, a typical drawback of many of these approaches is that they (1) do not necessarily *guarantee* that physical constraints on the solution are satisfied, which may lead to the failure of the scheme, (2) have free parameters which can require problem- and mesh-dependent tuning, a cost that cannot be afforded for large scale-resolving simulations, and/or (3) are not easily and efficiently implemented in the context of explicit DSEM on modern computing architectures. These issues motivate the development of shock capturing approaches for DSEM that can guarantee certain physical constraints are satisfied without requiring problem-dependent tunable parameters. In this chapter, we present an adaptive filtering approach for shock capturing in nodal DSEM to address these issues. By formulating physical constraints such as positivity and a local minimum entropy principle as constraints on the discrete solution, the filter strength is computed via a simple scalar optimization problem requiring only element-local information. Under some basic assumptions on the properties of the numerical scheme, the filtered solution is guaranteed to satisfy these constraints, resulting in an efficient and robust method for resolving discontinuous features without the use of problem-dependent tunable parameters. Furthermore, the proposed filtering approach does not appreciably degrade the efficiency and accuracy of the standard DSEM approach for smooth solutions.

The remainder of this chapter is organized as follows. Preliminaries regarding the methodology are presented in Section 3.1. The proposed filtering approach is then presented in Section 3.2 followed by numerical implementation details in Section 3.3. The results of numerical experiments on the Euler and Navier–Stokes equations are then given in Section 3.4, followed by conclusions drawn in Section 3.5.

3.1 Preliminaries

To aid in presenting the proposed adaptive filtering approach, some preliminaries will be introduced in this section that briefly touch on the topics of hyperbolic conservation laws, minimum entropy principles, discontinuous spectral element methods, and modal filtering.

3.1.1 Hyperbolic Systems and Entropy Principles

The work in this chapter primarily pertains to approximations of hyperbolic conservation laws (or hyperbolic components of mixed hyperbolic-parabolic conservation laws) of the form of Eq. (1.1). For simplicity, the domain is assumed to be periodic or the solution is compactly supported. If a solution of Eq. (1.1) is an entropy solution [69], an entropy inequality of the form

$$\partial_t \sigma(\mathbf{u}) + \boldsymbol{\nabla}{\cdot}\boldsymbol{\Sigma}(\mathbf{u}) \geqslant 0, \tag{3.1}$$

can be posed, for which the inequality holds given any entropy-flux pair $(\sigma, \boldsymbol{\Sigma})$ [70] that satisfies

$$\partial_{\mathbf{u}} \boldsymbol{\Sigma} = \partial_{\mathbf{u}} \sigma \partial_{\mathbf{u}} \mathbf{F}.$$

This notion of the entropy σ is generally referred to as a *numerical* entropy [71]. We utilize a formulation of Eq. (3.1) that is of opposite sign in comparison to the work of Tadmor [71] for consistency with a physical entropy. In smooth regions, the inequality is satisfied exactly (i.e., the entropy functional obeys a conservation law), whereas in the vicinity of shocks, the left-hand side attains a strictly positive value (i.e., an entropy source exists).

For certain systems, entropy solutions of Eq. (1.1) satisfy a minimum principle on the entropy, i.e.,

$$\sigma\left(\mathbf{u}(\mathbf{x}, t + \Delta t)\right) \geqslant \min_{\mathbf{x} \in \Omega} \sigma\left(\mathbf{u}(\mathbf{x}, t)\right), \tag{3.2}$$

for all $\Delta t > 0$ [71, 72]. Assuming a finite propagation speed in hyperbolic systems, it is possible to restrict this condition over a local domain of influence to form a local minimum principle on the entropy. For a given point $\mathbf{x}_0$, let D_0 be some local domain of influence over the interval $[t, t + \Delta t]$ – e.g., a closed d-ball centered on $\mathbf{x}_0$ with a radius of $\lambda_{\max}\Delta t$, where $\lambda_{\max}$ is some upper bound on the local maximum propagation speed of the system. A local minimum entropy principle can then be given as

$$\sigma\left(\mathbf{u}(\mathbf{x}_0, t + \Delta t)\right) \geqslant \min_{\mathbf{x} \in D_0} \sigma\left(\mathbf{u}(\mathbf{x}, t)\right). \tag{3.3}$$

For numerical approximations of hyperbolic conservation laws, it can be advantageous to find schemes that can enforce this condition in scenarios where enforcing Eq. (3.1) is not feasible as it still guarantees some notion of physicality to the solution and can help alleviate instabilities in the numerical scheme [73].

3.1.2 Spatial Discretization

For a nodal DSEM approximation of Eq. (1.1), we assume that the spatial scheme is chosen such that

$$\partial_t \overline{\mathbf{u}} = - \int_{\partial\Omega_k} \overline{\mathbf{F}}\left(\mathbf{x}\right)\ \mathrm{d}\mathbf{s} \approx - \sum_{j \in I} m_j \overline{\mathbf{F}}_j, \tag{3.4}$$

where m_j is the corresponding quadrature weight for the point $\mathbf{x}_j$ and $\overline{\mathbf{u}}$ is the element-wise mean defined as

$$\overline{\mathbf{u}} = \frac{1}{V_k} \int_{\Omega_k} \mathbf{u}(\mathbf{x})\, \mathrm{d}\mathbf{x} \qquad \text{and} \qquad V_k = \int_{\Omega_k} \mathrm{d}\mathbf{x}, \tag{3.5}$$

for some arbitrary element Ω_k. This relation is recovered for nodal discontinuous Galerkin approximations with appropriate quadrature and flux reconstruction schemes utilizing the equivalent discontinuous Galerkin correction functions [74]. The ability of DSEM to preserve desirable properties on the element-wise mean is well documented in the literature due to its equivalency to first-order Godunov methods [75–77]. Without presenting a comprehensive proof, we assume that properties such as positivity of certain convex functionals of the solution and a discrete local minimum entropy principle are satisfied by the element-wise mean if one utilizes explicit strong stability preserving time integration under some Courant-Friedrichs–Lewy (CFL) condition [78] with an appropriate choice for the Riemann solver. A more detailed description of these conditions and assumptions is presented in Section 3.3.

3.1.3 Modal Filtering

The approximation of the solution given by a nodal basis can be equivalently expressed by a modal expansion as

$$\mathbf{u}(\mathbf{x}) = \sum_{i=1}^{N_s} \widehat{\mathbf{u}}_i \psi_i(\mathbf{x}), \tag{3.6}$$

where $\psi_i(\mathbf{x})$ are a set of modal basis functions (e.g., Legendre polynomials, Koornwinder polynomials, etc.) and $\widehat{\mathbf{u}}_i$ are their corresponding modes. The modal basis is generally chosen such that the basis functions are orthogonal with respect to the inner product. With this formulation, a filtered solution can be defined as

$$\widetilde{\mathbf{u}}(\mathbf{x}) = \sum_{i=1}^{N_s} H_i\left(\widehat{\mathbf{u}}_i\right) \psi_i(\mathbf{x}), \tag{3.7}$$

where $H_i\left(\widehat{\mathbf{u}}_i\right)$ denotes some filtering operation applied to the modes. The filter function can be arbitrarily chosen, but must be dissipative, i.e.,

$$\left| H_i\left(\widehat{\mathbf{u}}_i\right) \right| \leq \left| \widehat{\mathbf{u}}_i \right|, \tag{3.8}$$

and conservative, i.e.,

$$\frac{1}{V_k} \int_{\Omega_k} \widetilde{\mathbf{u}}(\mathbf{x}) \, \mathrm{d}\mathbf{x} = \overline{\mathbf{u}}. \tag{3.9}$$

For many filtering approaches, the filter tends to be more dissipative for higher frequency modes as spurious oscillations tend to manifest as high-frequency modes in the solution [31].

3.2 Methodology

To enforce certain desirable properties of the systems in question (at least in a discrete sense), it is possible to formulate them as convex constraints on the solution. A common constraint for physical systems is that some convex functional $\Gamma\left(\mathbf{u}\right)$ is non-negative across the domain (e.g., density and pressure for the Euler and Navier–Stokes equations, water height in the shallow water equations, etc.). This constraint is given by the condition

$$\Gamma\left(\widetilde{\mathbf{u}}(\mathbf{x}_i)\right) \geq 0 \quad \forall \, i \in \{1, ..., N_s\}, \tag{3.10}$$

where the choice of functional(s) is dependent on the system in question. However, simply enforcing positivity of these functionals is usually not enough to ensure a well-behaved solution in the vicinity of a discontinuity. This constraint must generally be accompanied by a more restrictive condition, such as some local discrete minimum entropy principle on the solution, given by the condition

$$\sigma\left(\widetilde{\mathbf{u}}(\mathbf{x}_i)\right) \geq \sigma_{\mathrm{min}} \quad \forall\, i \in \{1, ..., N_s\}, \tag{3.11}$$

where $\sigma(\mathbf{u})$ is some convex entropy functional of the system in question and σ_{min} is some local minimum entropy to be defined in Section 3.2.2. Enforcing this condition on the entropy tends to alleviate many of the issues regarding spurious oscillations in the vicinity of discontinuities [73], although this is not always guaranteed if this condition is not enforced for *every* possible entropy-flux pair.

3.2.1 Adaptive Filtering

To enforce these properties, we define an adaptive filtering operation with the goal of satisfying the constraints on the discrete filtered solution without the need for problem-dependent tunable parameters. The specific choice of filter is not particularly important as long as the filter meets the following criteria:

- The filter is dependent on a single free parameter ζ (i.e., $H(\mathbf{u}) = H(\mathbf{u}, \zeta)$).
- The filter is conservative (i.e., $\int_{\Omega_k} \widetilde{\mathbf{u}}(\mathbf{x})\ \mathrm{d}\mathbf{x} = \int_{\Omega_k} \mathbf{u}(\mathbf{x})\ \mathrm{d}\mathbf{x}$).
- There exists a minimum (or maximum) value of ζ such that the filter recovers the unfiltered solution (i.e., $\widetilde{\mathbf{u}}(\mathbf{x}) = \mathbf{u}(\mathbf{x})$).
- There exists a maximum (or minimum) value of ζ such that the filter recovers the mean mode (i.e., $\widetilde{\mathbf{u}}(\mathbf{x}) = \overline{\mathbf{u}}$).

The objective of the adaptive filter is to apply the minimum amount of filtering to the solution such that these constraints are met. We assume that for some arbitrary system, the discretization is chosen such that $\overline{\mathbf{u}}$ satisfies Eqs. (3.10) and (3.11), an assumption that is explored in more detail in Section 3.3 for the specific systems in this work. From this assumption, it can be seen that there exists at least one value of ζ that recovers a filtered solution for which the constraints are satisfied exactly, i.e., Eqs. (3.10) and (3.11) are satisfied and at least one inequality becomes an equality. This value of ζ can be calculated through simple root-finding methods such as the bisection algorithm, and, in practice, is generally unique, although this is not guaranteed.

In this work, a second-order exponential filter [30] is chosen, given by the filter function

$$H_i\left(\widehat{\mathbf{u}}_i\right) = \widehat{\mathbf{u}}_i\, e^{-\zeta p_i^2}, \tag{3.12}$$

where p_i is defined as the maximal order of the modal basis function $\psi_i(\mathbf{x})$. For this choice of filter, setting $\zeta = 0$ recovers the unfiltered solution and setting $\zeta = \infty$ recovers the mean mode, the latter of which can be approximated in a computational sense as $\zeta = \mathcal{O}\left(-\log(\epsilon)\right)$ for some value of machine precision ϵ. With this formulation, we define the filter strength using the minimum value of ζ such that the solution abides by the positivity-preserving and discrete minimum entropy principle satisfying conditions, given by

$$\zeta = \underset{\zeta\, \geq\, 0}{\arg\min}\ \text{ s.t. }\ \left[\Gamma\left(\widetilde{\mathbf{u}}(\mathbf{x}_i)\right) \geq 0,\ \sigma\left(\widetilde{\mathbf{u}}(\mathbf{x}_i)\right) \geq \sigma_{\min}\ \forall\, i \in \{1, ..., N_s\}\right]. \tag{3.13}$$

From a computational perspective, convergence to a local minima of ζ is sufficient in the case that there exist multiple values of ζ such that the constraints are satisfied exactly. A description of the approach for computing ζ as well as the choice of functionals to enforce constraints upon for the various systems is given in Section 3.3. For sufficiently-resolved smooth solutions, the unfiltered solution is expected to already abide by these constraints [10], and therefore the standard DSEM approximation would be recovered. The proposed filtering approach is hereafter referred to as entropy filtering as the method effectively filters the modes of the solution that contribute to the violation of a minimum entropy principle.

Remark (Limiting)**.** *The proposed adaptive filtering operation can be considered somewhat similar to limiting-type approaches such as flux-corrected transport [28] and convex limiting [27]. In* *fact, for an unconventional filtering operation, given by*

$$H_i\left(\widehat{\mathbf{u}}_i\right) = \begin{cases} \widehat{\mathbf{u}}_i & \text{if } p_i = 0, \\ \zeta\widehat{\mathbf{u}}_i & \text{else,} \end{cases}$$

the filter recovers the approach of Zhang and Shu [75], a linear convex limiting operation between a low-order solution ($\zeta = 0$*) and a high-order solution (*$\zeta = 1$*). For a less trivial filter choice, this can be considered to be a nonlinear (and generally non-convex) limiting operation.*

3.2.2 Entropy Constraints

Although positivity-preserving constraints are generally unequivocal for most physical systems, the notion of a minimum entropy principle is more ambiguous. Many systems are endowed with a multitude of entropy functionals $\sigma(\mathbf{u})$ [71] and the choice of $\sigma_{\min}$ is not clearly defined. In this work, the minimum entropy principle is enforced on a numerical entropy that can be chosen arbitrarily from any entropy-flux pair $(\sigma, \boldsymbol{\Sigma})$ that satisfies Eq. (3.1) for the given system, the particular choice of which is posited to have a minor overall effect.

To calculate $\sigma_{\min}$, the discretization is assumed to be explicit in time under some standard CFL condition. With this assumption, the domain of influence of an arbitrary element Ω_k over a single temporal integration step can be considered to be strictly contained within the element and its direct Voronoi neighbors. Thus, $\sigma_{\min}$ can be defined using information only from an element and its direct neighbors. Let σ_*^k be defined as the minimum entropy within an element Ω_k, given as

$$\sigma_*^k = \min_{i \in \{1,\ldots,N_s\}} \sigma\left(\mathbf{u}_k(\mathbf{x}_i)\right), \tag{3.14}$$

where $\mathbf{u}_k$ denotes the solution within the element Ω_k. Furthermore, let $\mathcal{A}(k)$ be the set of element indices which are face-adjacent with Ω_k, including Ω_k itself. The local minimum entropy $\sigma_{\min}^k$

associated with the element Ω_k is then calculated as

$$\sigma_{\min}^k = \min_{i \in \mathcal{A}(k)} \sigma_*^i. \tag{3.15}$$

For elements adjacent to boundaries, the entropy of the boundary state is used as the adjacent entropy value. Assuming that these minimum entropy values are calculated prior to a temporal integration step and the filter is applied afterwards, this formulation of the entropy constraint enforces that the local discrete minimum entropy is non-decreasing in time across its domain of influence.

3.3 Implementation

3.3.1 Governing Equations and Constraints

The efficacy of the entropy filtering approach was evaluated on hyperbolic and mixed hyperbolic-parabolic conservation laws. For the hyperbolic system, the compressible Euler equations were chosen, given by Eq. (1.3). Positivity constraints were placed on the density and pressure,

$$\Gamma_1(\mathbf{u}) = \rho \qquad \text{and} \qquad \Gamma_2(\mathbf{u}) = P,$$

and the entropy functional was chosen as

$$\sigma = \rho \log(P\rho^{-\gamma}),$$

taken from the entropy-flux pair $(\sigma, \mathbf{v}\sigma)$. Assuming an explicit strong stability preserving temporal integration scheme under some standard CFL condition with a solution that initially satisfies these constraints discretely, the element-wise mean at the next temporal step will satisfy these constraints if the interface fluxes are computed using an entropy-stable positivity-preserving Riemann solver [77, 79] (e.g., Godunov methods [39], local Lax-Friedrichs flux [80], HLLC [81] with appropriate wavespeed estimates). These properties of the element-wise mean were shown for discontinuous Galerkin approximations in the works of Zhang and Shu [75], Zhang and Shu [76], Zhang et al. [77], and Chen and Shu [79].

Remark (Source terms)**.** *The extension of the proposed approach to hyperbolic systems with source terms is possible for discretizations that preserve the constraints on the element-wise mean. For positivity constraints, Zhang and Shu [82] showed these properties under a potentially more restrictive time step condition. For the entropy constraint, the contribution of the source term to the entropy over the temporal integration step would have to be evaluated to augment the* $\sigma_{\min}$ *value, or if the entropy source is strictly positive, one may forego this modification and apply the proposed filter approach at the expense of a more relaxed entropy constraint.*

For the mixed hyperbolic-parabolic system, the compressible Navier–Stokes equations were chosen, given by Eq. (1.4). As the assumption on the entropy of the element-wise mean is not necessarily satisfied with the mixed hyperbolic-parabolic discretization, the hyperbolic and parabolic components of the conservation law were isolated and treated separately using an explicit operator splitting approach [83]. Identical positivity and entropy constraints are used for the Navier–Stokes equations as for the Euler equations. To enforce an entropy constraint for this system, at each temporal integration stage, the hyperbolic step is computed first, after which the entropy filter can be applied as with purely hyperbolic systems. The parabolic component of the temporal update is then added to the filtered solution. A final check is performed to ensure that the mixed hyperbolic-parabolic solution retains the positivity-preserving properties of the hyperbolic step. In the rare occasion that it does not, the filter is applied again using only positivity constraints. For both components of the flux, the boundary conditions were modified to ensure consistency with the system being solved (e.g., no-slip boundary conditions for the parabolic step were replaced with slip boundary conditions for the hyperbolic step).

Remark (Navier–Stokes equations)**.** *The notion of a minimum entropy principle is satisfied by the Navier–Stokes equations when considering the thermodynamic entropy (see Tadmor [71], Section 3). If the discretization for the parabolic component can be formed such that the minimum entropy principle is satisfied on the element-wise mean, the operator splitting approach can be neglected and the filter can be applied on the full hyperbolic-parabolic step with a significant reduction in computational cost. Alternatively, one may neglect the operator splitting approach without mod-*

ifying the parabolic discretization to reduce the computational cost at the expense of the entropy constraint not necessarily being satisfied.

3.3.2 Discretization

The solution nodes were distributed along the Gauss–Legendre–Lobatto quadrature points for tensor-product elements and the α-optimized points [13] for simplex elements. Common interface flux values were computed using the HLLC Riemann solver [81] for the inviscid fluxes and the BR2 approach [43] for the viscous fluxes. Temporal integration was performed using a three-stage third-order strong stability preserving (SSP) Runge–Kutta scheme [84]. For the modal basis, orthogonal polynomials with respect to the unit measure were used (i.e., Legendre basis for tensor-product elements, Proriol-Koornwinder-Dubiner-Owens basis for triangles, etc.).

3.3.3 Filter Implementation

The implementation of the entropy filter was formulated as an element-wise scalar optimization problem. At each substage of the temporal integration method, a filtering operation was performed on the solution to enforce the positivity-preserving and minimum entropy constraints, the latter of which was computed using the solution at the previous substage. If the unfiltered solution satisfied the constraints, no filter was applied. Otherwise, the minimum necessary filter strength ζ was calculated via 20 iterations of a bisection approach. Faster convergence could be obtained using more sophisticated root bracketing methods such as the Brent or Illinois methods, but these approaches were not explored in this work. For the density and pressure constraints, a minimum value of $\rho_{\min} = P_{\min} = 10^{-8}$ was enforced to ensure a non-vacuum state for the Riemann solver, such that the positivity constraints were instead implemented as

$$\Gamma_1(\mathbf{u}) = \rho - \rho_{\min} \qquad \text{and} \qquad \Gamma_2(\mathbf{u}) = P - P_{\min}.$$

For the entropy constraint, a numerical tolerance of $\epsilon_\sigma = 10^{-4}$ was given, such that the constraint was instead implemented as

$$\sigma\left(\widetilde{\mathbf{u}}(\mathbf{x}_i)\right) \geq \sigma_{\min} - \epsilon_\sigma \quad \forall\, i \in \{1, ..., N_s\}. \tag{3.16}$$

This comparatively larger tolerance resulted in a slightly relaxed entropy constraint which was found to be beneficial for two reasons. First, the constraint was notably more prone to numerical precision issues due to the logarithm operation, particularly in the limit as $\rho \to \rho_{\min}$ or $P \to P_{\min}$ since small variations in these values could cause orders of magnitude more variation in the entropy. Secondly, marginally better resolution of flow features could be obtained by allowing slight undershoots in the entropy as strictly enforcing the entropy principle can degrade the accuracy of the solution [27, 73]. A more detailed description of the computational implementation including pseudo-code is presented in Chapter C.

3.4 Results

The proposed entropy filtering approach was evaluated on a series of numerical experiments for the Euler and Navier–Stokes equations within a high-order flux reconstruction framework. For brevity, the solution for these systems is expressed in terms of a vector of primitive variables as $\mathbf{q} = [\rho, \mathbf{v}, P]^T$.

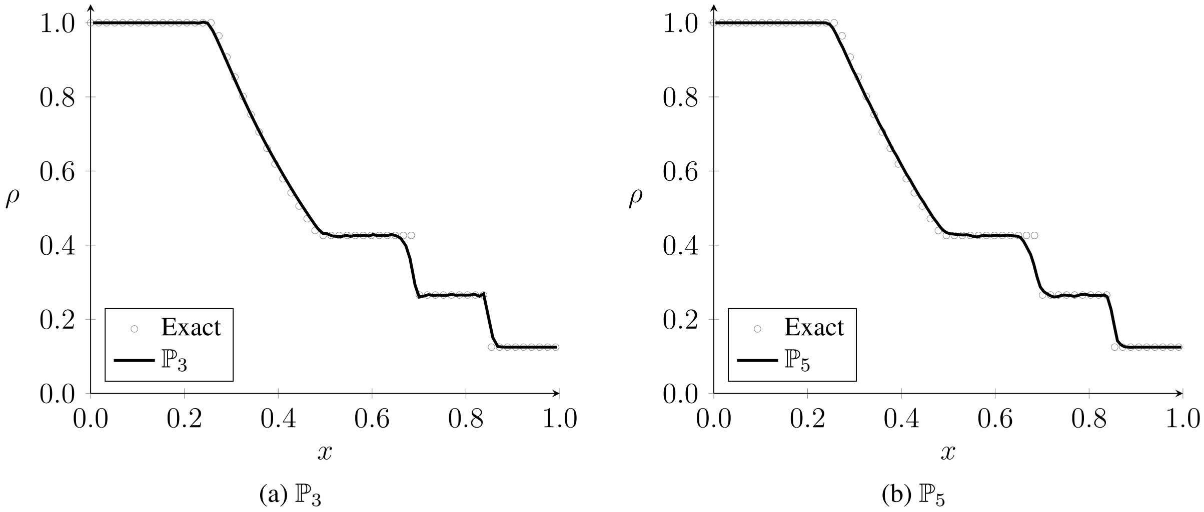
3.4.1 Euler Equations

3.4.1.1 Sod Shock Tube

For an initial evaluation of the shock capturing capabilities of the proposed approach in the context of the Euler equations, the canonical case of the Sod shock tube was considered [85]. The problem assesses the ability of the approach in resolving the three main features of the Riemann problem: shock waves, rarefaction waves, and contact discontinuities. The domain is set to $\Omega = [0, 1]$ and the initial conditions are given as

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 0.5, \\ \mathbf{q}_r, & \text{else,} \end{cases} \qquad \text{given} \quad \mathbf{q}_l = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 0.125 \\ 0 \\ 0.1 \end{bmatrix}.$$

The density profiles at $t = 0.2$ as predicted by a $\mathbb{P}_3$ and $\mathbb{P}_5$ FR approximation with 200 degrees of freedom are shown in Fig. 3.1. For both approximation orders, the results showed good agreement with the exact solution, with excellent resolution of the rarefaction wave and shock wave and minimal dissipation around the contact discontinuity. Furthermore, negligible spurious oscillations were observed in the vicinity of discontinuities. For a fixed number of degrees of freedom, marginally better results were obtained using the lower-order $\mathbb{P}_3$ approximation than the higher-order $\mathbb{P}_5$ approximation, particularly around the contact discontinuity. This effect can be attributed to the proportionally higher number of elements available for lower-order approximations, giving a more localized approach for the filter.



(a) $\mathbb{P}_3$ (b) $\mathbb{P}_5$

Figure 3.1: Density profile of the Sod shock tube problem at $t = 0.2$ computed using a $\mathbb{P}_3$ (left) and $\mathbb{P}_5$ (right) FR approximation with ∼200 degrees of freedom.

For a quantitative evaluation of the entropy filtering approach for discontinuous solutions, the convergence rates of the error against the exact solution were evaluated. For a given number of degrees of freedom M, the point-mean L^1 and L^2 norm of the density error was defined as

$$\epsilon_{\rho_1} = \frac{1}{M}\sum_{i=0}^{M-1} |\rho(x_i) - \rho_{\text{exact}}(x_i)| \quad \text{and} \quad \epsilon_{\rho_2} = \sqrt{\frac{1}{M}\sum_{i=0}^{M-1} (\rho(x_i) - \rho_{\text{exact}}(x_i))^2}, \tag{3.17}$$

respectively. The convergence rates of the density error with respect to the number of elements N for various approximation orders are shown in Table 3.1 and Table 3.2. The expected first-order convergence rate was generally obtained in both the L^1 and L^2 norm for all approximation orders, and the trend for the error was to decrease with increasing approximation order for a fixed number of elements. For a fixed number of degrees of freedom, shown in Table 3.3 and Table 3.4, marginally lower error was generally obtained with a lower-order approximation due to the previously mentioned effects.

N	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$	$\mathbb{P}_6$	$\mathbb{P}_7$
40	9.80×10^{-3}	8.57×10^{-3}	7.09×10^{-3}	8.08×10^{-3}	6.81×10^{-3}	6.96×10^{-3}
80	4.81×10^{-3}	4.30×10^{-3}	3.57×10^{-3}	4.15×10^{-3}	3.50×10^{-3}	3.61×10^{-3}
160	2.51×10^{-3}	2.33×10^{-3}	1.84×10^{-3}	2.16×10^{-3}	1.82×10^{-3}	1.99×10^{-3}
320	1.44×10^{-3}	1.30×10^{-3}	1.03×10^{-3}	1.18×10^{-3}	1.03×10^{-3}	1.14×10^{-3}
640	7.67×10^{-4}	6.03×10^{-4}	5.31×10^{-4}	6.32×10^{-4}	6.21×10^{-4}	6.38×10^{-4}
RoC	0.91	0.94	0.93	0.92	0.87	0.86

Table 3.1: Convergence of the L^1 norm of the density error with respect to mesh resolution N for the Sod shock tube problem at $t = 0.2$ with varying orders. Rate of convergence shown on bottom.

3.4.1.2 Shu-Osher Problem

To assess the effects of the entropy filter for more complex problems including shock waves and smooth oscillatory behavior, the case of Shu and Osher [29] was considered. The problem is

N	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$	$\mathbb{P}_6$	$\mathbb{P}_7$
20	3.34×10^{-3}	2.04×10^{-3}	2.09×10^{-3}	1.58×10^{-3}	1.45×10^{-3}	1.39×10^{-3}
40	1.84×10^{-3}	1.08×10^{-3}	1.15×10^{-3}	8.12×10^{-4}	8.81×10^{-4}	7.04×10^{-4}
80	8.02×10^{-4}	4.39×10^{-4}	5.06×10^{-4}	3.11×10^{-4}	3.39×10^{-4}	2.82×10^{-4}
160	3.28×10^{-4}	2.43×10^{-4}	2.30×10^{-4}	1.75×10^{-4}	1.87×10^{-4}	1.45×10^{-4}
320	2.05×10^{-4}	1.45×10^{-4}	1.39×10^{-4}	1.04×10^{-4}	1.06×10^{-4}	8.59×10^{-5}
RoC	1.05	0.98	1.01	1.01	0.98	1.03

Table 3.2: Convergence of the L^2 norm of the density error with respect to mesh resolution N for the Sod shock tube problem at $t = 0.2$ with varying orders. Rate of convergence shown on bottom.

M	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$	$\mathbb{P}_6$	$\mathbb{P}_7$
100	6.87×10^{-3}	9.51×10^{-3}	1.07×10^{-2}	1.27×10^{-2}	1.31×10^{-2}	1.72×10^{-2}
200	3.99×10^{-3}	5.37×10^{-3}	5.60×10^{-3}	7.34×10^{-3}	7.50×10^{-3}	9.70×10^{-3}
400	2.08×10^{-3}	2.72×10^{-3}	3.04×10^{-3}	3.88×10^{-3}	4.00×10^{-3}	5.65×10^{-3}
800	1.23×10^{-3}	1.46×10^{-3}	1.46×10^{-3}	2.08×10^{-3}	2.37×10^{-3}	2.90×10^{-3}
1600	6.60×10^{-4}	7.24×10^{-4}	9.00×10^{-4}	1.15×10^{-3}	1.24×10^{-3}	1.53×10^{-3}
RoC	0.85	0.93	0.91	0.88	0.85	0.87

Table 3.3: Convergence of the L^1 norm of the density error with respect to degrees of freedom M for the Sod shock tube problem at $t = 0.2$ with varying orders. Rate of convergence shown on bottom.

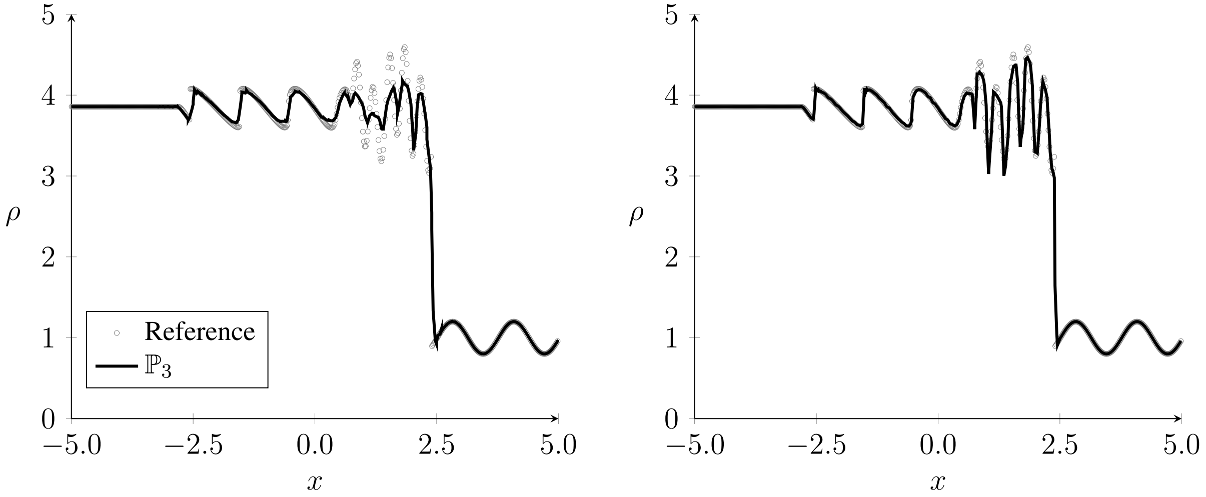
M	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$	$\mathbb{P}_6$	$\mathbb{P}_7$
100	1.21×10^{-3}	1.42×10^{-3}	1.95×10^{-3}	1.78×10^{-3}	1.94×10^{-3}	2.34×10^{-3}
200	6.62×10^{-4}	7.21×10^{-4}	7.89×10^{-4}	9.23×10^{-4}	1.07×10^{-3}	1.07×10^{-3}
400	2.72×10^{-4}	3.24×10^{-4}	4.26×10^{-4}	3.75×10^{-4}	4.76×10^{-4}	5.68×10^{-4}
800	1.81×10^{-4}	1.71×10^{-4}	1.58×10^{-4}	1.92×10^{-4}	2.78×10^{-4}	2.44×10^{-4}
1600	9.47×10^{-5}	7.39×10^{-5}	1.15×10^{-4}	1.09×10^{-4}	1.46×10^{-4}	1.17×10^{-4}
RoC	0.92	1.06	1.05	1.03	0.95	1.08

Table 3.4: Convergence of the L^2 norm of the density error with respect to degrees of freedom M for the Sod shock tube problem at $t = 0.2$ with varying orders. Rate of convergence shown on bottom.

solved on the domain $\Omega = [-5, 5]$ with the initial conditions

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant -4, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{given} \quad \mathbf{q}_l = \begin{bmatrix} 3.857143 \\ 2.629369 \\ 10.333333 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 1 + 0.2\sin 5x \\ 0 \\ 1 \end{bmatrix}.$$

The problem consists of a shock front propagating through a sinusoidally-perturbed density field, the interaction between which can induce instabilities in the flow field. However, these instabilities can be erroneously damped by overly dissipative shock capturing schemes. The predicted density profile at $t = 1.8$ computed using a $\mathbb{P}_3$ FR approximation with 100 and 200 elements is shown in Fig. 3.2. A reference solution was computed using a highly-resolved exact Godunov-type solver [26]. The results show good resolution of the leading and trailing shock waves, and no spurious oscillations were observed. The instabilities in the field aft of the leading shock were also well-resolved, particularly with increasing resolution.



(a) $N = 100$ (b) $N = 200$

Figure 3.2: Density profile of the Shu-Osher problem at $t = 1.8$ computed using a $\mathbb{P}_3$ FR approximation with 100 (left) and 200 (right) elements.

3.4.1.3 Isentropic Euler Vortex

For an extension to two-dimensional problems, the entropy filtering approach was initially applied to a smooth solution to verify the accuracy of the underlying DSEM was not detrimentally affected. The isentropic Euler vortex problem [66] was used as its analytic solution can be utilized to evaluate the convergence of the error. The initial conditions of the problem are given as

$$\mathbf{q}(\mathbf{x},0) = \begin{bmatrix} p^{\frac{1}{\gamma}} \\ V_x + \frac{S}{2\pi R}(y-y_0)\phi(r) \\ V_y - \frac{S}{2\pi R}(x-x_0)\phi(r) \\ \frac{1}{\gamma M^2}\left(1 - \frac{S^2 M^2(\gamma-1)}{8\pi^2}\phi(r)^2\right)^{\frac{\gamma}{\gamma-1}} \end{bmatrix}, \quad \text{where} \quad r = \|\mathbf{x}-\mathbf{x}_0\|_2 \quad \text{and} \quad \phi(r) = \exp\left(\frac{1-r^2}{2R^2}\right),$$

with the parameters $S = 13.5$ denoting the strength of the vortex, $R = 1.5$ the radius, $V_x = 0$, $V_y = 1$ the advection velocities, and $M = 0.4$ the freestream Mach number. The domain was set to $\Omega = [-10, 10]^2$, and a series of uniform quadrilateral meshes of size $N \times N$ was generated with periodic boundary conditions. After a single pass-through of the domain, the L^2 norm of the density error was calculated as

$$\epsilon_{\rho_2} = \sqrt{\frac{1}{|\Omega|}\int_{\Omega} (\rho - \rho_{\text{exact}})^2 \, \mathrm{d}\mathbf{x}}, \tag{3.18}$$

with the quadrature calculated on $(2p)^2$ Gauss-Legendre nodes. This error is tabulated in Table 3.5 for a series of experiments with varying mesh resolution and approximation order. The convergence rate of the error was generally in the range of p to $p+1$ for the varying approximation orders, on par with the theoretical rate of $p+1$. These findings indicate that the entropy filter does not appreciably degrade the accuracy of the underlying DSEM for smooth solutions.

To verify the computational efficiency of the entropy filter for smooth solutions where the unfiltered solution remains stable, the compute time for a solution without filtering and with a varying number of filter iterations was compared. The comparison was performed on an NVIDIA

N	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$	$\mathbb{P}_6$	$\mathbb{P}_7$
20	-	-	-	-	2.59×10^{-5}	5.37×10^{-6}
25	-	-	7.80×10^{-4}	7.73×10^{-5}	6.84×10^{-6}	9.85×10^{-7}
33	1.80×10^{-2}	1.79×10^{-3}	2.50×10^{-4}	1.11×10^{-5}	1.26×10^{-6}	9.58×10^{-8}
40	1.10×10^{-2}	7.58×10^{-4}	1.08×10^{-4}	2.86×10^{-6}	3.75×10^{-7}	1.97×10^{-8}
50	6.30×10^{-3}	3.02×10^{-4}	4.03×10^{-5}	7.52×10^{-7}	-	-
67	2.86×10^{-3}	1.05×10^{-4}	-	-	-	-
RoC	2.59	4.00	4.27	6.73	6.10	8.13

Table 3.5: Convergence of the L^2 norm of the density error with respect to mesh resolution N for the isentropic Euler vortex problem with varying approximation order. Rate of convergence shown on bottom.

V100 GPU over 5 flow-throughs of the domain on an $N = 40$ mesh with a $\mathbb{P}_3$ approximation and $\Delta t = 2{\cdot}10^{-4}$. For 5, 10, and 20 iterations of the filter, the relative computational cost increase (as measured by relative time-to-solution) was 1.1%, 1.7%, and 2.0%, respectively. These findings indicate that the entropy filter does not have a notable detrimental impact on the computational efficiency of the scheme for smooth solutions as it is primarily inactive. Furthermore, it also indicates that further improvements to the efficiency are possible through more sophisticated iterative solvers that require fewer iterations.

3.4.1.4 Double Mach Reflection

The double mach reflection problem of Woodward and Colella [86] was subsequently used to evaluate the ability of the entropy filter to resolve strong discontinuities in multiple dimensions. This case consists of a Mach 10 shock impinging on a 30 degree ramp and results in multiple strong shock-shock and shock-contact interactions. The problem is solved on the domain $\Omega = [0, 4] \times [0, 1]$ with the initial conditions

$$\mathbf{q}(\mathbf{x},0) = \begin{cases} \mathbf{q}_l, & \text{if } x < 1/6 + \tan(30^\circ) y, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{given} \quad \mathbf{q}_l = \begin{bmatrix} 8 \\ 7.14471 \\ -4.125 \\ 116.5 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 1.4 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

At the left boundary and the bottom boundary for $x < 1/6$, the solution was set to the post-shock state $\mathbf{q}_l$. No-slip adiabatic wall boundary conditions were applied for the bottom boundary for $x \geqslant \frac{1}{6}$. At the right boundary, the solution was set to the pre-shock state $\mathbf{q}_r$. For the top boundary, the exact solution is enforced, given as

$$\mathbf{q}(\mathbf{x},t)|_{y=1} = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 1/6 + \tan(30^\circ) y + \frac{10}{\cos(30^\circ)} t, \\ \mathbf{q}_r, & \text{else.} \end{cases}$$

The contours of density at $t = 0.2$ as predicted by a $\mathbb{P}_3$ FR scheme on a 2400×600 mesh are shown in Fig. 3.3. The results show sub-element resolution of discontinuities without the presence of spurious oscillations. Furthermore, the application of the filter did not excessively dissipate the Kelvin-Helmholtz instabilities along the contact line, indicating that the filter is not erroneously dissipating small-scale features. To verify this, the distribution of the filter parameter ζ is also shown in Fig. 3.3 overlaid on the isocontours of density. The filter was primarily active in the leading shock fronts, with minimal activation within the small-scale structures along contact line. Furthermore, even along the shock front, a relatively small value of ζ was required, significantly less than the value corresponding to the recovery of the mean mode.

3.4.1.5 Kelvin–Helmholtz Instability

To further evaluate the ability of the entropy filtering approach to resolve small-scale flow features in the vicinity of discontinuities, the roll-up of a Kelvin-Helmholtz instability was simulated.

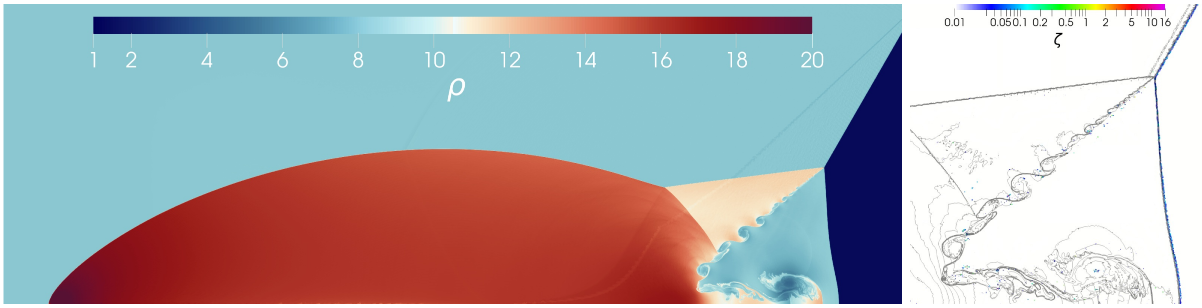


Figure 3.3: (Left) Contours of density on the subregion $[0, 3] \times [0, 1]$ for the double Mach reflection problem at $t = 0.2$ using a $\mathbb{P}_3$ FR approximation with a 2400×600 mesh. (Right) Enlarged view of the distribution of the filter parameter ζ overlaid on isocontours of density.

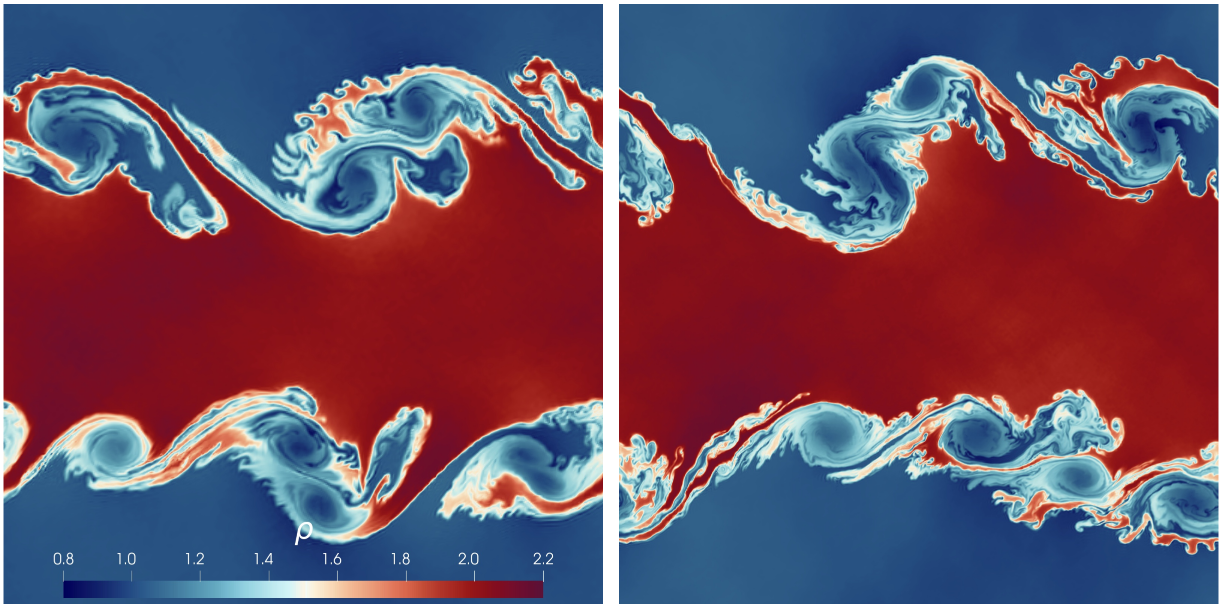
The problem is solved on the domain $\Omega = [-0.5, 0.5]^2$ with the initial conditions

$$\mathbf{q}(\mathbf{x}, 0) = \begin{cases} \mathbf{q}_l, & \text{if } |y| \leqslant 0.25, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{given} \quad \mathbf{q}_l = \begin{bmatrix} 2 \\ 0.5 \\ 0 \\ 2.5 \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 1 \\ -0.5 \\ 0 \\ 2.5 \end{bmatrix}.$$

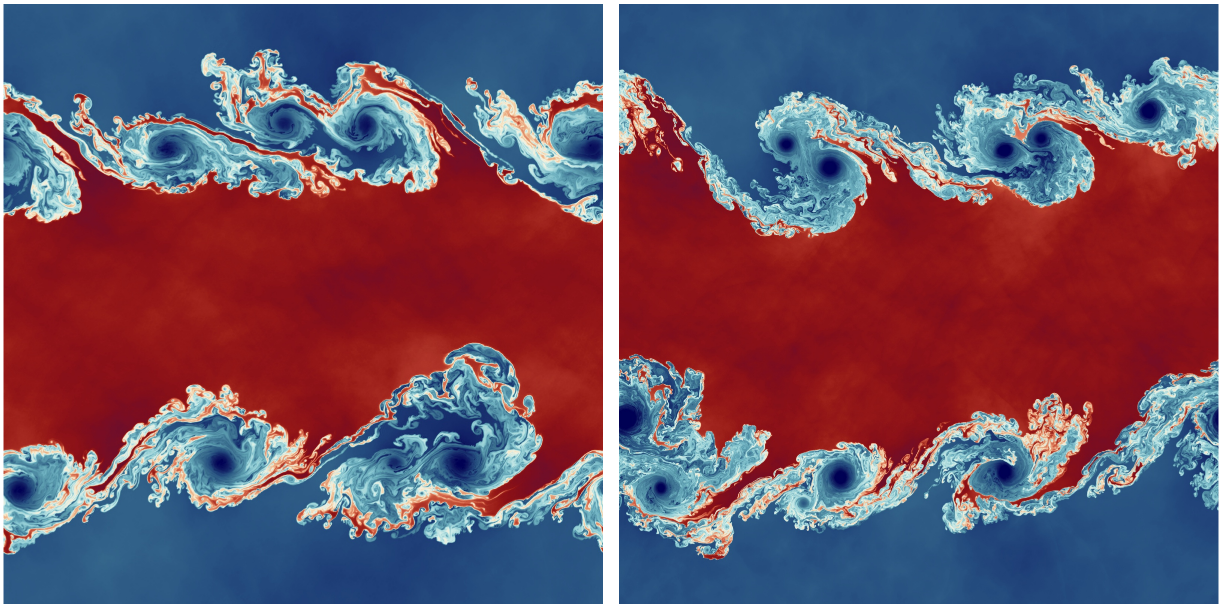
No explicit seeding of the instabilities was performed – the instabilities originate from roundoff errors in the solver. The contours of density at $t = 2$ as predicted by a $\mathbb{P}_4$ FR scheme with various mesh resolutions are shown in Fig. 3.4. The roll-up of the vortices was well-resolved by the approach, with complex small-scale vortical structures beginning to appear with increasing mesh resolution. No spurious oscillations were observed in the vicinity of the discontinuities, and more subtle features such as pressure waves were not excessively dissipated.

3.4.1.6 Mach 800 Astrophysical Jet

As a verification of the positivity-preserving properties of the filter for very extreme conditions, the case of high-speed astrophysical jets is considered. The test case, introduced by Balsara [87], consists of a Mach 800 jet in the presence of an ambient gas. The problem setup is identical to the work of Wu and Shu [88] in which a provably-positive third-order discontinuous Galerkin approach



(a) $N = 100^2$ (b) $N = 200^2$



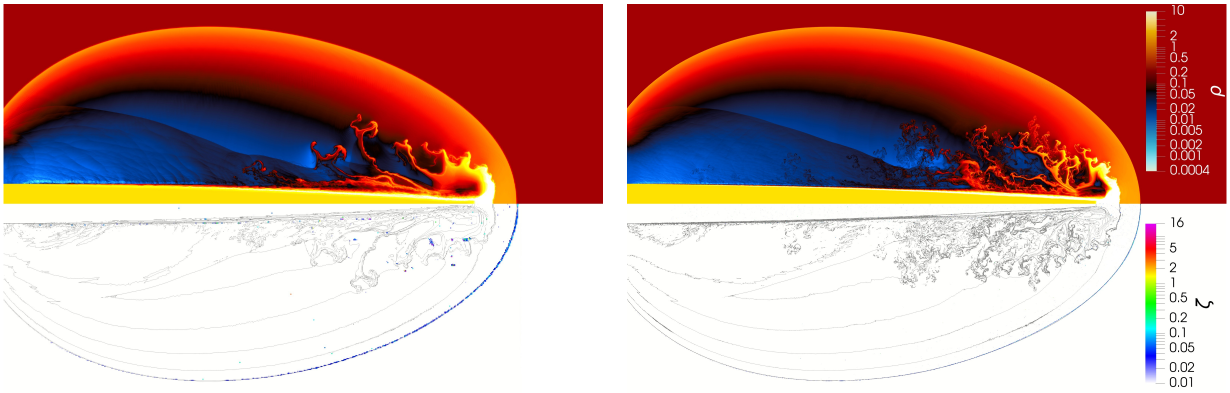
(c) $N = 400^2$ (d) $N = 800^2$

Figure 3.4: Contours of density for the Kelvin-Helmholtz instability problem at $t = 2$ using a $\mathbb{P}_4$ FR approximation with varying mesh resolutions.

is used, although the magnetic field is neglected in this case. A half domain $\Omega = [0, 0.5] \times [0, 1.5]$ is considered with symmetry (slip adiabatic) boundary conditions along the y-axis. The domain is filled with an ambient gas of density 0.1γ, zero velocity, and unit pressure. For the $y = 0$ boundary, the inlet region is defined on $x \leqslant 0.05$, and the solution is set to $\mathbf{q} = [\gamma, 0, 800, 1]^T$ which yields a Mach number of 800 with respect to the inflow gas. The remaining boundary conditions are set as free.

The contours of density at $t = 0.002$ as predicted by a $\mathbb{P}_3$ FR scheme with a coarse (200×600) and fine (800×2400) mesh are shown in Fig. 3.5. Excellent resolution of the leading shock wave was obtained, and the small-scale structures in the vicinity of the cocoon/jet interface were well-resolved even with the coarse mesh. Additionally, the distribution of the filter parameter ζ is shown in Fig. 3.5 overlaid on the isocontours of density. For both the coarse and fine mesh, the filter was primarily active at the leading shock region. Some regions of activation within the cocoon/jet interface were observed for the coarse mesh, but this behavior was reduced with increasing resolution, such that minimal activation away from the leading shock was observed with the fine mesh.

Given a similar case setup, a comparison can be made between the proposed approach on the coarse mesh and the mildly-magnetized results of Wu and Shu [88] (Fig. 6.i), which were obtained by the linear limiting approach of Zhang and Shu [75]. Significantly better resolution was obtained using the entropy filter, particularly with regards to the small-scale features near the cocoon/jet interface, although some of this may be attributed to the stabilizing effect of the magnetic field in their case. However, this suggests that the nonlinear limiting performed by the entropy filter offers noticeable advantages in comparison to a linear limiting approach with minimal computational overhead.



(a) $N = 200 \times 600$

(b) $N = 800 \times 2400$

Figure 3.5: Contours of density (top) and distribution of the filter parameter ζ (bottom) for the Mach 800 astrophysical jet problem at $t = 0.002$ using a $\mathbb{P}_3$ FR approximation with a 200×600 mesh (left) and 800×2400 mesh (right). Contours are reflected about the y-axis. View is reoriented such that the $+y$ direction is shown left-to-right.

3.4.2 Navier–Stokes Equations

3.4.2.1 Taylor–Green Vortex

The entropy filtering approach was then extended to turbulent compressible flows through the Navier–Stokes equations. To verify that the filter does not unnecessarily dissipate small-scale turbulent fluctuations in the absence of shocks, the proposed approach was applied to the subsonic Taylor–Green vortex at a Reynolds number of 1600, a canonical fluid dynamics problem for studying vortex dynamics and turbulent transition and decay [89]. The problem is solved on the periodic domain $\Omega = [-\pi, \pi]^3$ with the initial conditions

$$\mathbf{q}(\mathbf{x}, 0) = \begin{bmatrix} 1 \\ \sin(x)\cos(y)\cos(z) \\ -\cos(x)\sin(y)\cos(z) \\ 0 \\ P_0 + \frac{1}{16}\left(\cos(2x) + \cos(2y)\right)\left(\cos(2z+2)\right) \end{bmatrix},$$

where $P_0 = 1/\gamma M^2$ for a reference Mach number $M = 0.08$. Given the unit density and velocity, the dynamic viscosity μ is set to $1/1600$ to recover a Reynolds number of 1600.

The quantity of interest for this problem is the dissipation rate of the kinetic energy in the flow. The non-dimensional integrated kinetic energy can be defined as

$$K(t) = \frac{1}{V} \int_{\Omega} \frac{1}{2} \rho \mathbf{v} \cdot \mathbf{v} \, \mathrm{d}\mathbf{x}, \tag{3.19}$$

where $V = 8\pi^3$ is the volume of the domain. From this, a dissipation rate based on the kinetic energy can be calculated as

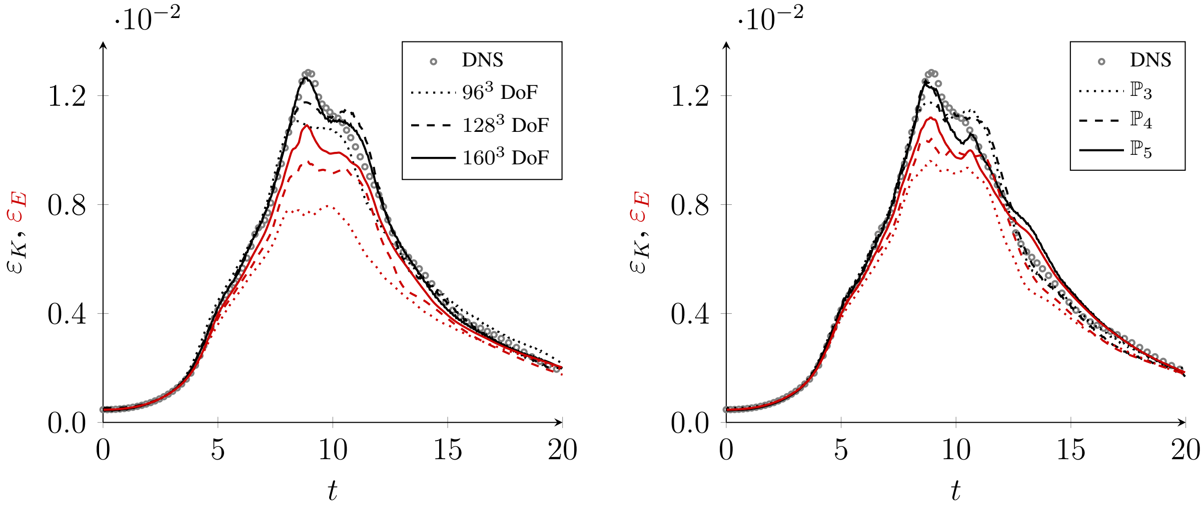
$$\varepsilon_K = \frac{\mathrm{d}K}{\mathrm{d}t}. \tag{3.20}$$

A similar measure of the dissipation can be obtained through a scaled form of the non-dimensional integrated enstrophy, defined as

$$\varepsilon_E = \frac{\beta}{V} \int_{\Omega} \frac{1}{2} \rho \boldsymbol{\omega} \cdot \boldsymbol{\omega} \, \mathrm{d}\mathbf{x}, \tag{3.21}$$

where $\boldsymbol{\omega}$ is the vorticity and $\beta = 2\mu$ is the scaling factor. For purely incompressible flows, these two quantities are equal, but for compressible flows, they differ by the contribution of the deviatoric strain and pressure dilatation to the dissipation. At low Mach numbers, the enstrophy-based dissipation can reasonably approximate the kinetic energy-based dissipation for well-resolved flows.

The prediction of these two quantities as computed by the entropy filtering approach with varying resolution and approximation order is shown in Fig. 3.6 in comparison to the DNS results of van Rees et al. [1]. For visualization of the kinetic energy-based dissipation, a moving-average smoothing operation was performed prior to computing the temporal derivative to reduce oscillations. For a $\mathbb{P}_3$ FR approximation, the number of degrees of freedom was varied from 96^3-160^3. Relatively good agreement was observed between the kinetic energy-based dissipation and the reference data across this entire range of resolution, with the most resolved case showing negligible deviation from the reference. For the enstrophy-based dissipation, the profiles evidently showed convergence to the reference with increasing resolution, and for a given resolution, showed similar results to unfiltered approaches [90]. When fixing the degrees of freedom to $\sim$120 and varying the approximation order, improvements in the prediction of both the kinetic energy- and enstrophy-based dissipation were observed with increasing approximation order. Additionally, nearly identical results were obtained without filtering (see Trojak et al. [91], Fig. 3b), which supports the presumption that the filter is predominantly inactive for turbulent flows without discontinuities.



(a) $\mathbb{P}_3$, 96^3-160^3 DoF

(b) 120 DoF, $\mathbb{P}_3$-$\mathbb{P}_5$

Figure 3.6: Dissipation measured by kinetic energy (black) and enstrophy (red) for the Taylor-Green vortex using a $\mathbb{P}_3$ FR approximation with varying DoF (left) and $\sim$120 DoF with varying approximation order (right). DNS results of van Rees et al. [1] (obtained by private communication) shown for reference.

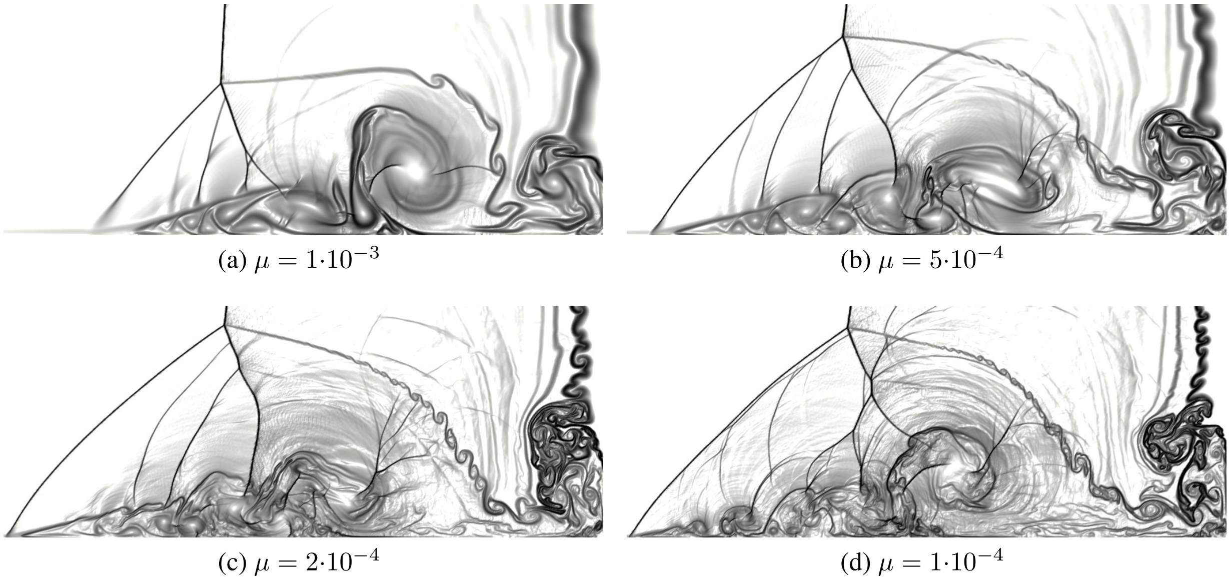
3.4.2.2 Viscous Shock Tube

To assess the efficacy of the entropy filtering approach for predicting shock-boundary layer interactions, a two-dimensional viscous shock tube was simulated. The problem, introduced by

Daru and Tenaud [92], is solved on the half-domain $\Omega = [0,1] \times [0,0.5]$ with the initial conditions

$$\mathbf{q}(\mathbf{x},0) = \begin{cases} \mathbf{q}_l, & \text{if } x \leqslant 0.5, \\ \mathbf{q}_r, & \text{else,} \end{cases} \quad \text{given} \quad \mathbf{q}_l = \begin{bmatrix} 120 \\ 0 \\ 0 \\ 120/\gamma \end{bmatrix}, \quad \mathbf{q}_r = \begin{bmatrix} 1.2 \\ 0 \\ 0 \\ 1.2/\gamma \end{bmatrix}.$$

For the top wall, a slip adiabatic wall boundary condition is applied to enforce symmetry. No slip adiabatic wall conditions are applied for the remaining walls. This case contains the standard features of the Riemann problem, namely a rarefaction wave, a contact discontinuity, and a shock wave, in addition to viscous interactions between these features and the wall. This interaction forms a complex lambda shock impinging on a viscous boundary layer, and as a result, presents a suitable test case for shock-boundary layer interactions.


(a) $\mu = 1{\cdot}10^{-3}$

(b) $\mu = 5{\cdot}10^{-4}$

(c) $\mu = 2{\cdot}10^{-4}$

(d) $\mu = 1{\cdot}10^{-4}$

Figure 3.7: Schlieren-type representation of the density gradient norm for the viscous shock tube problem using a $\mathbb{P}_4$ FR approximation with a 800×400 mesh for varying values of the dynamic viscosity μ. Contours shown at $t = 1$ on the subregion $[0.35, 1] \times [0, 0.25]$.

Numerical Schlieren results at $t = 1$ computed using a $\mathbb{P}_4$ FR approximation on a uniform 800×400 mesh are shown in Fig. 3.7 for varying values of the dynamic viscosity μ. As the Reynolds number was increased, progressively smaller-scale features became evident in the flow, particularly along the contact line and in the separation region. The discontinuities as well as the small vortical structures in the flow were both well-resolved. The predicted results show relatively good agreement with the results of Guermond et al. [93], although some small-scale features at higher Reynolds numbers were notably more pronounced in the present work. These results suggest that the entropy filter used in conjunction with an operator splitting approach is an effective tool for approximating mixed hyperbolic-parabolic systems that exhibit discontinuities as well as complex small-scale flow structures.

3.4.2.3 Transonic Delta Wing

As a final evaluation of the proposed approach for complex aeronautical applications including three-dimensional high Reynolds number flows computed on unstructured meshes, the test case of large eddy simulation around a transonic VFE-2 delta wing was considered. The geometry, introduced by Chu and Luckring [94], consists of a sharp leading edge delta wing with a sweep angle of 65° and a thickness to root chord ratio of 3.4%. To match the experimental setup of Konrath et al. [2], the Reynolds number is set to $3{\cdot}10^6$ based on a unit root chord (i.e., $2{\cdot}10^6$ based on the mean aerodynamic chord), and the freestream Mach number is set to 0.8, yielding locally supersonic turbulent flow on the suction side of the wing. For these geometry and flow conditions, the separation and roll-up of the primary vortex is triggered by the sharp leading edge which immediately transitions into turbulence. Secondary and tertiary vortices are then subsequently formed through near-wall viscous interactions. An angle of attack of 20.5° was chosen as these conditions yield strongly nonlinear behavior on the suction side as the primary vortex is on the precipice of vortex burst.

A second-order unstructured mesh was generated by extruding a triangular surface mesh in the near wall region and filling the remaining domain with tetrahedral elements. The resolution of the surface mesh was approximately uniform with an average edge length of $h = 0.003$, and

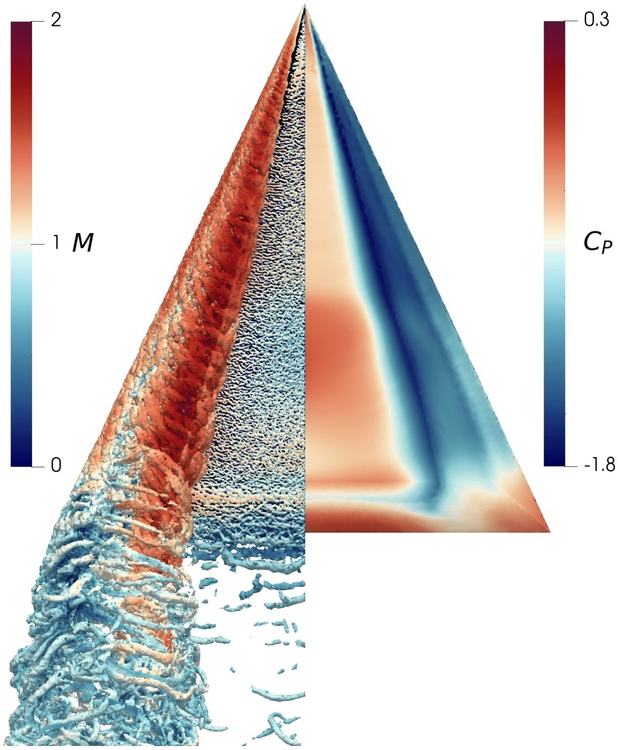
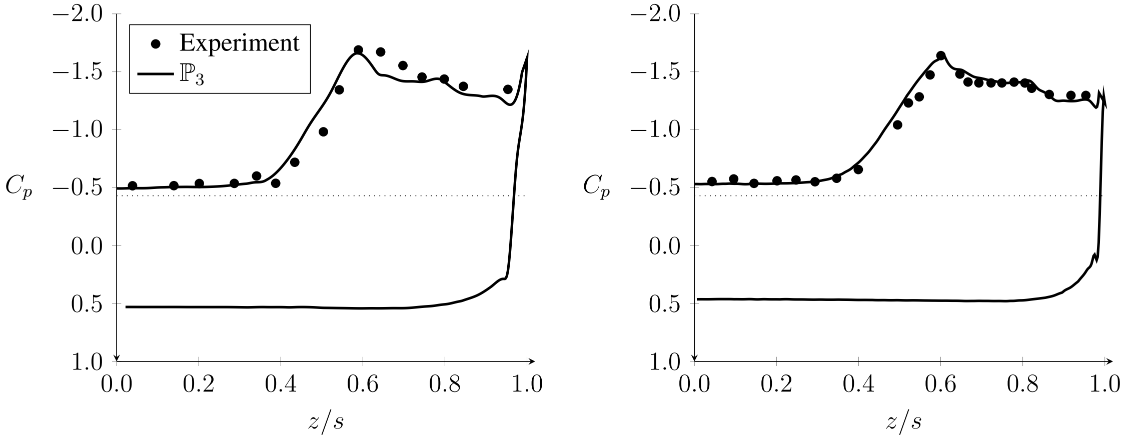


Figure 3.8: Isosurface of instantaneous Q-criterion colored by local Mach number (left) and time-averaged surface pressure coefficient contours (right) for the transonic VFE-2 delta wing computed using a $\mathbb{P}_3$ FR approximation.

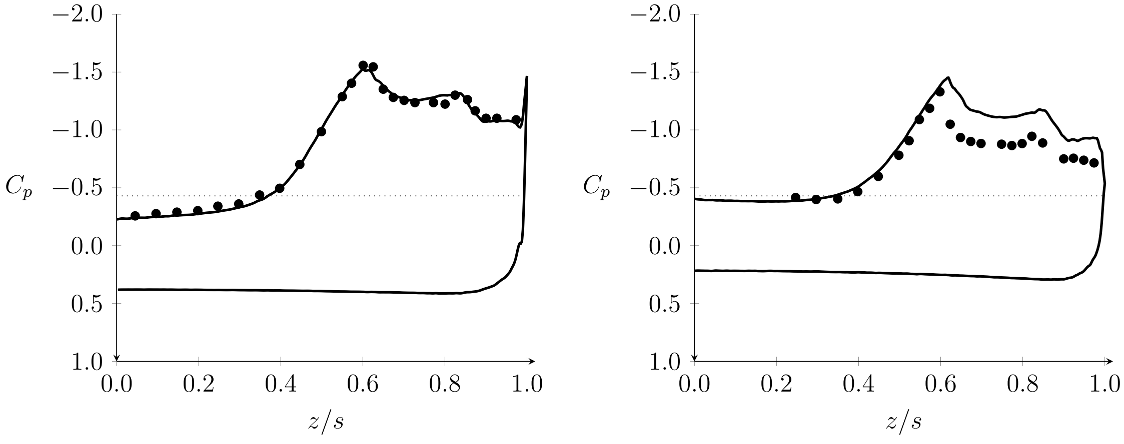
the near-wall prismatic region was extruded over a length of 0.02. To simplify the meshing, the sting and fairing from the experimental model were neglected. Since the separation and roll-up of the primary vortex is primarily governed by inviscid effects due to the sharp leading edge [95], the wall-normal resolution requirements are relaxed in comparison to standard wall-resolved large eddy simulation [96]. As such, the wall-normal resolution was set such that the y^+ value of the first solution point away from the wall was approximately 30 (i.e., the element y^+ was $\sim$120) based on equivalent flat plate conditions matching the freestream flow and root chord. In the separated region above the suction side, the resolution of the tetrahedral mesh was such that the average edge length was approximately $h = 0.01$. The overall mesh consisted of 4.1 million elements, yielding a total of 104 million degrees of freedom per component with a $\mathbb{P}_3$ approximation. At the wing, no slip adiabatic wall boundary conditions were applied, and characteristic Riemann invariant boundary conditions were used for the farfield.

At startup, the approximation order and Reynolds number were progressively increased until the initial transients were convected away. Averaging was then performed for a time period corresponding to five flows over root chord. The results of a $\mathbb{P}_3$ FR approximation are presented in Fig. 3.8, showing instantaneous Q-criterion isosurfaces colored by local Mach number and time-averaged surface pressure coefficient contours. The Q-criterion isosurfaces showed the canonical vortex structure of delta wing flows with the primary and secondary vortices evident in the flow. Due to the relatively high angle of attack, a local Mach number of approximately 2 was observed near the primary vortex core, and indication of the onset of vortex burst was seen near the trailing edge. Furthermore, the average surface pressure coefficient contours clearly showed the presence of a primary and secondary vortex, and good agreement was observed between the predicted results and the experimental results of Konrath et al. [2].

The surface pressure coefficient at various streamwise locations is presented in Fig. 3.9 in comparison to the experimental data. At $x/c = 0.2$, 0.4, and 0.6, excellent agreement was observed, both in terms of the location and the magnitude of the primary and secondary vortex pressure peaks. At $x/c = 0.8$, the location and magnitude of the primary vortex pressure peak was well-resolved, but a slight overprediction in the magnitude of the secondary vortex pressure peak was seen. This effect may be attributed to the strong sensitivity of the flow regime prior to vortex burst in conjunction with the slight discrepancies between the experimental and computational geometry at the trailing edge of the wing. Overall, the computational results were in-line with the experimental data, indicating that the entropy filtering approach can be successfully used for complex supersonic turbulent flows.



(a) $x/c = 0.2$ (b) $x/c = 0.4$



(c) $x/c = 0.6$ (d) $x/c = 0.8$

Figure 3.9: Surface pressure coefficient at various streamwise locations for the transonic VFE-2 delta wing computed using a $\mathbb{P}_3$ FR approximation. Experimental results of Konrath et al. [2] shown for reference. Dotted line denotes the sonic pressure coefficient. Spanwise extent is normalized by the local semispan s.

3.5 Conclusions

A novel adaptive filtering approach for shock capturing in discontinuous spectral element methods was presented in this chapter. The proposed entropy filtering approach formulates the filtering operation as an element-wise scalar optimization problem designed to enforce constraints such as positivity and a local discrete minimum entropy principle. As a result, this method is free of problem-dependent tunable parameters and can be easily implemented on arbitrary unstructured meshes with minimal computational effort. To evaluate the efficacy of the proposed filtering approach, a series of numerical experiments was performed on multi-dimensional hyperbolic and mixed hyperbolic-parabolic conservation laws such as the Euler and Navier–Stokes equations. Even for extremely strong shocks, sub-element resolution of the discontinuities was generally obtained with minimal spurious oscillations, and high-order accuracy was recovered for smooth solutions. Furthermore, for more complex problems including shock-vortex interactions and compressible turbulent flows, the filter was able to robustly resolve discontinuities without unnecessarily dissipating small-scale vortical structures in the flow. These results indicate that the proposed entropy filtering approach can be an effective and robust technique for shock capturing in discontinuous spectral element approximations of hyperbolic and mixed hyperbolic-parabolic conservation laws.

4. POSITIVITY-PRESERVING AND CONSERVATIVE APPROXIMATION OF THE POLYATOMIC BOLTZMANN–BGK EQUATION

The development and application of computational fluid dynamics methods have, for the most part, focused on solutions of the Euler and Navier–Stokes equations to give insight into the dynamics of fluid flow. These approaches naturally rely on the notion that the continuum assumption can be regarded as valid for the given fluid, such that transport phenomena can be adequately described solely through the macroscopic flow variables (e.g., density, velocity, etc). However, for many problems of interest, such as rarefied gases and hypersonic flows, the continuum assumption starts to break down as collisions between molecules become less frequent and regions of the flow significantly deviate from thermodynamic equilibrium. In such cases, it becomes necessary to revert to the governing equations of molecular gas dynamics which underpin the macroscopic behavior of the fluid. These methods, which derive from the kinetic theory of gases, can offer a more detailed description of non-equilibrium systems and flows outside of the continuum regime while seamlessly recovering the hydrodynamic equations in the asymptotic limit.

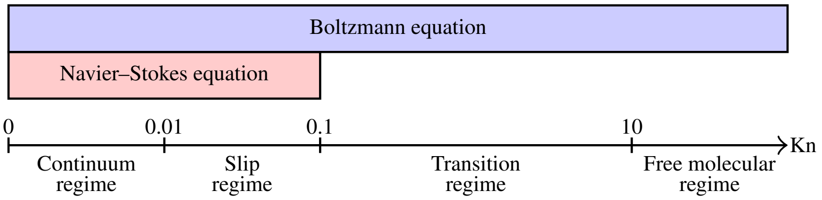


Figure 4.1: Validity of the governing equations across flow regimes of varying degrees of rarefaction characterized by the Knudsen number Kn.

To characterize gas dynamics at the molecular level, the Boltzmann equation can be used to provide a statistical description of particle transport and collision [32]. This significantly more general governing equation opens up the possibility of simulating a wider variety of flow regimes, shown in Fig. 4.1 with respect to the Knudsen number, the ratio of the particle mean free path to the characteristic length scale. In comparison to the Euler and Navier–Stokes equations, which solve for the nonlinear transport of the conserved flow variables, the Boltzmann equation simply solves for a scalar quantity, the *distribution function* or *probability density function*, for which the transport can be represented linearly. The complexity of this governing equation instead comes from two main sources: dimensionality and suitable modeling of the interactions among molecules. With regard to the former, the distribution function can be considered as a probability measure for particles existing at a given location with a certain velocity. As such, it must be defined over physical space and velocity space (and time), which can require discretization of up to 6-dimensional spaces. This problem is further compounded when attempting to model additional internal degrees of freedom (e.g., rotation, vibration, etc.), which can require discretizations of even higher dimensionality [97]. Furthermore, the proper discretization of the velocity space, which is, in theory, unbounded, is an open problem, with varying methods offering differing advantages in terms of accuracy, stability, and computational cost [98]. For the latter, the approximation of particle interactions in velocity space can be the most costly and algorithmically complex part of the Boltzmann equation, requiring integration over spaces of even higher dimension than the velocity space itself. This collision approximation also can, in many scenarios, cause the governing equation to become exceedingly stiff, further complicating attempts at efficient computation.

To reduce the computational cost associated with the direct computation of the collision term, an approximate model of the collision process can instead be used. One such approach, the BGK (Bhatnagar–Gross–Krook) [36] model, simplifies this process to a relaxation of a gas in non-equilibrium tending towards thermodynamic equilibrium, avoiding much of the computational cost associated with computing the collision integrals. This simplified model retains much of the behavior of the full collision operator, including its entropy-satisfying properties embedded in Boltzmann's H-theorem and convergence to the continuum approximation in the asymptotic limit, albeit with some corrections necessary to recover the proper transport coefficients. One additional drawback of the BGK model is that for general implementations, conservation of the macroscopic flow variables is only guaranteed if the integration in velocity space can be carried out exactly, which is usually not the case for discrete representations [99]. A common approach to remedy this issue is to resolve the velocity space finely enough such that the integration errors are negligible [100], but this comes at the expense of even further increasing the computational cost of solving the system.

A potential improvement in the cost of solving the Boltzmann equation can be obtained through the use of a higher-fidelity spatial discretization. The typical schemes of choice for implementations of the Boltzmann equation on general unstructured grids are low-order finite volume methods [3, 101] which are generally simple to implement and numerically robust. However, these benefits come with the drawback of low accuracy per degree of freedom and excessive numerical dissipation that makes them prohibitively expensive for scale-resolving simulations. Instead, the use of high-order DSEM may offer many advantages due to their accuracy and efficiency, and their extension to the approximation of the Boltzmann equation has been attempted in the recent works of Xiao [33] and Jaiswal [34] with very promising results. The drawback of these methods is that they are more algorithmically complex and typically far less robust, which can limit their potential as a method of reducing the computational cost associated with the Boltzmann equation.

The objective of this chapter is therefore to propose, develop, and validate an efficient numerical approach for accurately and robustly solving the Boltzmann–BGK equation with the goal of drastically reducing the computational cost and enabling its application to three-dimensional problems that were hitherto intractable. We utilize the FR scheme augmented with a positivity-preserving limiter that ensures the positivity of the distribution function as well as the macroscopic density/pressure. This spatial discretization is paired with a nodal discretization of the velocity space utilizing a modified form of the discrete velocity model of Mieussens [3] which guarantees the scheme remains conservative and well-balanced regardless of the resolution. Furthermore, we extend the approach to polyatomic molecules and general constitutive laws through an appropriate model for the internal degrees of freedom. The applicability of the proposed approach is shown in a series of complex numerical experiments, ranging from shock-dominated flows computed on unstructured grids without any numerical shock capturing algorithm to direct numerical simulation of three-dimensional compressible turbulent flows, the latter of which is, to the author's knowledge, the first instance of such a flow computed by directly solving the Boltzmann equation.

To this end, the organization of this chapter is as follows. Some preliminaries on the various properties of the Boltzmann–BGK equation and the associated discretization methods are presented in Section 4.1. The methodology of the proposed approach is then presented in Section 4.2, followed by the details of the numerical implementation in Section 4.3. The results of various numerical experiments are then shown in Section 4.4, and conclusions are drawn in Section 4.5.

4.1 Preliminaries

The work in this chapter pertains to approximations of the *polyatomic* Boltzmann equation with arbitrary constitutive laws, given as

$$\partial_t f(\mathbf{x}, \mathbf{v}, \zeta, t) + \mathbf{v}{\cdot}\nabla f = \mathcal{C}(f, f'), \tag{4.1}$$

where $\mathbf{x} \in \Omega^{\mathbf{x}}$ is the physical space for some physical domain $\Omega^{\mathbf{x}} \subseteq \mathbb{R}^d$ and spatial dimension d, $\mathbf{v} \in \Omega^{\mathbf{v}}$ is the associated velocity space for some velocity domain $\Omega^{\mathbf{v}} \in \mathbb{R}^m$ and velocity dimension $m \geq d$, $\zeta \in \mathbb{R}^+$ is the internal energy, $f(\mathbf{x}, \mathbf{v}, \zeta, t) \in \mathbb{R}$ is a scalar particle distribution function, and $\mathcal{C}(f, f')$ is some collision operator that accounts for intermolecular interactions [32]. This distribution function, describing the microscopic state of the system, gives a measure of the probability density of particles existing at a given location $\mathbf{x}$ traveling at a given velocity $\mathbf{v}$ (the combination of which is referred to as a point in *phase space*) with an internal energy ζ. The internal energy is given as a continuous variable that takes into account any additional degrees of freedom such as rotational and vibrational modes or unresolved velocity components [97]. From this distribution, one can gather information about the unique macroscopic state of the system (i.e., the conserved flow variables $\mathbf{Q}(\mathbf{x}, t)$) through its moments,

$$\mathbf{Q}(\mathbf{x}, t) = [\rho, \rho\mathbf{V}, E]^T = \int_{\mathbb{R}^d}\int_0^{\infty} f(\mathbf{x}, \mathbf{v}, \zeta, t)\, \boldsymbol{\psi}(\mathbf{v}, \zeta)\, \mathrm{d}\zeta\, \mathrm{d}\mathbf{v} \tag{4.2}$$

where $\boldsymbol{\psi}(\mathbf{v}, \zeta) := [1, \mathbf{v}, (\mathbf{v} \cdot \mathbf{v})/2 + \zeta]^T$ is the vector of collision invariants. We utilize a slightly modified notation here where the lowercase symbol $\mathbf{v}$ refers to the microscopic velocity and the uppercase symbol $\mathbf{V}$ refers to the macroscopic velocity, such that $\rho\mathbf{V}$ is the macroscopic momentum vector.

4.1.1 Bhatnagar-Gross-Krook Operator

Due to its high-dimensional nature, the computational cost of solving the scalar Boltzmann equation is considerably higher than the associated conservation laws for the conserved variables $\mathbf{Q}(\mathbf{x}, t)$. This is further exacerbated by the complexity of the collision operator $C(f, f')$, which in its full form can require integration over spaces of even higher dimension. To ameliorate this cost, an approximate form of the collision operator was introduced by Bhatnagar et al. [36] by modeling the collision of particles as a multidimensional relaxation process of a gas tending towards thermodynamic equilibrium. The collision operator for the BGK (Bhatnagar–Gross–Krook) model can be given as

$$\mathcal{C}(f, f') \approx \frac{g(\mathbf{x}, \mathbf{v}, \zeta, t) - f(\mathbf{x}, \mathbf{v}, \zeta, t)}{\tau}, \tag{4.3}$$

where $g(\mathbf{x}, \mathbf{v}, \zeta, t)$ is the equilibrium distribution function and τ is the collision time scale. It can be shown that if the following compatibility condition is satisfied,

$$\int_{\mathbb{R}^d} \int_0^\infty g(\mathbf{x}, \mathbf{v}, \zeta, t)\ \boldsymbol{\psi}(\mathbf{v}, \zeta)\ \mathrm{d}\zeta\ \mathrm{d}\mathbf{v} = \int_{\mathbb{R}^d} \int_0^\infty f(\mathbf{x}, \mathbf{v}, \zeta, t)\ \boldsymbol{\psi}(\mathbf{v}, \zeta)\ \mathrm{d}\zeta\ \mathrm{d}\mathbf{v}, \tag{4.4}$$

the resulting system is conservative with respect to the conserved flow variables.

For the polyatomic case, the equilibrium distribution function is given as

$$g(\mathbf{x}, \mathbf{v}, \zeta, t) = g_{\mathbf{v}}(\mathbf{x}, \mathbf{v}, t) \times g_\zeta(\mathbf{x}, \zeta, t), \tag{4.5}$$

where $g_{\mathbf{v}}$ is the equilibrium distribution function of the monatomic case and $g_\zeta(\mathbf{x}, \zeta, t)$ is an additional term used to augment the equilibrium distribution function to account for internal degrees of freedom. We will assume that $g_\zeta(\mathbf{x}, \zeta, t)$ is properly normalized such that its inclusion does not affect the moments, i.e.,

$$\int_{\mathbb{R}^d} \int_0^\infty g(\mathbf{x}, \mathbf{v}, \zeta, t)\ \boldsymbol{\psi}(\mathbf{v}, \zeta)\ \mathrm{d}\zeta\ \mathrm{d}\mathbf{v} = \int_{\mathbb{R}^d} g_{\mathbf{v}}(\mathbf{x}, \mathbf{v}, t)\ \boldsymbol{\psi}(\mathbf{v})\ \mathrm{d}\mathbf{v} = \mathbf{Q}(\mathbf{x}, t). \tag{4.6}$$

From the H-theorem, the monatomic equilibrium state is generally taken to be the state that minimizes the entropy $H(z)$, i.e.,

$$g_{\mathbf{v}} = \arg\min_z H(z), \tag{4.7}$$

where

$$H(z) = \int_{\mathbb{R}^d} z \log(z)\ \mathrm{d}\mathbf{v}. \tag{4.8}$$

In the *continuous* case, the distribution function in velocity space that minimizes this entropy and satisfies Eqs. (4.2) and (4.5) is a Maxwellian of the form

$$g_{\mathbf{v}}(\mathbf{x}, \mathbf{v}, t) = \frac{\rho(\mathbf{x}, t)}{[2\pi\theta(\mathbf{x}, t)]^{d/2}} \exp\left[-\frac{\|\mathbf{v} - \mathbf{V}(\mathbf{x}, t)\|_2^2}{2\theta(\mathbf{x}, t)}\right], \tag{4.9}$$

where $\theta = P/\rho$ is a scaled temperature. We utilize the notation $\|\cdot\|_2^2$ to denote the squared norm along the spatial dimension and $g(\mathbf{Q}(\mathbf{x}, t))$ to denote the Maxwellian corresponding to the conserved variables $\mathbf{Q}(\mathbf{x}, t)$.

4.1.2 Internal Degrees of Freedom

In its standard form, the Boltzmann–BGK approach assumes a monatomic molecule in which the only degrees of freedom are its translational components. This limits the number of degrees of freedom, n, to the dimensionality of the velocity space (i.e., $n = m$), which in turn, constrains the specific heat ratio by the relation $\gamma = 1 + 2/n$. This constraint on the degrees of freedom imposes several difficulties for the application of the approach to practical problems. First, there is significant interest in the simulation of real gases, many of which are polyatomic in nature, for which the standard approach cannot be used. Secondly, even for monatomic molecules, certain flow conditions such as high temperatures can induce additional kinetic effects such as vibration which cannot be accounted for without additional degrees of freedom. Lastly, although this drawback is less critical, the simulation of monatomic flows with spatial homogeneity (e.g., one-dimensional flows) requires the discretization of higher-dimensional velocity spaces to recover the proper specific heat ratio, unnecessarily increasing the computational cost.

To simultaneously account for the additional degrees of freedom of polyatomic gases, ancillary kinetic modes such as rotation and vibration, and arbitrary constitutive laws, Baranger et al. [97] introduced a model for internal degrees of freedom for the molecule. In this approach, an additional dimension is considered to account for the internal energy of the molecule ζ, and the equilibrium distribution function is modified by the factor

$$g_\zeta(\mathbf{x}, \zeta, t) = \Lambda(\delta) \left(\frac{\zeta}{\theta(\mathbf{x}, t)} \right)^{\frac{\delta}{2} - 1} \frac{1}{\theta(\mathbf{x}, t)} \exp \left(- \frac{\zeta}{\theta(\mathbf{x}, t)} \right), \tag{4.10}$$

where $\delta \geq 0$ is the number of internal degrees of freedom and $\Lambda(\delta) = 1/\Gamma(\delta/2)$ is a normalization factor to ensure Eq. (4.6) is satisfied. With this model, the total degrees of freedom are then $n = m + \delta$, such that a specific heat ratio of $\gamma = 1 + 2/(m + \delta)$ can be recovered. For the monatomic case, the internal degrees of freedom are simply set as $\delta = 0$ and the ζ domain is neglected. Note that δ does not necessarily have to be an integer nor a constant (see Baranger et al. [97], Section 3), which allows for the implementation of arbitrary constitutive laws for multi-physics applications (e.g., temperature-dependent internal degrees of freedom).

4.1.3 Velocity and Internal Energy Spaces

Discretizations for the velocity space generally fall into one of two approaches: nodal form or moment form. In the former approach, the infinite velocity space is truncated onto a finite domain and represented by a set of discrete points. The associated transport equation, Eq. (4.1), is then solved for each one of these discrete points. In the latter approach, the distribution function and the velocity space are represented by a set of global basis functions, and the transport equation is solved for the entire velocity space using an appropriate treatment of the inner product in Eq. (4.1). (e.g., orthogonal projection). Due to their orthogonality with respect to the measure $\exp(-x^2)$ on the real domain, Hermite polynomials are a natural choice for the basis functions, typically allowing for a more computationally efficient representation but at the expense of robustness, particularly for flows deviating from equilibrium [98]. However, for the internal energy domain, finding a set of suitable global basis functions becomes less trivial.

The focus of this work is on the former approach – nodal discretizations of the velocity and internal energy spaces – as this method is the most general, encompassing arbitrary distribution functions and enabling the straightforward implementation of additional physics such as the inclusion of body forces. For this approach, the discretization for the resulting system of equations becomes greatly simplified, reducing to a set of linear advection equations that are essentially independent across the spatial and velocity/internal energy domains and share a dependency only through the nonlinear source term. However, to accurately resolve the underlying physics, it must be ensured that the resolution is fine enough and the truncated velocity domain encompasses enough of the predominant behavior of the distribution function. Therefore, the proper discretization of the velocity/internal energy domains becomes a critical factor in the accuracy of the Boltzmann approach, and, as will be shown in Section 4.2.4, can also affect the numerical properties of the scheme such as conservation and well-balancing.

4.1.4 Limiting Behavior

The Boltzmann–BGK model is a significantly more general approach than the equations governing fluid flow in the continuum limit, and it can be shown to recover these hydrodynamic equations in the asymptotic limit. The primary non-dimensional quantity characterizing the Boltzmann equation is the Knudsen number, defined as the ratio

$$Kn = \frac{\lambda}{L_{\text{ref}}}, \tag{4.11}$$

where L_{ref} is the characteristic length scale and λ is the particle mean free path. In the limit as $Kn \to 0$, the continuum approximation can be regarded as valid, whereas for higher Kn, the flow enters the rarefied regime. The Knudsen number can be related to Mach number M and Reynolds number Re as

$$Kn = \sqrt{\frac{\gamma\pi}{2}}\frac{M}{Re}. \tag{4.12}$$

The Knudsen number can similarly be expressed in terms of the macroscopic flow variables as

$$Kn = \sqrt{\frac{\gamma\pi}{2}}\frac{\mu}{\rho c_s L_{\text{ref}}}. \tag{4.13}$$

In the low Knudsen number limit, the BGK approximation converges to the hydrodynamic limit governed by the Euler/Navier–Stokes equations with the dynamic viscosity given as

$$\mu = \tau P, \tag{4.14}$$

which defines the collision time in terms of the Knudsen number and the macroscopic flow quantities as

$$\tau = Kn\sqrt{\frac{2}{\gamma\pi}}\frac{\rho c_s L_{\text{ref}}}{P} = \sqrt{\frac{2\gamma}{\pi}}\frac{Kn L_{\text{ref}}}{c_s}. \tag{4.15}$$

In the low Mach number limit, the variations in pressure become small, such that for a fixed collision time, the dynamic viscosity remains essentially constant. However, for higher Mach numbers where the pressure, and therefore the viscosity, is expected to vary more strongly, the collision time can be set adaptively based on the macroscopic flow state to recover a more physically consistent viscosity coefficient that varies as a function of temperature instead of pressure [3]. This method is briefly explored in this work. Furthermore, due to the use of a single relaxation time, momentum relaxation and thermal relaxation are assumed to be identical, such that the model recovers a unit Prandtl number ($Pr = 1$). Relatively simple modifications to the equilibrium distribution function can be performed to recover other values for the Prandtl number [102, 103], but they are omitted in this work.

4.2 Methodology

A description of the proposed numerical approach for discretizing the Boltzmann–BGK equation is presented in this section, including details on the positivity-preserving high-order spatial discretization, the discretely conservative velocity and internal energy discretizations, boundary conditions for the system, and a guideline for the required resolution levels. A schematic of this discretization is presented in Fig. 4.2.

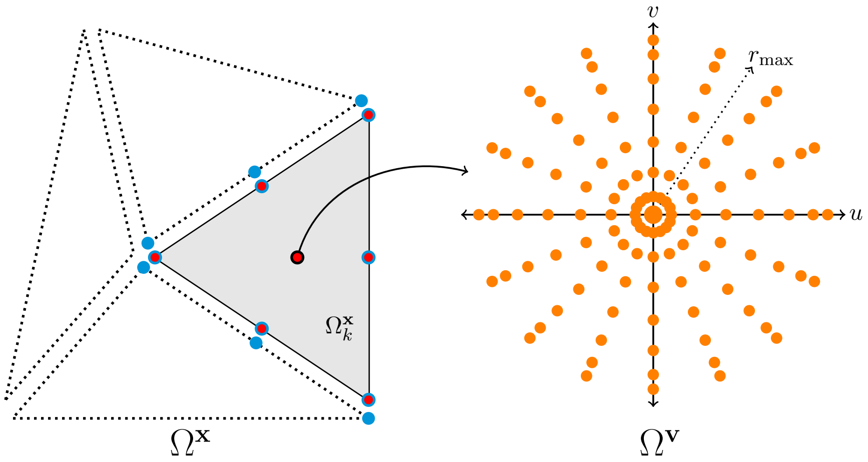


Figure 4.2: Schematic of a two-dimensional phase space discretization using an unstructured spatial domain $\Omega^{\mathbf{x}}$ with $\mathbb{P}_2$ elements and a velocity domain $\Omega^{\mathbf{v}}$ with $N_r = 8$, $N_\phi = 16$ and no internal degrees of freedom. Circles denote the spatial solution nodes (red), interface flux nodes (blue), and velocity space nodes (orange), respectively.

4.2.1 Spatial Discretization

For a fixed location $\mathbf{v}_0$ in velocity space, the evolution of the particle distribution function is equivalent to a linear advection equation with a nonlinear source term S. With a slight abuse of notation, we let $f(\mathbf{x}, \zeta, t) = f(\mathbf{x}, \mathbf{v}_0, \zeta, t)$ for some arbitrary $\mathbf{v}_0 \in \Omega^{\mathbf{v}}$, which yields the transport equation

$$\partial_t f(\mathbf{x}, \zeta, t) + \boldsymbol{\nabla} \cdot \mathbf{F}(f) = S, \qquad \mathbf{f}(f) = \mathbf{v}_0 f. \tag{4.16}$$

Here, $\mathbf{v}_0$ represents some constant advection velocity and $\mathbf{f}(f)$ is the flux. Due to the linearity of the flux, the discontinuous flux $\mathbf{f}^D(\mathbf{x})$ can be simply represented as

$$\mathbf{f}^D(\mathbf{x}) = \mathbf{v}_0 f(\mathbf{x}). \tag{4.17}$$

Due to the point-wise equivalency of the Boltzmann equation to the linear advection equation, the common interface flux can be computed using a simple upwinding approach as

$$\overline{F}_i = \begin{cases} v_n f_i^-, & \text{if } v_n > 0 \\ v_n f_i^+, & \text{else,} \end{cases} \tag{4.18}$$

where $v_n = \mathbf{v}_0 \cdot \mathbf{n}_i$. With this discretization and an approximation of the source term S, the system can be evolved using a suitable temporal integration approach.

4.2.2 Velocity Discretization

For the velocity discretization, the velocity space was truncated onto a finite domain $\Omega^{\mathrm{v}} \subset \mathbb{R}^d$ and represented nodally by N_v discrete points, shown on the right-hand side of Fig. 4.2. As the distribution functions tend to have a radially-symmetric nature in many cases, particularly for flows near equilibrium, it is convenient for $m > 1$ to represent the space in polar/spherical coordinates as

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} r\cos\phi + U_0 \\ r\sin\phi + V_0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} r\sin\psi\cos\phi + U_0 \\ r\sin\psi\sin\phi + V_0 \\ r\cos\psi + W_0 \end{bmatrix}, \tag{4.19}$$

where $r \in (0, r_{\max}]$ is the radial distance for some maximum radial extent $r_{\max}$, $\phi \in [0, 2\pi)$ is the polar angle, $\psi \in [0, \pi)$ is the azimuthal angle, and $\mathbf{V}_0 = [U_0, V_0, W_0]$ is a vector of velocity offsets. The offsets were computed as the component-wise average of the minima and maxima of the initial macroscopic velocities in the domain.

$$\mathbf{V}_0 = \frac{1}{2}\left(\max_{\mathbf{x}\in\Omega^{\mathbf{x}}}\left[\mathbf{V}(\mathbf{x},0)\right] + \min_{\mathbf{x}\in\Omega^{\mathbf{x}}}\left[\mathbf{V}(\mathbf{x},0)\right]\right). \tag{4.20}$$

The two primary parameters in the velocity space discretization are the radial extent of the domain and the distribution of the nodes. For the former, the choice of the radial extent plays a large role in the properties of the scheme – too small of an extent can cause inaccurate predictions of the underlying physics due to the loss of information regardless of the resolution of the discretization, whereas too large of an extent can degrade the efficiency of the discretization while also restricting the maximum permissible explicit time step due to unnecessarily large particle velocities. As described in Evans et al. [100], a standard approach is to prescribe the radial extent as some factor of the thermal velocity (i.e., speed of sound). In this work, we take a similar approach but apply an *a priori* case-dependent metric to automatically account for variation in the velocity distributions between problems. The radial extent is computed as

$$r_{\max} = k\max_{\mathbf{x}\in\Omega^{\mathbf{x}}}\left[c_s(\mathbf{x},0)\right] + \frac{1}{2}\|\delta\mathbf{V}\|_2, \tag{4.21}$$

where $\|\delta\mathbf{V}\|_2$ is the component-wise maximum difference in the macroscopic velocities in the domain, i.e.,

$$\delta\mathbf{V} = \max_{\mathbf{x}\in\Omega^{\mathbf{x}}}\left[\mathbf{V}(\mathbf{x},0)\right] - \min_{\mathbf{x}\in\Omega^{\mathbf{x}}}\left[\mathbf{V}(\mathbf{x},0)\right]. \tag{4.22}$$

The parameter k is chosen such that the contribution of the distribution function outside of the domain is negligible. Therefore, k is computed such that the relative magnitude of the distribution function outside of the velocity domain compared to the maximum is of $\mathcal{O}(\epsilon_{\mathbf{v}})$, where $\epsilon_{\mathbf{v}}$ is some small user-defined constant to be presented in Section 4.3. If the distribution function within the domain is assumed to be initially Maxwellian, then k can be analytically computed as

$$k = \sqrt{-\frac{2}{\gamma}\log(\epsilon_{\mathbf{v}})}, \tag{4.23}$$

which guarantees that the relative magnitude of the initial distribution function outside of velocity domain is *at least* a factor of ϵ_{v} smaller than the maxima everywhere in the domain. Note that this choice of domain extent does not take into account flows whose temperature/velocity are expected to drastically differ from their initial conditions. In those scenarios, it may be necessary to verify the sensitivity of the solution on the domain extent *a posteriori*.

The distribution of nodes was chosen with polar/spherical integration in mind without *a priori* knowledge of the expected distribution function. Since there should be no bias for any particular angle in the polar/azimuthal distribution, a constant spacing with equal weighting was used. For N_ϕ polar and N_ψ azimuthal nodes, the nodal locations and weights were given as

$$\phi_i = i\Delta\phi \,\forall\, i \in \{0, ..., N_\phi - 1\}, \qquad \Delta\phi = \frac{2\pi}{N_\phi}, \qquad w_i^\phi = \Delta\phi, \tag{4.24}$$

$$\psi_i = (i + {}^1\!/\!{}_2)\Delta\psi \,\forall\, i \in \{0, ..., N_\psi - 1\}, \qquad \Delta\psi = \frac{\pi}{N_\psi}, \qquad w_i^\psi = \Delta\psi. \tag{4.25}$$

respectively. Note that one end is open for the polar distribution and both ends are open for the azimuthal distribution to account for periodicity in the respective components and to avoid coincident nodal points. For the radial distribution, the Gauss-Legendre nodes x^g and weights w^g were chosen as they provide satisfactory integration properties, avoid the singularity at $r = 0$, and cluster nodes at both the head and tail of the distribution. By applying a suitable normalization, the radial nodal locations and weights can be given as

$$r_i = r_{\max}\frac{x_i^g + 1}{2}, \qquad w_i^r = \frac{r_{\max}}{2} w_i^g J(r_i), \tag{4.26}$$

where $J(r_i)$ is the Jacobian of the transformation from polar/spherical coordinates to Cartesian coordinates. For a velocity domain dimension m, the Jacobian is calculated as the surface area of an m-ball of radius r_i, e.g.,

$$J(r_i) = \begin{cases} 2\pi r_i, & \text{if } m = 2, \\ 4\pi r_i^2, & \text{if } m = 3. \end{cases} \tag{4.27}$$

The dimensionality of the velocity space can then be calculated as $N_v = N_r N_\phi$ for $m = 2$ and $N_v = N_r N_\phi N_\psi$ for $m = 3$.

In the case of $m = 1$, the nodal distribution can be given across $[-r_{\max}, r_{\max}]$. To maintain consistency with the polar/spherical nodal distributions and to retain the clustering of the nodes around the head and tail, the nodal distribution was computed as the concatenation of two Gauss-Legendre nodal distributions of size $N_v/2$ on $[-r_{\max}, 0]$ and $[0, r_{\max}]$, respectively, which is identical to the two-dimensional case with $N_\phi = 2$. As such, for the one-dimensional cases in this work, N_v is set to an even integer.

4.2.3 Internal Energy Discretization

For the internal energy discretization with $\delta > 0$, a similar methodology as with the radial velocity discretization was applied. The internal energy nodal locations and weights are given in terms of the normalized Gauss-Legendre quadrature as

$$\zeta_i = \zeta_{\max} \frac{x_i^g + 1}{2} \ \forall\, i \in \{1, ..., N_\zeta\}, \qquad w_i^r = \frac{\zeta_{\max}}{2} w_i^g. \tag{4.28}$$

For the choice of $\zeta_{\max}$, similar arguments as with $r_{\max}$ can be made, but in this case, the upper bound does not have an impact on the maximum permissible explicit time step. To calculate $\zeta_{\max}$, a similar approach was taken as with $r_{\max}$ by solving for a value such that the relative contribution of the internal energy distribution outside of the domain is of $\mathcal{O}(\epsilon_\zeta)$, where ϵ_ζ is again some small user-defined constant to be presented in Section 4.3. However, unlike for the velocity domain, there does not exist a closed-form solution to this problem for general values of δ, and therefore it must be computed numerically. For several values of δ and ϵ_ζ, the ratio $\zeta_{\max}/\theta_{\max}$ was computed and presented in Table 4.1.

δ	$\epsilon_\zeta = 10^{-2}$	$\epsilon_\zeta = 10^{-4}$	$\epsilon_\zeta = 10^{-6}$	$\epsilon_\zeta = 10^{-8}$	$\epsilon_\zeta = 10^{-10}$	$\epsilon_\zeta = 10^{-12}$	$\epsilon_\zeta = 10^{-14}$
2	4.605	9.210	13.816	18.421	23.026	27.631	32.236
3	5.453	10.380	15.175	19.916	24.628	29.320	33.999
4	6.471	11.667	16.627	21.488	26.295	31.067	35.815
5	7.656	13.065	18.165	23.133	28.026	32.870	37.680

Table 4.1: Numerically computed values of the relative domain extent $\zeta_{\max}/\theta_{\max}$ for varying values of the tolerance ϵ_ζ and internal degrees of freedom δ.

4.2.4 Discrete Velocity Model

For the Boltzmann–BGK equation, accurate integration over the velocity and internal energy spaces is of key importance to the properties of the numerical scheme such as conservation and well-balancing. However, issues can arise when these infinite, continuous spaces are represented discretely. If $\mathbf{M}$ is a discrete nodal integration operator over $\Omega^{\mathrm{v}} \times \Omega^{\zeta}$ with strictly non-negative entries and $\mathbf{x}$ represents some discrete nodal values, i.e.,

$$\mathbf{M}\left(\mathbf{x}\right) \approx \int_{\mathbb{R}^d} \int_0^\infty x \, \mathrm{d}\zeta \, \mathrm{d}\mathbf{v}, \tag{4.29}$$

then it can be seen that in the limit of infinite resolution and infinite domain size, the discrete integral will converge towards the continuous integral if the integration operator is consistent. However, for finite approximations, this equality does not hold in general, and therefore for some discrete equilibrium distribution function g,

$$\mathbf{M}\left(\boldsymbol{\psi} \otimes \mathbf{g}(\mathbf{Q})\right) \neq \mathbf{Q}. \tag{4.30}$$

Without this equality, the moments of the equilibrium distribution function do not yield the same conserved variables used to create the equilibrium distribution function, which results in a scheme that is not conservative and is not well-balanced (i.e., a constant solution in thermodynamic equilibrium does not remain constant nor in thermodynamic equilibrium).

For nodal discretizations of the velocity and internal energy spaces, a standard approach is to discretize these spaces with enough resolution such that the integration errors are negligible []. However, due to the high-dimensional nature of the governing equations, this combinatorial explosion results in computational costs that severely limit the applicability of these methods. As an alternative approach, Mieussens [3] introduced a discrete velocity model (DVM) for the Boltzmann–BGK equation to address these issues. In the DVM approach, a discrete equilibrium distribution function $\mathbf{g}$ is sought that satisfies the *discrete* compatibility condition,

$$\mathbf{M}\left(\boldsymbol{\psi} \otimes \mathbf{g}\right) = \mathbf{M}\left(\boldsymbol{\psi} \otimes \mathbf{f}\right) = \mathbf{Q}, \tag{4.31}$$

and minimizes the discrete entropy H',

$$\mathbf{g} = \arg\min_{\mathbf{z}} H'(\mathbf{z}), \qquad H'(\mathbf{z}) = \mathbf{M}\left(\mathbf{z}\log\mathbf{z}\right). \tag{4.32}$$

It was shown in Mieussens [3] that there exists a unique solution $\mathbf{g}'$ to this problem, given by a *modified* Maxwellian of the form

$$\mathbf{g}' = \mathbf{g}(\mathbf{Q}'), \tag{4.33}$$

where $\mathbf{g}(\mathbf{Q}')$ is a discrete Maxwellian corresponding to a modified set of conserved variables $\mathbf{Q}' \neq \mathbf{Q}$ that converge to $\mathbf{Q}$ in the limit of infinite resolution. However, in general, the values of $\mathbf{Q}'$ that satisfy the discrete compatibility condition do not have a closed form solution, and therefore they must be found iteratively. This can be cast as a root-finding problem, where $\mathbf{Q}'$ is the root of

$$\mathbf{M}\left(\boldsymbol{\psi} \otimes \mathbf{g}(\mathbf{Q}')\right) - \mathbf{Q} = 0. \tag{4.34}$$

Note that due to the dimensional reduction caused by the discrete integration operator, this nonlinear optimization problem is only of the same dimension as $\mathbf{Q}$ (i.e., $d + 2$), irrespective of the dimensionality of g. Furthermore, the Jacobian $\mathbf{J}$ of the root-finding problem can be analytically calculated, allowing for the use of efficient root-finding algorithms such as Newton's method with minimal overhead.

To simplify the presentation, we will use the same notation as Mieussens [3], where the equilibirum distribution function is given with respect to a vector $\boldsymbol{\alpha}(\mathbf{Q}) = [\alpha_1, \alpha_2, ..., \alpha_{d+2}]^T$ instead of $\mathbf{Q}$ as

$$g(\boldsymbol{\alpha}) = \alpha_1 \exp\left[-\alpha_2\left(\sum_{i=1}^{d}(\mathbf{v}_i - \alpha_{i+2})^2\right)\right], \tag{4.35}$$

where for $d = 3$,

$$\alpha_1 = \frac{\rho}{(2\pi\theta)^{d/2}}, \quad \alpha_2 = \frac{1}{2\theta}, \quad \alpha_3 = U, \quad \alpha_4 = V, \quad \alpha_5 = W. \tag{4.36}$$

The Jacobian can then be given as

$$\mathbf{J}(\boldsymbol{\alpha}) = \mathbf{M}\left(\boldsymbol{\Theta} \otimes \boldsymbol{\psi} \otimes \mathbf{g}(\boldsymbol{\alpha})\right), \tag{4.37}$$

where

$$\boldsymbol{\Theta} = \frac{\partial g(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}\frac{1}{g(\boldsymbol{\alpha})}, \tag{4.38}$$

such that for $d = 3$,

$$\boldsymbol{\Theta}_1 = \frac{1}{\alpha_1}, \tag{4.39}$$

$$\boldsymbol{\Theta}_2 = -(\mathbf{u} - \alpha_3)^2 - (\mathbf{v} - \alpha_4)^2 - (\mathbf{w} - \alpha_5)^2 + \frac{\delta - 4\zeta\alpha_2}{2\alpha_2}, \tag{4.40}$$

$$\boldsymbol{\Theta}_3 = 2\alpha_2(\mathbf{u} - \alpha_3), \tag{4.41}$$

$$\boldsymbol{\Theta}_4 = 2\alpha_2(\mathbf{v} - \alpha_4), \tag{4.42}$$

$$\boldsymbol{\Theta}_5 = 2\alpha_2(\mathbf{w} - \alpha_5). \tag{4.43}$$

Note that due to the inclusion of the internal energy components, $\boldsymbol{\Theta}$ differs from the expression in

Mieussens [3] by the quantity

$$\frac{\delta - 4\zeta\alpha_2}{2\alpha_2} = \frac{\partial g_\zeta(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}} \frac{1}{g_\zeta(\boldsymbol{\alpha})}. \tag{4.44}$$

For the case of $\delta = 0$, this term is simply neglected as $g_\zeta = 1$. The root can then be found using Newton's method through the relation

$$\boldsymbol{\alpha}_{n+1} = \boldsymbol{\alpha}_n - \mathbf{J}(\boldsymbol{\alpha}_n)^{-1} \left[\mathbf{M}\left(\boldsymbol{\psi} \otimes \mathbf{g}(\boldsymbol{\alpha}_n)\right) - \mathbf{Q}\right], \tag{4.45}$$

with the initial guess $\boldsymbol{\alpha}_0$ taken as

$$\boldsymbol{\alpha}_0 = \boldsymbol{\alpha}(\mathbf{Q}). \tag{4.46}$$

In practice, since $\boldsymbol{\alpha}_0$ is generally a good guess for even moderately resolved velocity and internal energy spaces, very few iterations are required to bring the residual down to machine precision levels, with 1-2 iterations generally being sufficient.

4.2.5 Positivity-Preserving Limiter

As the distribution function f represents the number density of particles, it is inherently a non-negative quantity. However, as the high-order spatial scheme does not preserve a maximum principle on its own, a distribution function that is non-negative at $t = 0$ may not remain non-negative for $t > 0$. In addition to the fact that negative values of f are non-physical, this effect can result in stability problems. It is trivial to show that if f is strictly non-negative, the resulting density and specific internal energy are guaranteed to be non-negative given a discrete integration operator $\mathbf{M}$ with all positive values. The issue arises when f is no longer non-negative – then the density or specific internal energy values may become negative for which the equilibrium Maxwellian distribution is severely ill-behaved, causing the scheme to diverge. Therefore, for the purpose of robustness, it is necessary to modify the spatial scheme to ensure that f remains non-negative.

In this work, this property is enforced similarly to the work of Jaiswal [34] by using the "squeeze" limiter of Zhang and Shu [75]. Within each element Ω_k, the spatial average of the distribution function for each velocity and internal energy node $\mathbf{v}_q$ and ζ_r is computed as

$$\bar{f}_{q,r} = \sum_{i=1}^{N_s} m_i f_{q,r}(\mathbf{x}_i), \tag{4.47}$$

where the subscripts q, r denote the q-th velocity node and r-th internal energy node, respectively, and m_i are a set of quadrature weights such that

$$\sum_{i=1}^{N_s} m_i z(\mathbf{x}_i) = \frac{\int_{\Omega_k} z(\mathbf{x})\,\mathrm{d}\mathbf{x}}{\int_{\Omega_k} \mathrm{d}\mathbf{x}}, \tag{4.48}$$

for any polynomial $z(\mathbf{x})$ of degree $\leq p$. In Zhang and Shu [75] and Zhang and Shu [76], it was shown that the maximum principle is preserved for $\bar{f}$ under a standard CFL condition with strong-stability preserving temporal integration and an upwind interface flux. Due to the positivity of the equilibrium distribution function g in the source term of the Boltzmann–BGK equation, this property can be extended to the element-wise mean of high-order flux reconstruction approximations of the Boltzmann–BGK equation under the condition $\Delta t \leq \tau$ as this results in a convex combination of two positive states g and f. With the maximum principle property on $\bar{f}$, a conservative, positivity-preserving, and formally high-order reconstruction of f can be given as

$$\hat{f}_{q,r}(\mathbf{x}_i) = \bar{f}_{q,r}(\mathbf{x}_i) + \beta_{q,r}\left[f_{q,r}(\mathbf{x}_i) - \bar{f}_{q,r}\right], \tag{4.49}$$

where

$$\beta_{q,r} = \min\left[\left|\frac{\bar{f}_{q,r}}{\bar{f}_{q,r} - f_{q,r}^{\min}}\right|, 1\right] \quad \text{and} \quad f_{q,r}^{\min} = \min f_{q,r}(\mathbf{x}_i)\ \forall\, i \in \{1, ..., N_s\}. \tag{4.50}$$

This limiting is performed on a per-node basis in the velocity and internal energy spaces prior to each evaluation of the time derivative of the distribution function. Note that unlike some systems where an arbitrarily small positive constant must be given as a minimum to prevent numerical complications (e.g., Euler equations), f is allowed to attain a zero value, such that there is no need to prescribe an arbitrary minimum.

4.2.6 Discontinuity Capturing

Although the limiting approach in Section 4.2.5 is guaranteed to preserve the positivity of density and internal energy, the use of a high-order spatial discretization may still result in the introduction of spurious oscillations in the vicinity of discontinuities without proper treatment. For the Boltzmann equation, this is generally less of a problem as the discontinuous structures can in fact be resolved with some finite thickness at the particle mean free path level [33]. However, if the numerical resolution is not fine enough such that these structures are resolved, they can behave as discontinuities which can lead to numerical instabilities in the solution. For many high-order numerical schemes for both continuum gas dynamics and kinetics, it is commonplace to modify the parameters of the system in the vicinity of discontinuities such that they can be resolved on the order of the mesh scale (e.g., artificial viscosity [33]). However, as the focus of this work is on *directly resolving* flow physics, we instead pose some resolution requirements of the scheme such that the structures can be resolved without any additional treatment.

For a given numerical resolution, there exists some maximal mesh scale $h_{\max}$ such that features of a smaller scale cannot be resolved. For a high-order discontinuous spectral element method, computing this scale is more ambiguous due to the presence of multiple solution points within an element and inhomogeneity in the distribution of solution points. In this work, we compute this scale as the maximal distance between any solution point and its Voronoi neighbors within the element (see Section 2.4 of Guermond et al. [10]). To resolve "discontinuous" structures, it is necessary for their thickness Δ to be greater than this maximal mesh scale, i.e.,

$$\Delta \geq h_{\max}. \tag{4.51}$$

To form an *a priori* estimate of this thickness, we follow the methodology of Xiao [33], where the thickness is calculated in terms of the molecular mean free path λ as

$$\Delta \sim 10\lambda, \tag{4.52}$$

using the relation that the thickness of a weak shock wave is on the order of 10 mean free paths. Note that for strong shocks or other discontinuities, this relation may vary but it is generally of the same order of magnitude [4]. A mesh Knudsen number may then be calculated as

$$Kn_h = \frac{\lambda}{h_{\max}} = \frac{KnL_0}{h_{\max}}. \tag{4.53}$$

From Eq. (4.51) and Eq. (4.52), an approximate resolution requirement can be posed in terms of this mesh Knudsen number as

$$Kn_h \geq 1/10. \tag{4.54}$$

In practice, this requirement is checked *a priori* based on the initial Knudsen number but may also be observed over the course of the simulation based on the local mesh scale and local Knudsen number. Note that this requirement does not *guarantee* that the resulting scheme will be monotone in the vicinity of discontinuities, particularly due to the ambiguity in calculating $h_{\max}$ and the shock thickness, but it does form a suitable estimate such that the predicted results are generally well-behaved.

4.2.7 Boundary Conditions

The enforcement of boundary conditions in the solution of the Boltzmann equation is nontrivial, and there exist a variety of approaches for dealing with this problem. In this work, the boundary conditions are *weakly* enforced via the standard FR approach of forming a boundary state $f^+(\mathbf{v})$ that is used in conjunction with the interior solution values $f^-(\mathbf{v})$ to calculate the interface flux at the boundary. As the interface flux is computed using an upwinding approach, the boundary state can only affect the incoming particles at the boundary while the outgoing particles are essentially left unmodified.

The presentation of these boundary conditions will generally follow the work of Evans et al. [100]. The most straightforward boundary condition to enforce is a Neumann-type boundary condition, which can be implemented by simply setting the boundary state as

$$f^{+}(\mathbf{v}) = f^{-}(\mathbf{v}). \tag{4.55}$$

This boundary condition essentially assumes that the boundary has no effect on the solution, such that the correction terms associated with the interface flux are zero. For Dirichlet-type boundary conditions, the implementation becomes slightly more involved. It can be assumed in most scenarios that the boundary state can be represented in terms of the macroscopic variables $\mathbf{Q}$. The simplest case for this boundary condition is to assume that the boundary state is in thermodynamic equilibrium, such that the distribution function can be represented as a Maxwellian. Therefore, the boundary state can be set as

$$f^{+}(\mathbf{v}) = g(\mathbf{Q}'), \tag{4.56}$$

where $\mathbf{Q}'$ is the modified Maxwellian used to ensure a conservative scheme. Since the boundary state is typically fixed, this modified Maxwellian can be pre-computed.

Wall boundary conditions for the Boltzmann equation are significantly more complex, and in most cases, it is still an open problem as to what the proper approach is for representing interactions of the particles with the wall. Some theoretical development on wall boundary conditions for the Boltzmann equation is presented in Williams [104], and numerical applications of various approaches for wall boundary conditions are shown in Evans et al. [100]. The focus of this work is strictly on specular wall boundary conditions for convex surfaces which model the wall interaction as a reflection of the incoming particles, mimicking an effect that is somewhat similar to a slip-wall boundary condition for continuum approximations. For this boundary condition, the boundary state is set as

$$f^{+}(\mathbf{v}) = f^{-}(\mathbf{v} - 2(\mathbf{v}\cdot\mathbf{n})\mathbf{n}), \tag{4.57}$$

where $\mathbf{n}$ is the outward-facing normal direction of the wall. In general, the implementation of this boundary condition requires interpolation in velocity space, but for simple geometries, the implementation can be vastly simplified by use of a velocity space that is symmetric with respect to the wall-tangent directions.

4.3 Implementation

For the spatial discretization, the solution nodes were placed on the Gauss–Lobatto quadrature nodes for tensor-product elements and α-optimized points [13] for simplex elements, but any closed nodal set would essentially be identical due to the linearity of the flux. Temporal integration was performed via an explicit, fourth-order, four-stage Runge–Kutta scheme. Although the positivity property of the element-wise mean required by the positivity-preserving limiter theoretically requires strong stability preserving temporal integration, this is almost never an issue in practice, and it is therefore advantageous to utilize more efficient temporal schemes. The time step was chosen as the minima of the collision time scale τ and the maximum allowable time step by the Courant—Friedrichs—Lewy (CFL) condition Δt_{CFL}, i.e.,

$$\Delta t = \min\left(\tau, \Delta t_{CFL}\right), \tag{4.58}$$

where Δt_{CFL} was computed using the estimate of Cockburn and Shu [105] as

$$\Delta t_{CFL} = \frac{CFL}{2p+1}\frac{h_{\min}}{c_{\max}}. \tag{4.59}$$

Here, $CFL = 0.5$ is the chosen CFL number, p is the order of the scheme, $h_{\min}$ is the minimum element edge length in the domain, and $c_{\max} = r_{\max}$ is the maximum particle velocity in the domain. For certain problems, such as ones in the zero Mach number limit or ones that converge to a steady state, it may be advantageous to utilize implicit time stepping due to the stiffness of the source term. However, as this work is primarily focused on problems tending towards direct numerical simulation, where the time step limit imposed by the source term usually does not drastically differ from the CFL-based time step, explicit time stepping was deemed preferable due to its significant benefits in terms of computational efficiency on massively-parallel GPU computing architectures.

To simplify the presentation, the initial conditions are given in terms of the primitive variables $\mathbf{q} = [\rho, \mathbf{V}, P]^T$. The initial distribution function was set as the *modified* Maxwellian corresponding to the initial macroscopic variables through 5 iterations of the DVM, after which the DVM residual was generally on the order of machine precision. Throughout the simulations, a fixed number of iterations (2) was used for the DVM, which will be later shown to be sufficient to ensure conservation of the macroscopic flow variables. The extent of the velocity domain was computed as presented in Section 4.2.2 and the velocity tolerance was set as $\epsilon_{\mathbf{v}} = 10^{-15}$ as the estimated extent did not vary strongly with the tolerance (e.g., a velocity tolerance of $\epsilon_{\mathbf{v}} = 10^{-6}$ would only decrease the extent by 37%). However, for cases with internal degrees of freedom, since the extent of the internal energy domain varied more strongly with the internal energy tolerance (as shown in Table 4.1), the internal energy tolerance was instead set as $\epsilon_{\zeta} = 10^{-6}$. The effects of this truncation will be explored in Section 4.4. The velocity discretization was generally chosen with the goal of minimizing the necessary resolution, particularly for larger-scale problems, usually through qualitative convergence studies performed by incrementally increasing the resolution and comparing with established results.

Unless otherwise stated, the Knudsen number was computed with respect to a unit reference length and the discrete maximum of the initial wavespeed in the domain, and the velocity dimension m was set equal to the spatial dimension d. Furthermore, the collision time τ was set as constant based on the Knudsen number in almost all scenarios, except for problems where a variable collision time model is explicitly defined. For some problems, comparisons were made between the Boltzmann–BGK approach and a standard Navier–Stokes approach implemented within an identical codebase. The Navier–Stokes results were computed using the HLLC [81] Riemann solver for the inviscid fluxes and the BR2 approach of Bassi and Rebay [43] for the viscous fluxes, and a unit Prandtl number and identical specific heat ratio was used for consistency with the Boltzmann–BGK approach.

4.4 Results

4.4.1 Smooth Pulse Propagation

As an initial evaluation of the accuracy of the proposed scheme, the propagation of a smooth density pulse is considered. The problem is defined on a one-dimensional periodic domain $\Omega^x = [0, 1]$ and the initial conditions are given as

$$\mathbf{q}(x,0) = \begin{bmatrix} \rho \\ U \\ P \end{bmatrix} = \begin{bmatrix} 1 + \exp\left(-\beta\left(x - 0.5\right)^2\right) \\ 1 \\ 1 \end{bmatrix}, \tag{4.60}$$

where $\beta = 100$ is a parameter controlling the pulse width. The problem was investigated with $Kn = 10^{-1}, 10^{-2}$, and 10^{-3} based on a unit characteristic length and the maximum wavespeed in the domain, with an example of the resulting density profiles shown in Fig. 4.3. Additionally, the effects of including internal degrees of freedom were explored. Given a single velocity dimension ($m = 1$), a comparison was made between the monatomic case ($\delta = 0$, $\gamma = 3$) and the polyatomic case mimicking a diatomic molecule in three dimensions ($\delta = 4$, $\gamma = 1.4$).

4.4.1.1 Spatial Convergence

To verify the high-order spatial accuracy of the flux reconstruction approach, the convergence of the density error was evaluated. After one flow-through of the domain ($t = 1$), the L^∞ norm of the density error was calculated as

$$\epsilon_{\rho,\infty} = \|\rho(x,t) - \rho_{\text{ref}}(x,t)\|_{\infty,\Omega^x}, \tag{4.61}$$

where the extremum was computed on the discrete solution nodes and ρ_{ref} is the reference density computed with a highly-resolved numerical scheme. For this case, the reference simulation was performed using a $\mathbb{P}_5$ scheme with 100 elements. To isolate the effects of the spatial discretization error from the velocity discretization error, a very high resolution, $N_v = 128$, was used for the

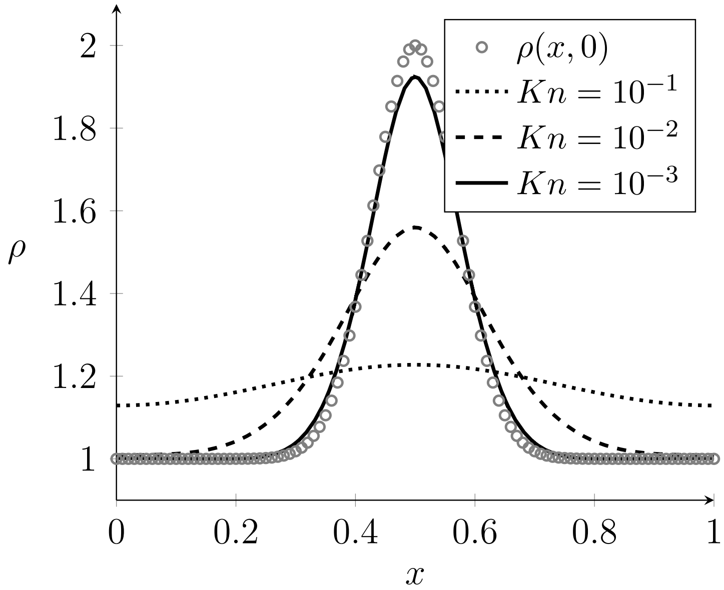


Figure 4.3: Density profile of the smooth pulse propagation problem at $t = 1$ with varying Knudsen numbers computed using a $\mathbb{P}_5$ scheme with 20 elements, the discrete velocity model (DVM) with $N_v = 32$, and $\delta = 0$ ($\gamma = 3$).

velocity space for all tests with $N_\zeta = N_v$ for the polyatomic case. The convergence of the density error for $Kn = 10^{-1}, 10^{-2}$, and 10^{-3} is shown in Table 4.2, Table 4.3, and Table 4.4 for the monatomic case and Table 4.5, Table 4.6, and Table 4.7 for the polyatomic case, respectively. For almost all cases, the expected theoretical $p + 1$ convergence rate was recovered or exceeded, although some variation is expected due to the lack of an analytic solution. These results confirm the high-order spatial accuracy afforded by the flux reconstruction scheme.

4.4.1.2 Velocity Convergence

To evaluate the effects of the velocity/internal energy discretization as well as the discrete velocity model on the solution, the density error was similarly analyzed while varying the velocity/internal energy resolution. To isolate the effects of the velocity/internal energy discretization error from the spatial discretization error, a highly-resolved $\mathbb{P}_5$ scheme with 20 elements was used for all tests. The L^∞ norm of the density error with respect to the resolution N_v, N_ζ after one flow-through of the domain is shown in Fig. 4.4 for $Kn = 10^{-1}, 10^{-2}$, and 10^{-3} with and with-

Δx	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$
$1/4$	3.37×10^{-2}	1.97×10^{-3}	1.20×10^{-4}	6.61×10^{-5}
$1/8$	7.03×10^{-4}	5.42×10^{-5}	2.61×10^{-6}	7.24×10^{-8}
$1/12$	4.52×10^{-5}	1.49×10^{-6}	5.51×10^{-8}	1.69×10^{-9}
$1/16$	1.84×10^{-5}	4.75×10^{-7}	1.10×10^{-8}	2.81×10^{-10}
$1/20$	9.04×10^{-6}	1.91×10^{-7}	3.61×10^{-9}	7.27×10^{-11}
RoC	5.26	5.98	6.71	8.63

Table 4.2: Convergence of the L^∞ norm of the density error with respect to mesh resolution Δx and varying approximations orders for the smooth pulse propagation problem with $Kn = 10^{-1}$ and $\delta = 0$ ($\gamma = 3$). Rate of convergence shown on bottom.

Δx	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$
$1/4$	1.43×10^{-1}	1.63×10^{-2}	3.33×10^{-3}	1.20×10^{-3}
$1/8$	1.59×10^{-2}	9.56×10^{-4}	1.30×10^{-4}	1.15×10^{-5}
$1/12$	2.85×10^{-3}	1.82×10^{-4}	1.54×10^{-5}	1.34×10^{-6}
$1/16$	9.67×10^{-4}	5.76×10^{-5}	3.90×10^{-6}	2.88×10^{-7}
$1/20$	4.55×10^{-4}	2.36×10^{-5}	1.21×10^{-6}	8.12×10^{-8}
RoC	3.63	4.06	4.93	5.92

Table 4.3: Convergence of the L^∞ norm of the density error with respect to mesh resolution Δx and varying approximations orders for the smooth pulse propagation problem with $Kn = 10^{-2}$ and $\delta = 0$ ($\gamma = 3$). Rate of convergence shown on bottom.

Δx	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$
$1/4$	3.54×10^{-1}	1.32×10^{-1}	4.69×10^{-2}	1.57×10^{-2}
$1/8$	1.03×10^{-1}	1.24×10^{-2}	2.16×10^{-3}	2.59×10^{-4}
$1/12$	3.25×10^{-2}	1.98×10^{-3}	2.25×10^{-4}	1.83×10^{-5}
$1/16$	1.08×10^{-2}	6.20×10^{-4}	4.76×10^{-5}	4.53×10^{-6}
$1/20$	4.37×10^{-3}	2.32×10^{-4}	1.76×10^{-5}	1.39×10^{-6}
RoC	2.71	3.96	4.98	5.85

Table 4.4: Convergence of the L^∞ norm of the density error with respect to mesh resolution Δx and varying approximations orders for the smooth pulse propagation problem with $Kn = 10^{-3}$ and $\delta = 0$ ($\gamma = 3$). Rate of convergence shown on bottom.

Δx	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$
$1/4$	3.76×10^{-3}	4.91×10^{-4}	1.98×10^{-4}	2.96×10^{-5}
$1/8$	7.46×10^{-4}	5.55×10^{-5}	2.80×10^{-6}	1.59×10^{-7}
$1/12$	2.22×10^{-4}	1.19×10^{-5}	4.67×10^{-7}	1.88×10^{-8}
$1/16$	9.68×10^{-5}	3.91×10^{-6}	1.05×10^{-7}	3.61×10^{-9}
$1/20$	5.00×10^{-5}	1.63×10^{-6}	3.64×10^{-8}	9.58×10^{-10}
RoC	2.70	3.55	5.31	6.35

Table 4.5: Convergence of the L^∞ norm of the density error with respect to mesh resolution Δx and varying approximations orders for the smooth pulse propagation problem with $Kn = 10^{-1}$ and $\delta = 4$ ($\gamma = 1.4$). Rate of convergence shown on bottom.

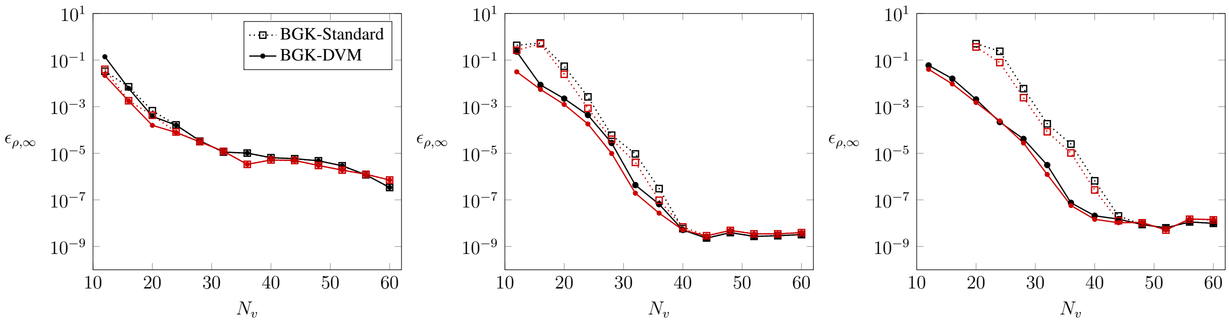
Δx	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$
$1/4$	3.06×10^{-2}	5.36×10^{-3}	5.93×10^{-3}	5.28×10^{-4}
$1/8$	5.92×10^{-3}	4.87×10^{-4}	5.59×10^{-5}	4.10×10^{-6}
$1/12$	1.88×10^{-3}	9.30×10^{-5}	6.53×10^{-6}	4.08×10^{-7}
$1/16$	8.19×10^{-4}	2.94×10^{-5}	1.68×10^{-6}	9.55×10^{-8}
$1/20$	4.33×10^{-4}	1.19×10^{-5}	5.80×10^{-7}	2.93×10^{-8}
RoC	2.65	3.81	5.72	6.07

Table 4.6: Convergence of the L^∞ norm of the density error with respect to mesh resolution Δx and varying approximations orders for the smooth pulse propagation problem with $Kn = 10^{-2}$ and $\delta = 4$ ($\gamma = 1.4$). Rate of convergence shown on bottom.

Δx	$\mathbb{P}_2$	$\mathbb{P}_3$	$\mathbb{P}_4$	$\mathbb{P}_5$
$1/4$	5.07×10^{-2}	9.73×10^{-3}	5.09×10^{-3}	1.52×10^{-3}
$1/8$	6.34×10^{-3}	3.75×10^{-4}	2.97×10^{-5}	6.76×10^{-6}
$1/12$	2.10×10^{-3}	7.56×10^{-5}	4.11×10^{-6}	6.39×10^{-7}
$1/16$	9.09×10^{-4}	1.95×10^{-5}	1.16×10^{-6}	2.68×10^{-7}
$1/20$	4.62×10^{-4}	8.86×10^{-6}	5.24×10^{-7}	1.24×10^{-7}
RoC	2.90	4.37	5.71	5.89

Table 4.7: Convergence of the L^∞ norm of the density error with respect to mesh resolution Δx and varying approximations orders for the smooth pulse propagation problem with $Kn = 10^{-3}$ and $\delta = 4$ ($\gamma = 1.4$). Rate of convergence shown on bottom.

out the DVM and internal degrees of freedom, respectively. For $Kn = 10^{-1}$, the results of the DVM and the standard approach were essentially identical over the resolution range. However, for smaller Kn, the difference between the standard approach and the DVM became more evident. For $Kn = 10^{-2}$ and $Kn = 10^{-3}$, the use of the DVM resulted in 1-3 orders of magnitude lower error than the standard approach over a large range of resolution, with this disparity becoming more evident with lower Kn. Furthermore, the use of the DVM sometimes resulted in a lower minimal stable resolution requirement for the velocity/internal energy spaces, with an example being the case of $Kn = 10^{-3}$ where the DVM was stable even with lowest resolution ($N_v, N_\zeta = 12$) while the standard approach diverged for $N_v, N_\zeta < 20$.



(a) $Kn = 10^{-1}$ (b) $Kn = 10^{-2}$ (c) $Kn = 10^{-3}$

Figure 4.4: Convergence of the L^∞ norm of the density error with respect to velocity space resolution N_v computed using a $\mathbb{P}_5$ scheme with 20 elements and $Kn = 10^{-1}$ (left), $Kn = 10^{-2}$ (middle), and $Kn = 10^{-3}$ (right) with and without the discrete velocity model (DVM). Results for $\delta = 0$ ($\gamma = 3$) and $\delta = 4$ ($\gamma = 1.4$) shown in black and red, respectively. Diverged solutions not plotted.

To verify the conservation properties of the DVM in comparison to the standard approach, the mass conservation error, defined as

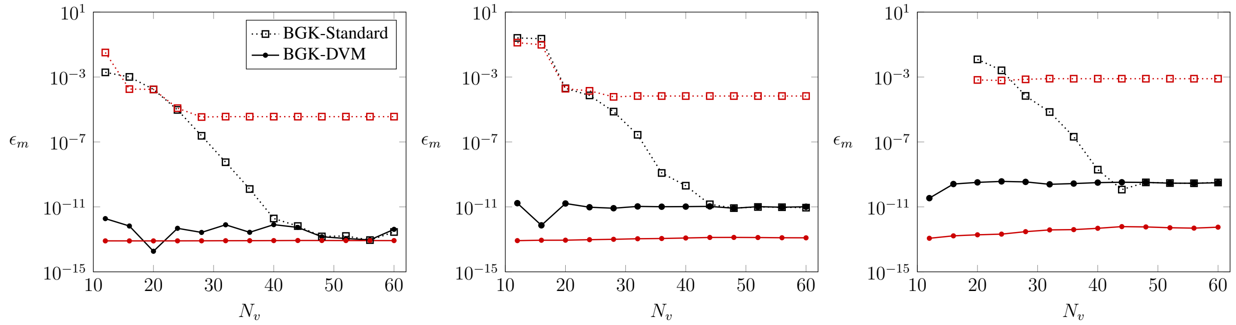
$$\epsilon_m = \left| \int_{\Omega^x} \rho(x,t)\, \mathrm{d}x - \int_{\Omega^x} \rho(x,0)\, \mathrm{d}x \right|, \tag{4.62}$$

was similarly analyzed after one flow-through of the domain. The convergence of this mass conservation error with respect to the velocity/internal energy resolution is shown in Fig. 4.5 for $Kn = 10^{-1}, 10^{-2}$, and 10^{-3}. In all cases, with just two iterations of the DVM, the mass conservation error was essentially negligible regardless of the resolution, varying from 10^{-13} to 10^{-10} across the various problem setups. However, without the DVM, the conservation error was significant, particularly at lower resolutions where errors of $\mathcal{O}(10^{-2})$ and larger were observed. For the monatomic case, where the domain bounds were chosen such that the values of the distribution function outside of the domain were essentially negligible ($\epsilon_{\mathbf{v}}$), the conservation error without the DVM did eventually converge to the DVM results with increasing resolution, but the resolution required to match the DVM results was significantly larger, such that the velocity/internal energy discretization required a factor of $3-4$ times as many nodes *per dimension*. In contrast, for the polyatomic case, the bounds of the internal energy domain were chosen such that value of the distribution function outside of it was small but non-negligible (ϵ_{ζ}) to decrease the resolution requirements. In this case, the mass conservation error with the DVM was still negligible, but without the DVM, the error was significant and did *not* converge with increasing resolution. This effect in the standard approach can be attributed to the error introduced by truncating the infinite domain. As a result, the DVM offers even more advantages in the overall computational cost as it allows for the use of smaller domain sizes which can be better resolved for the same number of discrete nodes.

4.4.2 Double Expansion Wave

A double expansion wave was subsequently considered as a verification of the entropy-satisfying and positivity-preserving properties of the scheme. This problem, known as the *123 problem* (or Test 2) in Toro [26], is solved on the domain $\Omega^x = [0, 1]$ with the initial conditions

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_L, & \text{if} x \leq 0.5, \\ \mathbf{q}_R, & \text{else,} \end{cases} \qquad \text{where} \quad \mathbf{q}_L = \begin{bmatrix} 1 \\ -2 \\ 0.4 \end{bmatrix}, \quad \mathbf{q}_R = \begin{bmatrix} 1 \\ 2 \\ 0.4 \end{bmatrix}. \tag{4.63}$$

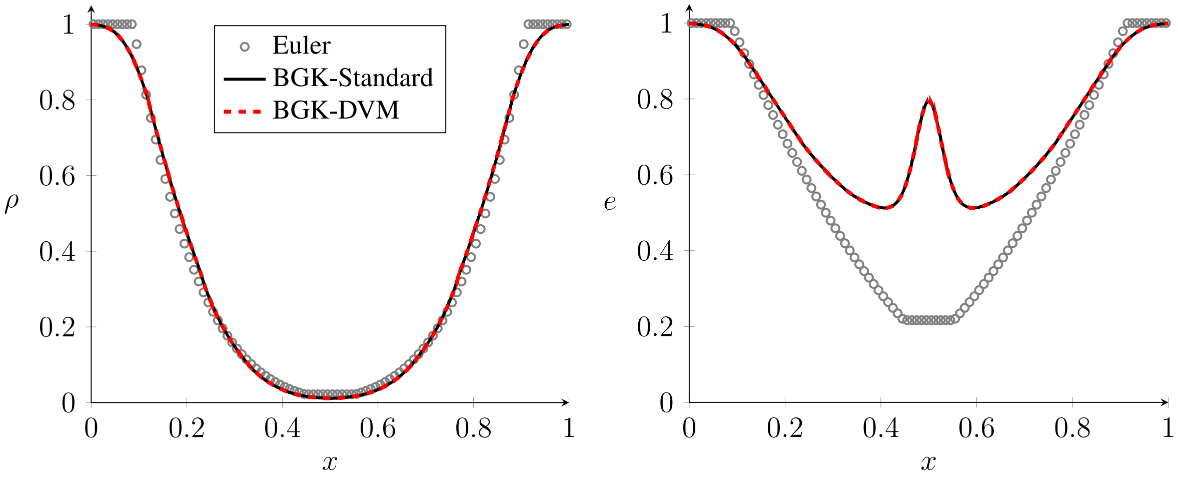


(a) $Kn = 10^{-1}$ (b) $Kn = 10^{-2}$ (c) $Kn = 10^{-3}$

Figure 4.5: Convergence of the mass conservation error with respect to velocity space resolution N_v computed using a $\mathbb{P}_5$ scheme with 20 elements and $Kn = 10^{-1}$ (left), $Kn = 10^{-2}$ (middle), and $Kn = 10^{-3}$ (right) with and without the discrete velocity model (DVM). Results for $\delta = 0$ $(\gamma = 3)$ and $\delta = 4$ $(\gamma = 1.4)$ shown in black and red, respectively. Diverged solutions not plotted.

The initial discontinuous state develops into two outrunning expansion waves with a low density/pressure region in the center. The resulting symmetric wave structure is difficult to resolve, particularly in the center region, with even robust first-order numerical schemes for gas dynamics frequently failing [26].

The problem was simulated using a $\mathbb{P}_3$ scheme with $N_e = 100$ elements and $N_v = N_\zeta = 32$ with Dirichlet boundary conditions. For consistency with the original problem setup, the polyatomic case was chosen, with $\delta = 4$ such that $\gamma = 1.4$. The predicted density and specific internal energy $e = \theta/(\gamma - 1)$ at $t = 0.15$ is shown in Fig. 4.6 and Fig. 4.7 for $Kn = 10^{-2}$ and $Kn = 10^{-3}$, respectively, based on the initial macroscopic state. For both Knudsen numbers, the density profile was well-resolved even in the low-density center region, with convergence to the exact Euler results qualitatively observed with decreasing Kn and near identical behavior with and without the DVM. For the specific internal energy profile, which is generally much more difficult to accurately resolve in this region as errors are amplified by the low density and pressure, the predicted results showed good agreement in the decreasing Kn limit on the exterior of the expansion waves. However, the center region shows a spike in the internal energy profile, indicative of spurious physical entropy generation in the center. This effect is quite common for gas dynamics solvers, with almost all schemes producing this energy spike (see Toro [26], Section 6.4). Interestingly enough, we also observe this effect when solving the Boltzmann equation, which should, in theory, be entropy-satisfying.

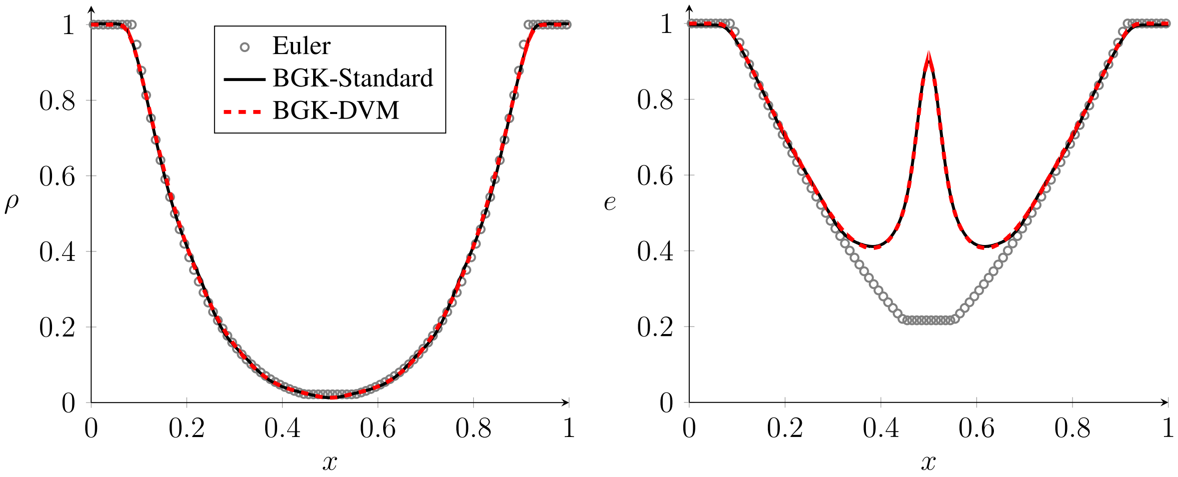


(a) Density

(b) Specific Internal Energy

Figure 4.6: Density (left) and specific internal energy (right) profile for the double expansion wave problem at $t = 0.15$ with $Kn = 10^{-2}$ computed using a $\mathbb{P}_3$ scheme with 100 elements, $N_v = N_\zeta = 32$, and $\delta = 4$ ($\gamma = 1.4$). Results shown in comparison to the exact Euler profiles.

To discern the cause of this spurious entropy generation, we explore several possibilities. One may attribute this problem to the effects of a non-zero Knudsen number, but Fig. 4.6 and Fig. 4.7 show that although the profiles in the expansion region tend toward the Euler results with decreasing Kn, the magnitude of the spike actually *increases* with decreasing Kn. Another explanation may be that this is an artifact of the mesh resolution, which is generally the case for gas dynamics solvers but would certainly raise questions about the entropy-satisfying properties of the proposed scheme. To explore this possibility, the problem was solved on a progression of meshes with increasing resolution, both with a fixed $Kn = 10^{-3}$ and with a proportionally decreasing $Kn = h/10$, where $h = 1/N_e$ is the mesh size. The predicted specific internal energy profiles for these cases is shown in Fig. 4.8. It can be seen that for the decreasing Kn case (i.e., constant mesh



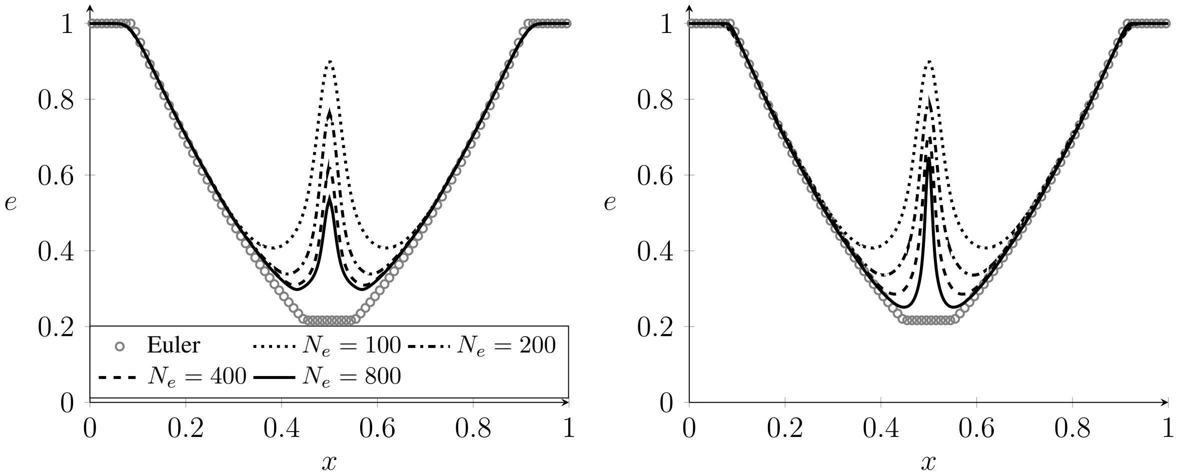
(a) Density

(b) Specific Internal Energy

Figure 4.7: Density (left) and specific internal energy (right) profile for the double expansion wave problem at $t = 0.15$ with $Kn = 10^{-3}$ computed using a $\mathbb{P}_3$ scheme with 100 elements, $N_v = N_\zeta = 32$, and $\delta = 4$ ($\gamma = 1.4$). Results shown in comparison to the exact Euler profiles.

Knudsen number), the profiles converge to the Euler results, but the spike does not proportionally decrease. In contrast, the fixed Kn case (e.g., decreasing mesh Knudsen number) shows a significant decrease in the energy spike with increasing resolution, indicating that the problem is more associated with the mesh Knudsen number than the resolution itself.

We posit that this effect is related to a violation of the resolution requirements proposed in Section 4.2.6. Although it may seem that for the given parameters, the resolution condition in Eq. (4.54) is easily satisfied, the assumptions posed in Section 4.2.6 are not valid for the initial conditions — regardless of the Knudsen number of the problem, the initial discontinuity thickness given by Eq. (4.63) is exactly zero. Therefore, the actual mesh Knudsen number for the initial state is zero, such that the scales are not properly resolved and Eq. (4.54) is not actually satisfied. This issue is exacerbated by the stationary nature of the center region in symmetric Riemann problems, which does not benefit from "correcting" effects such as compression in shocks or numerical dissipation in traveling contact discontinuities. To verify this claim, the previous mesh convergence study was run again utilizing the case of a decreasing Knudsen ($Kn = h/10$), but the discontinuous original initial condition was modified to a *continuous* initial condition with a finite discontinuity



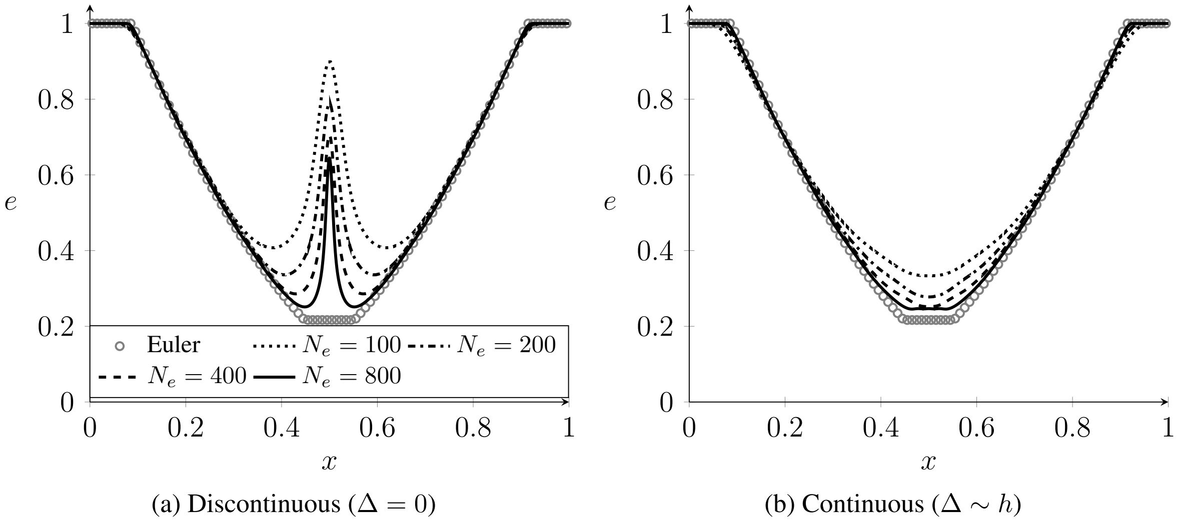
(a) Fixed Kn (b) Decreasing Kn

Figure 4.8: Convergence of the specific internal energy profile with respect to mesh resolution N_e for the double expansion wave problem at $t = 0.15$ computed using a $\mathbb{P}_3$ scheme, $N_v = N_\zeta = 32$ with the discrete velocity model, and $\delta = 4$ ($\gamma = 1.4$). Left: Discontinuous initial conditions ($\Delta = 0$). Right: Continuous initial conditions ($\Delta \sim h$). Results shown in comparison to the exact Euler profiles.

thickness Δ. The modified initial velocity profile was chosen as

$$U(x,0) = \frac{U_R - U_L}{2}\tanh\left[\frac{1}{h}(x-0.5)\right] + \frac{U_R + U_L}{2} = 2\tanh\left[\frac{1}{h}(x-0.5)\right], \tag{4.64}$$

such that the discontinuity thickness Δ is of $\mathcal{O}(h)$ which satisfies the resolution requirements at $t = 0$ and recovers the Euler equations in the $Kn \to 0$, $h \to 0$ limit. A comparison of the predicted specific internal energy profile for the discontinuous and continuous initial conditions is shown in Fig. 4.9. As it can clearly be seen, modifying the initial conditions to be consistent with the system in question completely removes the spurious spike in specific internal energy (and consequently, physical entropy) while converging towards the Euler results in the $Kn \to 0$, $h \to 0$ limit. Note that the modification was only necessary due to a pathological test case known to be sensitive to this issue — for practical purposes (and for the remaining numerical experiments in this work), the modification to the initial conditions was neglected as its effect was generally negligible in comparison to numerical approximation errors.



(a) Discontinuous ($\Delta = 0$)

(b) Continuous ($\Delta \sim h$)

Figure 4.9: Convergence of the specific internal energy profile with respect to mesh resolution N_e and $Kn = h/10$ for the double expansion wave problem at $t = 0.15$ computed using a $\mathbb{P}_3$ scheme, $N_v = N_\zeta = 32$ with the discrete velocity model, and $\delta = 4$ ($\gamma = 1.4$). Left: Discontinuous initial conditions ($\Delta = 0$). Right: Continuous initial conditions ($\Delta \sim h$). Results shown in comparison to the exact Euler profiles.

4.4.3 Normal Shock Structure

As a verification of the proposed method for more complex flow regimes and a comparison of the Boltzmann-BGK method to continuum approaches, the structure of a normal shock in argon was computed over a range of Mach numbers. For this problem, the domain is set as $\Omega^x = [-25, 25]$ and the initial conditions are given as

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_L, & \text{if} x \leq 0, \\ \mathbf{q}_R, & \text{else}, \end{cases} \qquad \text{where} \quad \mathbf{q}_L = \begin{bmatrix} 1 \\ M\sqrt{\gamma} \\ 1 \end{bmatrix}. \tag{4.65}$$

The left-hand state $\mathbf{q}_L$ is given in terms of the incoming (upstream) Mach number M, and the right-hand state $\mathbf{q}_R$ can be computed through the Rankine–Hugoniot conditions as

$$\frac{\rho_R}{\rho_L} = \frac{(\gamma+1)M^2}{(\gamma-1)M^2+2}, \qquad \frac{U_R}{U_L} = \frac{(\gamma-1)M^2+2}{(\gamma+1)M^2}, \qquad \text{and} \qquad \frac{P_R}{P_L} = \frac{2\gamma M^2-(\gamma-1)}{\gamma+1}. \tag{4.66}$$

For comparison with experimental data and other numerical approaches, the specific heat ratio γ was set as $5/3$ (i.e., $\delta = 2$) to mimic a monatomic molecule (argon) in three dimensions. The quantity of interest for this problem is the inverse ratio of the shock thickness Δ, defined as

$$\Delta = \frac{\rho_R - \rho_L}{\max\left[\partial_x \rho(x)\right]}, \tag{4.67}$$

relative to the upstream mean free path λ_L.

Since the shock structure is highly sensitive to the transport coefficients, the use of a constant collision time τ, for which the viscosity would depend on pressure, can result in inaccurate predictions. One can instead adapt the collision time based on the solution to recover more physically consistent transport coefficients similarly to the approach of Mieussens [3] as

$$\tau(\mathbf{Q}) = \tau_{ref} \frac{\rho_{ref}\theta_{ref}^{1-\omega}}{\rho\theta^{1-\omega}}. \tag{4.68}$$

Here, ω refers to the viscosity law exponent of the gas, taken as $\omega = 0.81$ for argon, and the reference quantities are taken from the incoming macroscopic state $\mathbf{Q}_L$. The reference collision time τ_{ref} was computed by setting the Knudsen number as $Kn = 1$ based on a unit reference length and the incoming macroscopic state $\mathbf{Q}_L$, such that $\lambda_L = Kn$.

The shock structure was computed using a $\mathbb{P}_3$ scheme with 100 elements with $N_v = N_\zeta = 32$ which allows for the structure to be fully resolved based on the Knudsen number and the spatial resolution. The boundary conditions were set to Dirichlet on both ends. To compute the shock thickness, the corrected gradients were used, with the common interface solution taken as the mean of the interior and exterior values similarly to the BR1 approach of Bassi and Rebay [106]. The solution was advanced to a final time of $t = 100$, after which the L^∞ norm of the temporal residual was generally converged to values less than 10^{-5}. The computed inverse thickness ratio λ_L/Δ is shown in Fig. 4.10 over a range of Mach numbers in comparison to the Navier–Stokes results of Mieussens [3] and experimental results of Alsmeyer [4], Linzer and Hornig [5], and Camac [6]. The predictions of the present Boltzmann–BGK approach were in much better agreement with the experimental data than the Navier–Stokes results, properly predicting both the critical Mach number where this inverse thickness ratio is at its maxima and the decay of the ratio with increasing Mach number. The predicted inverse thickness ratio was generally on the higher end of the experimental data but still within the range of uncertainty between the various experiments.

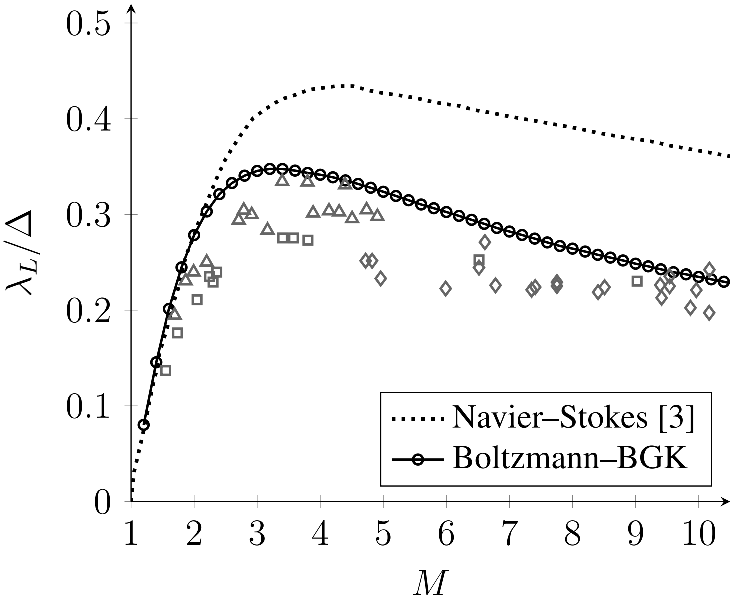
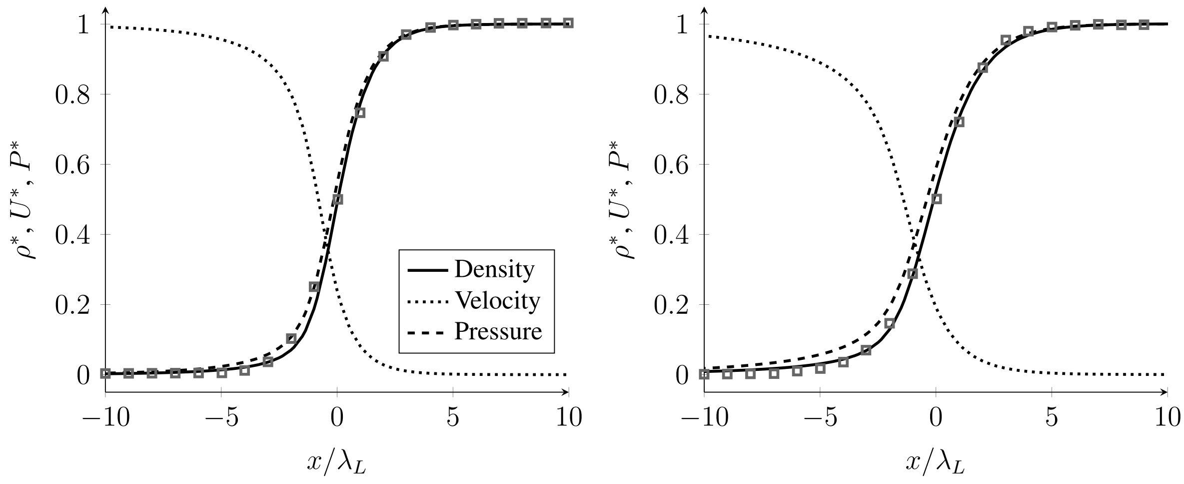


Figure 4.10: Inverse thickness ratio with respect to inflow Mach number for a stationary normal shock in argon ($\delta = 2$, $\gamma = 5/3$) computed using a $\mathbb{P}_3$ scheme with 100 elements and $N_\upsilon = N_\zeta = 32$. Navier–Stokes results of Mieussens [3] and experimental results of Alsmeyer [4] (squares), Linzer and Hornig [5] (triangles), and Camac [6] (diamonds) shown for reference.

For a more direct comparison of the shock structures, the predicted normalized density profiles (ρ^*) as well as normalized velocity (U^*) and pressure (P^*) profiles at two Mach numbers, $M = 3.8$

and $M = 9.0$, were compared to the experimental density profiles of Alsmeyer [4] in Fig. 4.11. At $M = 3.8$, the predicted density profile was generally in good agreement with the experiment with respect to both the thickness and the structure, although some minor undershoot in the upstream section of the shock was observed. At $M = 9.0$, even better agreement with the experiment was observed, with the predicted density profiles matching up almost exactly. The benefits of the kinetic approach to predicting shock structures are clearly evident in these results.



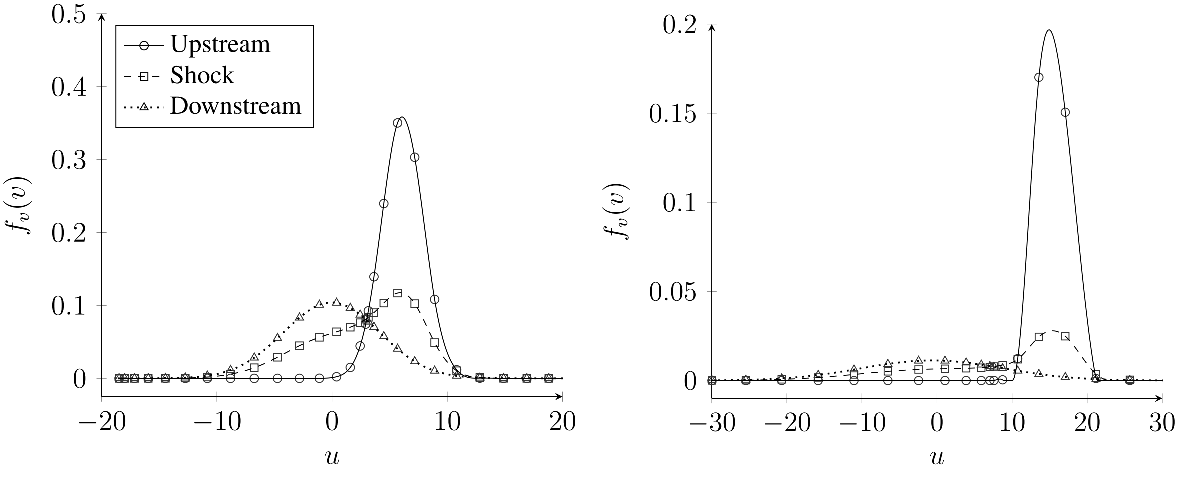
(a) $M = 3.8$ (b) $M = 9.0$

Figure 4.11: Normalized density, velocity, and pressure profiles for a stationary normal shock in argon ($\delta = 2$, $\gamma = 5/3$) at $M = 3.8$ (left) and $M = 9.0$ (right) computed using a $\mathbb{P}_3$ scheme with 100 elements and $N_v = N_\zeta = 32$. Experimental density profiles of Alsmeyer [4] (square markers) shown for reference.

To observe the effects of the various flow conditions on the distribution function, the profiles of the distribution function at various locations in the flow were visualized for $M = 3.8$ and $M = 9.0$. Since the velocity/internal energy domain is two-dimensional in this case, the velocity component $f_v(v)$ was extracted by taking the discrete maxima over the internal energy domain, i.e.,

$$f_v(v) = \max_{\zeta} f(u, \zeta). \tag{4.69}$$

The predicted distribution function profiles for $M = 3.8$ and $M = 9.0$ are shown in Fig. 4.12 at the upstream ($x = -25$), shock ($x = 0$), and downstream ($x = 25$) locations. The discrete values of the distribution function are shown as markers, but for presentation purposes, a spline interpolation of these values is also shown. The profiles indicate that the *a priori* approach for choosing the extent of the velocity domain is sufficient as all of the statistically significant variation in the profiles was completely captured within the domain. For both Mach numbers, the upstream and downstream distribution functions appeared Maxwellian, indicating the flow is in thermodynamic equilibrium, whereas at the shock showed clear non-equilibrium behavior. The distribution functions both also showed the increase of the thermal velocity across the shock. These results appear to agree qualitatively with the results of Mieussens [3] obtained with a three-dimensional velocity space at a slightly different Mach number.



(a) $M = 3.8$

(b) $M = 9.0$

Figure 4.12: Velocity component of the distribution function at various spatial locations for a stationary normal shock in argon ($\delta = 2$, $\gamma = 5/3$) at $M = 3.8$ (left) and $M = 9.0$ (right) computed using a $\mathbb{P}_3$ scheme with 100 elements and $N_v = N_\zeta = 32$. Nodal values shown as markers, spline interpolation shown as lines.

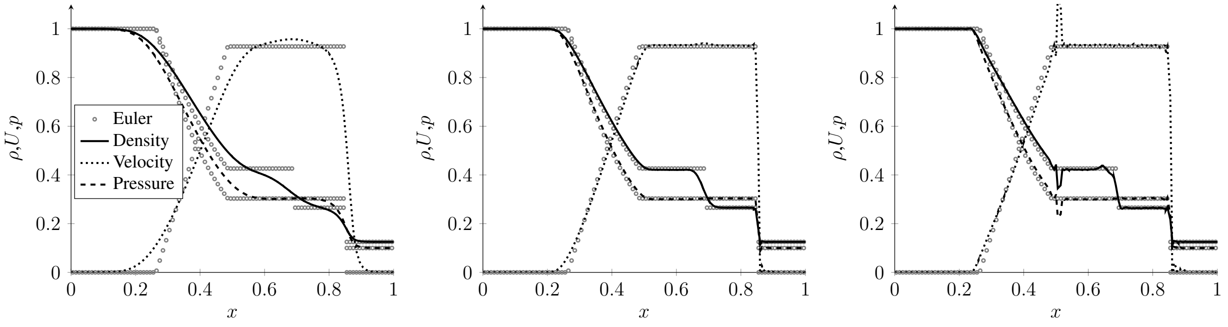
4.4.4 Sod Shock Tube

Extensions to unsteady gas dynamics with discontinuities was performed through the Sod shock tube [85]. This one-dimensional test case exhibits the three main features of the Riemann problem, namely a shock wave, a contact discontinuity, and a rarefaction wave, which makes it a suitable test for the discontinuity-resolving properties of the proposed scheme. The problem is computed on the domain $\Omega^x = [0, 1]$ and the initial conditions are given as

$$\mathbf{q}(x,0) = \begin{cases} \mathbf{q}_L, & \text{if} x \leq 0.5, \\ \mathbf{q}_R, & \text{else}, \end{cases} \qquad \text{where} \quad \mathbf{q}_L = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad \mathbf{q}_R = \begin{bmatrix} 0.125 \\ 0 \\ 0.1 \end{bmatrix}, \tag{4.70}$$

with Dirichlet boundary conditions on both ends. The specific heat ratio γ was set to 1.4 ($\delta = 4$) for consistency with the original problem setup.

To evaluate the validity of the resolution requirements posed in Section 4.2.6, the parameters of the scheme were fixed while modulating the mesh Knudsen number. A $\mathbb{P}_3$ scheme with 50 elements was used with $N_v = N_\zeta = 16$ and mesh Knudsen numbers of $Kn_h = 1$, $1/10$, and $1/100$ based on the initial macroscopic state. The predicted macroscopic solution profiles are shown in Fig. 4.13 at $t = 0.2$ for these respective mesh Knudsen numbers. For $Kn_h = 1$, the solution was excessively diffused, although the locations and magnitudes of most discontinuities were reasonably predicted. With $Kn_h = 1/10$ (the proposed resolution limit), the solution was in good agreement with the exact Euler results, with excellent resolution of the location and magnitude of the discontinuities and minimal observable spurious oscillations. The contact discontinuity was more diffused compared to the shock, but this behavior is expected due to the compressive nature of shock waves. When the mesh Knudsen number was lowered past the proposed limit to $Kn_h = 1/100$, spurious oscillations began to appear in the predicted profiles, severely degrading the accuracy of the solution. However, the location and magnitude of the discontinuities were still reasonably well predicted, indicating that the prediction of the governing physics is not severely degraded even with a numerically ill-behaved solution.



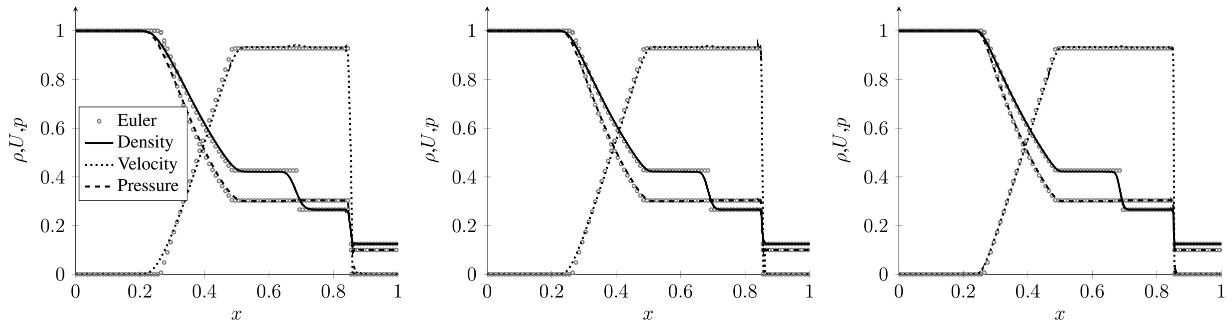
(a) $Kn_h = 1$ (b) $Kn_h = 1/10$ (c) $Kn_h = 1/100$

Figure 4.13: Density, velocity, and pressure profiles for the Sod shock tube problem at $t = 0.2$ with varying mesh Knudsen number Kn_h computed using a $\mathbb{P}_3$ scheme with 50 elements, $N_v = N_\zeta = 16$, and $\delta = 4$ ($\gamma = 1.4$). Results shown in comparison to the exact Euler profiles.

The proposed resolution limit was further explored with a mesh convergence study. The mesh Knudsen number was fixed at $Kn_h = 1/10$ while the mesh resolution was varied with $N_e = 50$, 100, and 200. The predicted solution profiles showed excellent convergence to the exact Euler results with increasing resolution (i.e., decreasing Knudsen number). Sub-element resolution of the shock was generally observed, whereas the contact was generally resolved over 1-2 elements. As mentioned in Section 4.2.6, the resolution limit does not guarantee monotonicity of the solution in the vicinity of a discontinuity. This was observed in the predicted velocity profiles with very minor overshoots in the vicinity of the shock. However, the predicted density and pressure profiles were (visually) monotonic in the vicinity of the discontinuities. These results indicate that the mesh Knudsen number offers a suitable indicator metric for the resolution of discontinuous structures.

4.4.5 Mach 3 Forward Facing Step

The proposed approach was then extended to large-scale problems on two-dimensional unstructured grids with complex flow physics including strong shocks and boundary interactions through the forward facing step problem of Woodward and Colella [86]. This problem originally consists of



(a) $N_e = 50$ (b) $N_e = 100$ (c) $N_e = 200$

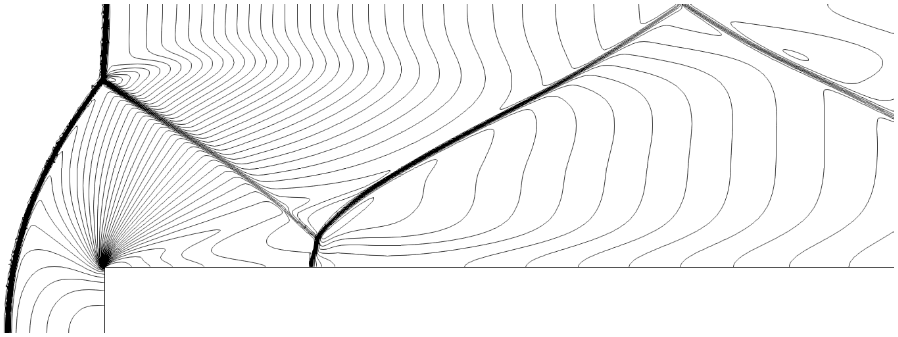
Figure 4.14: Density, velocity, and pressure profiles for the Sod shock tube problem at $t = 0.2$ with a fixed mesh Knudsen number $Kn_h = 1/10$ computed using a $\mathbb{P}_3$ scheme with varying mesh resolution, $N_v = N_\zeta = 16$, and $\delta = 4$ ($\gamma = 1.4$). Results shown in comparison to the exact Euler profiles.

an inviscid Mach 3 flow in a wind tunnel with a step perturbation that forms strong shocks, rarefaction fans, and contact discontinuities. The domain is set as $\Omega^{\mathbf{x}} = [0,3]\times[0,1]\backslash[0.6,3]\times[0,0.2]$ with the initial conditions set as $\mathbf{q}(\mathbf{x},0) = [1,1,0,1/(\gamma M^2)]^T$ for a specific heat ratio $\gamma = 1.4$ ($\delta = 3$) and Mach number $M = 3$. Dirichlet boundary conditions were applied on the inlet ($x = 0$) and Neumann boundary conditions were applied on the outlet ($x = 3$). The remaining boundaries were set to specular wall boundary conditions. To vastly simplify the implementation of the specular wall boundary conditions, the velocity offsets were set as $\mathbf{V}_0 = \mathbf{0}$ and the polar resolution N_ϕ was set as a multiple of 4, such that the velocity space was symmetric about the wall normal directions and the reflection operator in Eq. (4.57) could be easily and efficiently implemented without requiring interpolation in velocity space.

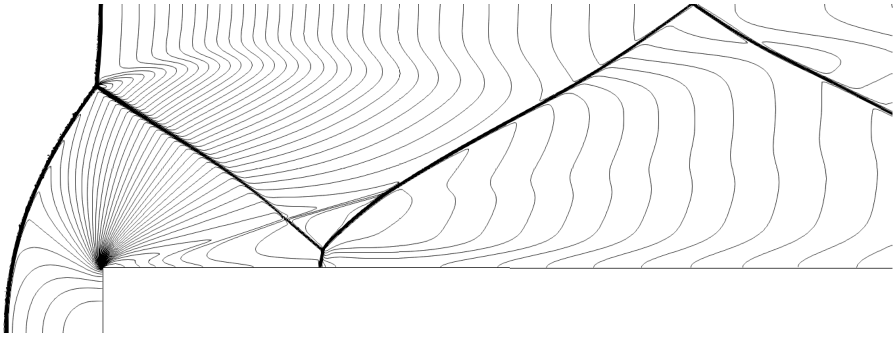
The problem was solved using a $\mathbb{P}_4$ scheme on a series of meshes of increasing resolution with $N_r = N_\phi = 16$ and $N_\zeta = 8$. The coarse, medium, and fine meshes were generated using approximately homogeneous unstructured triangles with characteristic mesh lengths of $h = 1/50$, $1/100$, and $1/200$, respectively, resulting in $2.3{\cdot}10^4$, $7.6{\cdot}10^4$, and $3.1{\cdot}10^5$ elements. For the given resolution in the spatial, velocity, and internal energy domains, this resulted in approximately 703 million total degrees of freedom for the coarse mesh, 2.38 billion total degrees of freedom for the medium mesh, and 9.65 billion total degrees of freedom for the fine mesh. The Knudsen number was set based on the initial conditions as $Kn = 1{\cdot}10^{-2}$, $2{\cdot}10^{-3}$, and $2{\cdot}10^{-3}$ for the coarse, medium, and fine meshes, respectively, such that the mesh Knudsen number was approximately constant and the flow structures could be fully resolved. The approximate computational cost of the finest case was 480 GPU-hours (15 hours on 32 NVIDIA V100 GPUs), which, given such a large number of degrees of freedom, shows the efficiency of the proposed approach.

The predicted density profiles at $t = 12$, represented through isocontours equispaced on the range $[0, 5]$, are shown in Fig. 4.15 for the various meshes and problem setups. The results show the canonical structure of the forward-facing step, with a strong shock reflected from the step, a rarefaction fan centered on the corner, and a Mach stem emanating from the top wall. For each case, the discontinuities in the flow were very well resolved and showed no spurious oscillations in the vicinity of the shocks. Furthermore, the thickness of the discontinuities decreased accordingly with decreasing mesh size and Knudsen number. With decreasing Knudsen number, it can be seen that the results are converging to predictions in the Euler limit (see Dumbser et al. [107], Fig. 8). These effects were most noticeable in the "liftoff" of the Mach stem emanating from the bottom wall, which should decrease to zero in the absence of physical and numerical viscosity, and the behavior of the contact line emerging from the triple point. In the Euler limit, the onset of Kelvin–Helmholtz instabilities should become evident along this contact line, but with a finite Knudsen number, these instabilities can be stabilized due to viscous effects.

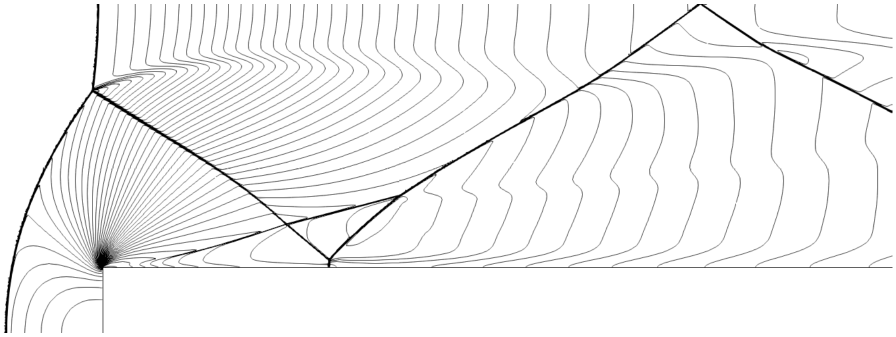
To evaluate the ability of the proposed approach to resolve shock structures, the contours of density in the region of the leading shock are shown in Fig. 4.16 in comparison to the computational mesh. It can be seen that the flow fields show quite impressive sub-element resolution of the shock, effectively predicting a monotonic shock structure over a length of only 1-2 spatial solution points. Furthermore, this behavior was maintained with decreasing mesh scale by fixing the mesh Knudsen number. The level of resolution of shocks relative to the mesh scale afforded by the present Boltzmann–BGK approach *without any numerical shock capturing scheme* rival that of a high-order discontinuous spectral element method augmented with highly-resolved *a posteriori*



(a) Coarse mesh, $Kn = 1{\cdot}10^{-2}$

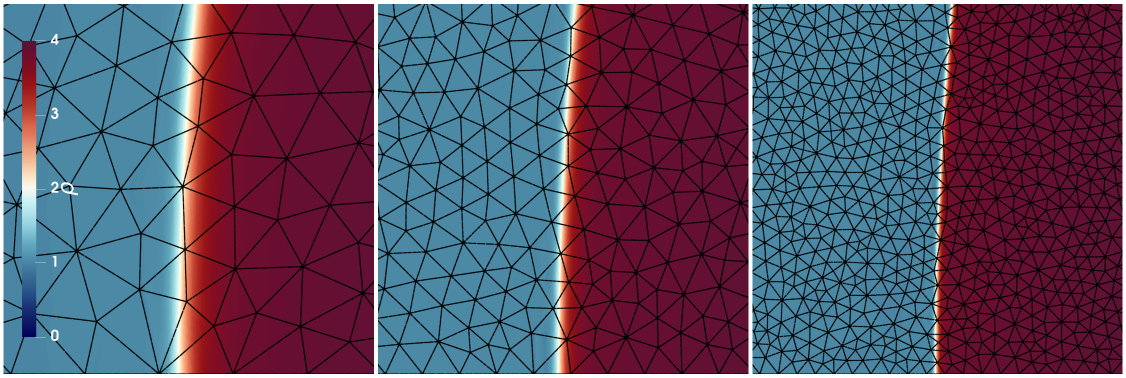


(b) Medium mesh, $Kn = 5{\cdot}10^{-3}$



(c) Fine mesh, $Kn = 2{\cdot}10^{-3}$

Figure 4.15: Isocontours of density (equispaced on the range $[0, 5]$) for the Mach 3 forward facing step problem at $t = 12$ computed using a $\mathbb{P}_4$ scheme with $N_r = N_\phi = 16$, $N_\zeta = 8$, and $\delta = 3$ ($\gamma = 1.4$). Top row: Coarse mesh ($2.3{\cdot}10^4$ elements) with $Kn = 1{\cdot}10^{-2}$. Middle row: Medium mesh ($7.6{\cdot}10^4$ elements) with $Kn = 5{\cdot}10^{-3}$. Bottom row: Fine mesh ($3.1{\cdot}10^5$ elements) with $Kn = 2{\cdot}10^{-3}$.



(a) Coarse mesh, $Kn = 1{\cdot}10^{-2}$ (b) Medium mesh, $Kn = 5{\cdot}10^{-3}$ (c) Fine mesh, $Kn = 2{\cdot}10^{-3}$

Figure 4.16: Enlarged view of the contours of density with mesh overlay near the leading shock for the Mach 3 forward facing step problem at $t = 12$ computed using a $\mathbb{P}_4$ scheme on a coarse (left), medium (middle), and fine (right) mesh.

WENO-type reconstruction [107]. These results indicate that the proposed approach can be a very effective tool for directly resolving shock structures in more complex fluid flows.

4.4.6 Kelvin–Helmholtz Instability

As an initial evaluation of the Boltzmann–BGK approach in predicting turbulent flow phenomena, the onset and development of a two-dimensional Kelvin–Helmholtz instability was computed, where two counter-flowing fluids of varying density exhibit a fluid instability which transitions into complex vortical flow. The domain is set as periodic on $\Omega^{\mathbf{x}} = [-0.5, 0.5]^2$ and the initial conditions are given as

$$\mathbf{q}(x,y,0) = \begin{cases} \mathbf{q}_L, & \text{if } |y| \leq 0.25, \\ \mathbf{q}_R, & \text{else}, \end{cases} \qquad \text{where} \quad \mathbf{q}_L = \begin{bmatrix} 2 \\ 0.5 \\ V(x,y) \\ 6.25 \end{bmatrix}, \quad \mathbf{q}_R = \begin{bmatrix} 1 \\ -0.5 \\ V(x,y) \\ 6.25 \end{bmatrix}, \tag{4.71}$$

which yields an Atwood number of $1/3$ and a maximum Mach number of 0.2. To seed instabilities in the flow, a single mode is excited similarly to Springel [108] by setting the initial vertical velocity component as

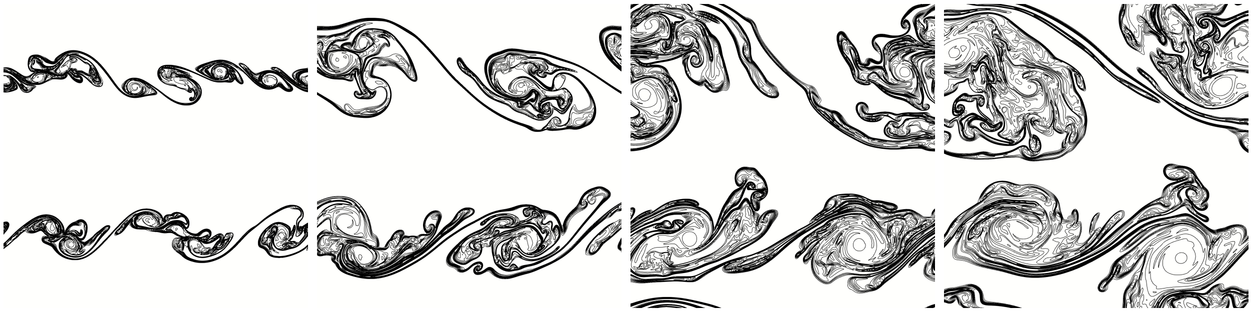
$$V(x,y) = \kappa \sin(\omega \pi x) \left[\exp\left(-\frac{(y-0.25)^2}{2\sigma^2} \right) + \exp\left(-\frac{(y+0.25)^2}{2\sigma^2} \right) \right], \tag{4.72}$$

where the parameters $\kappa = 0.02$, $\omega = 128$, and $\sigma = 0.0035$ dictate the perturbation strength, frequency, and vertical decay, respectively. The collision time τ was set constant according to a reference dynamic viscosity $\mu = \tau P = 10^{-5}$ based on the initial static pressure. Due to the low Mach number, the variations in pressure, and therefore viscosity, were relatively low.

To verify the ability of the proposed scheme to predict the onset of flow instabilities consistent with the hydrodynamic equations, a comparison between the solution computed with the Boltzmann–BGK equation and the solution computed with the Navier–Stokes equations was performed. For both cases, the solution was computed on a uniform 400^2 quadrilateral mesh with a $\mathbb{P}_3$ scheme, and to reduce the overall computational cost, the monatomic case was chosen ($\delta = 0$, $\gamma = 2$). For the velocity space discretization of the Boltzmann–BGK approach, the resolution was set as $N_r = N_\phi = 32$, yielding approximately 2.62 billion degrees of freedom. To maintain consistency between the two methods, the Navier–Stokes approach was computed with $Pr = 1$, $\gamma = 2$, and $\mu = 10^{-5}$, and the previously-developed entropy filtering method was used to stabilize the high-order scheme in the vicinity of discontinuities [109].

The predicted density profiles, represented as equispaced isocontours on the range $[1, 2]$, at unit time intervals are shown in Fig. 4.17 for both the Boltzmann–BGK and Navier–Stokes approaches. It can be seen that even though transport in the Boltzmann–BGK equation is linear, the predicted results evidently show the development of the instability into complex vortical flow. In fact, in the earlier time intervals, the large-scale coherent structures of the flow were almost identically computed by the Navier–Stokes and the Boltzmann–BGK approaches. The two flow fields did start to diverge at later times, but this is expected due to the chaotic nature of the Kelvin–Helmholtz instability. Furthermore, the range of scales of the flow features were generally similar between the two approaches, indicating that the turbulent length scales are properly resolved by the Boltzmann–BGK approach. Some discrepancies in the smallest scales did exist between the two approaches, with the Boltzmann–BGK approach showing marginally smaller flow features. This may be attributed to the numerical dissipation introduced by the discontinuity capturing approach for the Navier–Stokes method which artificially inflates the physical viscosity and the minor fluctuations in viscosity of the Boltzmann–BGK approach related to low pressure regions in the vortex cores. However, the marked similarities between the predicted flow fields still indicate that the Boltzmann–BGK approach can accurately resolve multi-scale nonlinear flow phenomena in a manner that is consistent with predictions in the continuum limit.

(a) Boltzmann–BGK



(b) Navier–Stokes

Figure 4.17: Isocontours of density (equispaced on the range $[1, 2]$) for the Kelvin–Helmholtz instability problem at $t = 1$ (left), $t = 2$ (middle-left), $t = 3$ (middle-right), and $t = 4$ (right) computed using a $\mathbb{P}_3$ scheme with 400^2 elements and $\mu = 10^{-5}$. Top row: Boltzmann–BGK with $N_r = N_\phi = 32$ and $\delta = 0$ ($\gamma = 2$). Bottom row: Navier–Stokes with $Pr = 1$ and $\gamma = 2$.

4.4.7 Compressible Taylor–Green Vortex

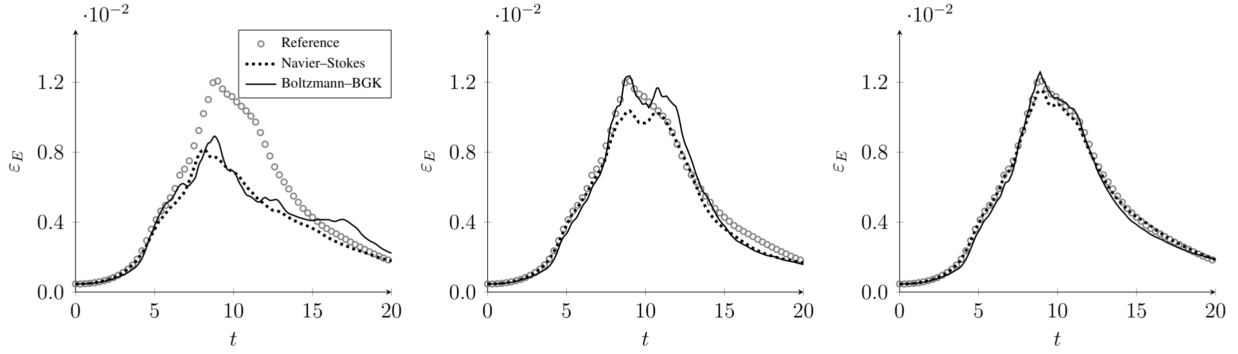
As the final and most rigorous test case, the transition to turbulence of a three-dimensional compressible Taylor–Green vortex was studied. To the author's knowledge, this is the first direct numerical simulation of three-dimensional turbulent flows performed by directly solving the Boltzmann equation, made possible by the efficiency of the proposed approach. This canonical test case consists of an initially laminar flow that transitions as time evolves, forming large coherent vortical structures that eventually break down to small-scale isotropic turbulence. The problem is solved the periodic domain $\Omega^{\mathbf{x}} = [-\pi, \pi]^3$, and the initial conditions are given as

$$\mathbf{q}(x,y,z,0) = \begin{bmatrix} 1 \\ \sin(x)\cos(y)\cos(z) \\ -\cos(x)\sin(y)\cos(z) \\ 0 \\ P_0 + \frac{1}{16}\left(\cos(2x) + \cos(2y)\right)\left(\cos(2z+2)\right) \end{bmatrix}. \tag{4.73}$$

The reference pressure was set as $P_0 = 1/(\gamma M^2)$, where $\gamma = 5/3$ and $M = 0.2$. The monatomic case ($\delta = 0$) was chosen to make the computational cost tractable and the Mach number was slightly increased over its typical value of $M = 0.08$ to decrease the stiffness of the source term. The collision time was set constant as $\tau = \mu/P_0$, where the reference dynamic viscosity was set as $\mu_{ref} = 1/1600$ corresponding to a Reynolds number of 1600. Through *a posteriori* analysis, the variation in the viscosity due to the fixed collision time was deemed to be negligible due to the low Mach number.

The problem was solved on a series of increasingly-resolved meshes and compared to a standard Navier–Stokes approach. As a point of reference, direct numerical simulation was performed by solving the flow field with the same Navier–Stokes approach ($\gamma = 5/3$, $Pr = 1$) on a $N_e = 64^3$ mesh with a $\mathbb{P}_3$ scheme, a suitable level of resolution for capturing all statistically significant physical scales [17]. Comparison was then performed on a series of meshes with $N_e = 16^3$, 32^3, and

48^3, also discretized with a $\mathbb{P}_3$ scheme. For the Boltzmann–BGK approach, the velocity domain was discretized with $N_r = 16$, $N_\phi = 16$, and $N_\psi = 8$, which will later be shown to be sufficiently resolved. For the finest case, where the resolution level corresponded to approximately 14.5 billion degrees of freedom, the computational cost was on the order of 3000 GPU-hours (72 hours on 40 NVIDIA V100 GPUs).



(a) $N_e = 16^3$ (b) $N_e = 32^3$ (c) $N_e = 48^3$

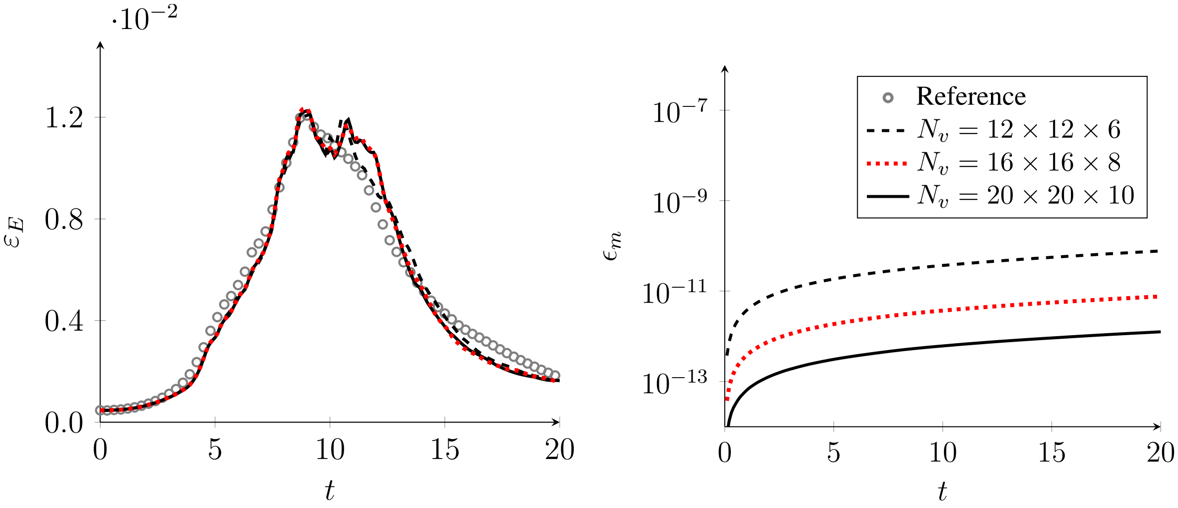
Figure 4.18: Dissipation measured by enstrophy for the compressible Taylor-Green vortex at $Re = 1600$ computed using a $\mathbb{P}_3$ scheme with 16^3 (left), 32^3 (middle), and 48^3 (right) elements. Boltzmann–BGK results computed with $N_r = 16$, $N_\phi = 16$, $N_\psi = 8$, and $\delta = 0$ ($\gamma = 5/3$). Navier–Stokes results computed with $Pr = 1$ and $\gamma = 5/3$.

The standard metric of comparison for the Taylor–Green vortex is the dissipation ε_E measured by a scaled form of the non-dimensional volume-integrated enstrophy, given as

$$\varepsilon_E = \frac{\beta}{V} \int_{\Omega^{\mathbf{x}}} \frac{1}{2} \rho \boldsymbol{\omega} \cdot \boldsymbol{\omega} \, \mathrm{d}\mathbf{x}, \tag{4.74}$$

where $\boldsymbol{\omega}$ is the vorticity, $\beta = 2\mu$ is the scaling factor, and $V = 8\pi^3$ is the domain volume. The viscosity was taken as a constant corresponding to the reference value μ_{ref}, but the differences to the enstrophy values computed using the pressure-dependent viscosity $\mu = \tau P$ were confirmed to be negligible. The profiles of the enstrophy-based dissipation rate are shown in Fig. 4.18 for the various meshes. For the coarsest mesh, the enstrophy was underpredicted for both the Boltzmann–

BGK and the Navier–Stokes approaches in comparison to the reference results due to the low levels of resolution, However, the profiles showed relatively good agreement between the two approaches, with the Boltzmann–BGK results showing some more oscillations in the profile and an overprediction of the enstrophy at later times. When the resolution was increased to the medium mesh, the enstrophy profile of the Boltzmann–BGK approach showed much better agreement with the reference results, with better agreement at the enstrophy peak than the Navier–Stokes results at the same resolution but with a spurious secondary peak at a later time. For the finest mesh, the results of the Boltzmann–BGK approach converged excellently to the reference results with nearly an identical enstrophy profile. Furthermore, these results were consistent with the work of Cao et al. [110] which utilized a gas kinetic scheme derived from the Chapman—Enskog expansion of the Boltzmann–BGK equation that directly solves for the conserved flow variables.



(a) (b)

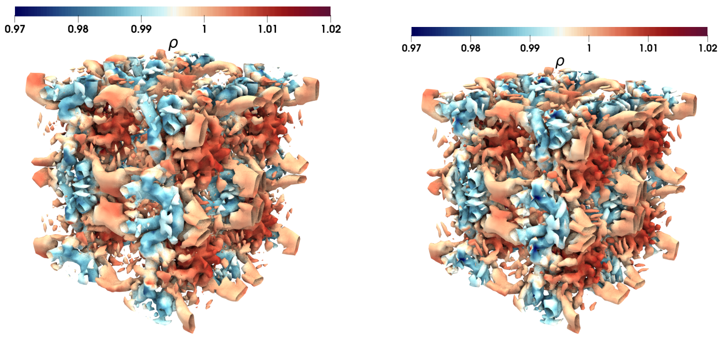
Figure 4.19: Convergence of dissipation measured by enstrophy (left) and mass conservation error (right) with respect to velocity space resolution $N_v = N_r \times N_\phi \times N_\phi$ for the Boltzmann-BGK approximation of the compressible Taylor–Green vortex at $Re = 1600$ computed using a $\mathbb{P}_3$ scheme with the DVM and $N_e = 32^3$. Default velocity space resolution shown in red.

To verify that the results were converged in the velocity domain and the scheme remained conservative, the spatial resolution was fixed at $N_e = 32^3$ and the velocity resolution was modulated.

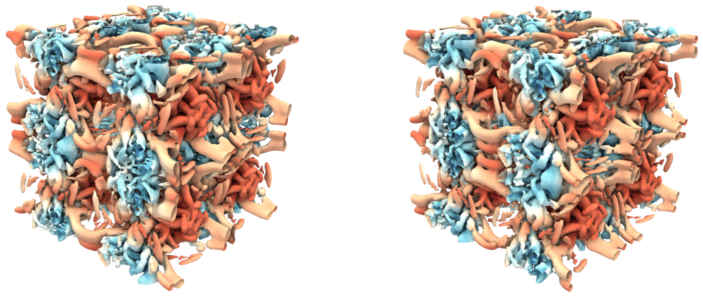
The predicted enstrophy profiles and the mass conservation error for the various velocity resolution levels are shown in Fig. 4.19. In comparison to the standard resolution used in this study ($N_v = 16 \times 16 \times 8$), lowering the resolution to $N_v = 12 \times 12 \times 6$ resulted in a very similar enstrophy profiles, with just a slightly deviation in the second spurious enstrophy peak. When the resolution was increased to $N_v = 20 \times 20 \times 10$, the results were essentially indistinguishable from the standard resolution, indicating that the results were well-converged in the velocity domain for the choice of velocity resolution used for the experiments. Furthermore, regardless of the choice of the resolution, the scheme remained conservative, with a mass conservation error of less than 10^{-10} purely as a result of the linear accumulation of mass conservation error roughly on the order of machine precision per time step.

A comparison between the flow fields was then performed to confirm that the turbulent structures in the flow were accurately predicted. The isosurfaces of Q-criterion (with $Q = 0.2$) colored by density are shown for the various meshes in Fig. 4.20 at $t = 10$ where the enstrophy is near its peak. It can be seen that the dominant flow structures were accurately predicted by the Boltzmann–BGK approach in comparison to the reference results, even at the lowest mesh resolution, and these results were very similar to the Navier–Stokes predictions at the same resolution. When the resolution was increased, the predicted results converged excellently to the reference results, with effectively identical prediction of the flow structures and the density distribution within the domain.

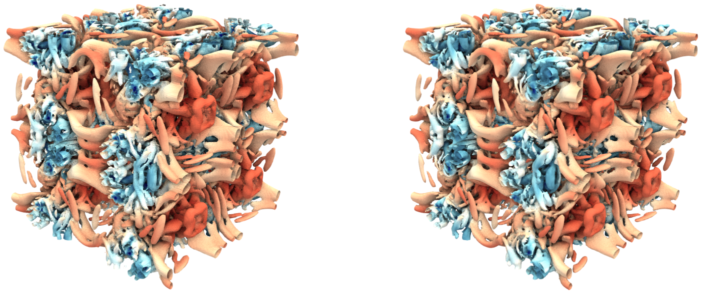
As a final verification of the ability of the Boltzmann–BGK approach in predicting turbulent flow phenomena, the behavior of the turbulent kinetic energy spectra was examined. Similarly to the Q-criterion isosurfaces, this analysis was performed near the enstrophy peak at $t = 10$. The spectra $E(k)$ was computed by taking the three-dimensional fast Fourier transform of the square of the velocity magnitude ($\mathbf{V} \cdot \mathbf{V}$) interpolated onto a uniform grid of the same number of degrees of freedom and then performing a spherical averaging operation with the number of bins set identical to the maximum resolvable wavenumber in one-dimension ($N_e(p+1)/2$). The predicted energy spectra are shown in Fig. 4.21 for the various meshes. It can be seen that the predicted spectra



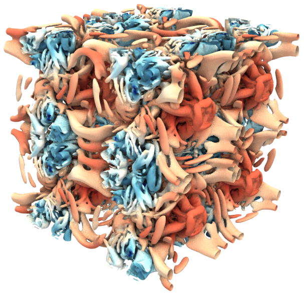
(a) Boltzmann-BGK, $N_e = 16^3$ (b) Navier–Stokes, $N_e = 16^3$



(c) Boltzmann-BGK, $N_e = 32^3$ (d) Navier–Stokes, $N_e = 32^3$

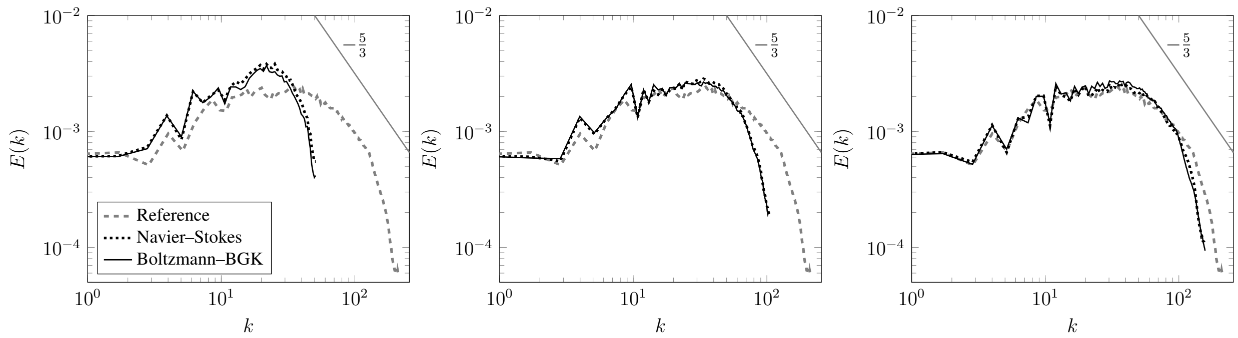


(e) Boltzmann-BGK, $N_e = 48^3$ (f) Navier–Stokes, $N_e = 48^3$



(g) Reference

Figure 4.20: Isosurfaces of Q-criterion $= 0.2$ colored by density at $t = 10$ for the compressible Taylor–Green vortex at $Re = 1600$ computed using a $\mathbb{P}_3$ scheme with varying mesh resolution. Left column: Boltzmann–BGK results with $N_r = 16$, $N_\phi = 16$, $N_\psi = 8$, and $\delta = 0$ ($\gamma = 5/3$). Right column: Navier–Stokes results with $Pr = 1$ and $\gamma = 5/3$.



(a) $N_e = 16^3$ (b) $N_e = 32^3$ (c) $N_e = 48^3$

Figure 4.21: Turbulent kinetic energy spectra at $t = 10$ for the compressible Taylor-Green vortex computed using a $\mathbb{P}_3$ scheme with 16^3 (left), 32^3 (middle), and 48^3 (right) elements. Boltzmann–BGK results computed with $N_r = 16$, $N_\phi = 16$, $N_\psi = 8$, and $\delta = 0$ ($\gamma = 5/3$). Navier–Stokes results computed with $Pr = 1$ and $\gamma = 5/3$.

converged to the reference results with increasing resolution. Furthermore, for a given resolution, the Boltzmann–BGK approach remarkably predicts the energy spectra almost identically to the Navier–Stokes approach. These results present a validation of the Boltzmann–BGK approach for three-dimensional turbulent flows and indicate that the approach can accurately resolve turbulent flow phenomena in a consistent manner with respect to the hydrodynamic equations.

4.5 Conclusions

A novel numerical approach for the solving the polyatomic Boltzmann equation with a BGK collision model was presented in this chapter [111]. The proposed approach utilizes a combination of an efficient implementation of a high-order FR spatial discretization augmented with a positivity-preserving limiter and a discrete velocity model ensuring the conservation and well-balancing properties of the scheme. As a result, the computation of three-dimensional problems on arbitrary domains becomes entirely feasible on the current generation of high-performance computing architectures. The method was validated on a series of canonical flow problems in both the continuum and rarefied regime, showing high-order spatial accuracy and discrete conservation as well as sub-element resolution of shock waves without the need of any *ad hoc* numerical shock capturing techniques. Furthermore, the results of the simulation of a three-dimensional compressible

Taylor–Green vortex, which is, to the author's knowledge, the first such case of a three-dimensional turbulent flow simulated via directly solving the Boltzmann equation, demonstrate the ability of the proposed approach as a method for direct numerical simulation of compressible turbulence. The ability of directly solving the Boltzmann equation opens a new path for the study of transition to turbulent and shock boundary layer interactions in high speed flow since the evolution of the probability density function encodes the dynamics of the flow in a form that is not accessible from the solution of the compressible Navier–Stokes equations. Furthermore, this approach allows for a unified framework for simulating flows in both the continuum and rarefied regime.

5. CONCLUSION

Various algorithmic developments for high-order DSEM were proposed in this work to tackle problems in continuum fluid mechanics and molecular gas dynamics. The effects of high-order discretizations on hybrid turbulence models were first explored through simulations of turbulence using the Partially-averaged Navier–Stokes approach. The approach was applied to two canonical flows, the wall-bounded periodic hill and a circular cylinder at a Reynolds number of 3900. It was found that this method particularly benefited from the reduced numerical dissipation and increased accuracy of DSEM in comparison to approaches without models and that the results were less sensitive to the resolution-control parameter.

Extensions of DSEM to flows with discontinuous features was then performed by introducing an adaptive entropy-based filtering approach as a shock capturing technique. By formulating convex invariants such as positivity of density and pressure and a local minimum entropy principle as discrete constraints on the solution, the filter strength was computed through an element-wise nonlinear optimization process. The proposed filtering approach, which can be implemented on general unstructured grids and does not require any tunable parameters, could robustly resolve strong discontinuities without excessively dissipating away small-scale flow features. The efficacy of the method was shown in numerical experiments on hyperbolic and mixed hyperbolic-parabolic conservation laws such as the Euler and Navier–Stokes equations for problems including extreme shocks, shock-vortex interactions, and complex compressible turbulent flows.

Finally, the simulation of non-equilibrium and rarefied flows was performed by utilizing DSEM for the approximation of the nonlinear Boltzmann equation. Through the combination of a positivity-preserving limiter and a discrete velocity model, the resulting scheme guaranteed the positivity of the macroscopic density and pressure as well as discrete conservation and well-balancing regardless of the velocity resolution. The approach was validated on a series of large-scale complex numerical experiments, ranging from shock-dominated flows computed on unstructured grids to direct numerical simulation of three-dimensional compressible turbulent flows, the latter of which is the first instance of such a flow computed by directly solving the Boltzmann equation. The results showed the ability of the scheme to directly resolve shock structures without any *ad hoc* numerical shock capturing method and correctly approximate turbulent flow phenomena in a consistent manner with the hydrodynamic equations.

Future work can further improve upon these proposed developments and address additional challenges of DSEM. For hybrid turbulence modeling, limiters can be implemented to enhance the robustness of the approximation of the turbulence variables, and the resolution control parameter can be adaptively chosen by estimating the level of resolution in the flow field through a modal decomposition of the velocity field within elements. For shock capturing within the entropy filtering framework, the proposed approach can be improved by increasing the locality of the filtering kernel to mimic that of methods such as artificial viscosity and subcell approaches, modifying the approach to remove the need for operator splitting in mixed systems, and applying a more robust method for calculating the entropy tolerance. Finally, the approximation of the Boltzmann equation would benefit from implicit time integration schemes and the implementation of an adaptive discretization in velocity space.

REFERENCES

[1] W. M. van Rees, A. Leonard, D. Pullin, and P. Koumoutsakos, "A comparison of vortex and pseudo-spectral methods for the simulation of periodic vortical flows at high Reynolds numbers," *Journal of Computational Physics*, vol. 230, no. 8, pp. 2794–2805, Apr. 2011.

[2] R. Konrath, C. Klein, R. Engler, and D. Otter, "Analysis of PSP results obtained for the VFE-2 65° delta wing configuration at sub- and transonic speeds," in *44th AIAA Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics, Jan. 2006.

[3] L. Mieussens, "Discrete velocity model and implicit scheme for the BGK equation of rarefied gas dynamic," *Mathematical Models and Methods in Applied Sciences*, vol. 10, no. 08, pp. 1121–1149, Nov. 2000.

[4] H. Alsmeyer, "Density profiles in argon and nitrogen shock waves measured by the absorption of an electron beam," *Journal of Fluid Mechanics*, vol. 74, no. 3, pp. 497–513, Apr. 1976.

[5] M. Linzer and D. F. Hornig, "Structure of shock fronts in argon and nitrogen," *Physics of Fluids*, vol. 6, no. 12, p. 1661, 1963.

[6] M. Camac, "Argon shock thickness," *Physics of Fluids*, vol. 7, no. 7, p. 1076, 1964.

[7] B. Vermeire, F. Witherden, and P. Vincent, "On the utility of GPU accelerated high-order methods for unsteady flow simulations: A comparison with industry-standard tools," *Journal of Computational Physics*, vol. 334, pp. 497–521, Apr. 2017.

[8] A. Ern and J.-L. Guermond, *Theory and Practice of Finite Elements*. Springer New York, 2004.

[9] G. Petrov, "Application of Galerkin's method to the problem of stability of flow of a viscous fluid," *J. Appl. Math. Mech*, vol. 4, no. 3, pp. 3–12, 1940.

[10] J.-L. Guermond, R. Pasquetti, and B. Popov, "Entropy viscosity method for nonlinear conservation laws," *Journal of Computational Physics*, vol. 230, no. 11, pp. 4248–4267, May

2011.

[11] A. N. Brooks and T. J. Hughes, “Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations,” *Computer Methods in Applied Mechanics and Engineering*, vol. 32, no. 1-3, pp. 199–259, Sep. 1982.

[12] W. H. Reed and T. R. Hill, “Triangular mesh methods for the neutron transport equation.” in *National topical meeting on mathematical models and computational techniques for analysis of nuclear systems*, April 1973, dataset: LANLOPAC.

[13] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods*. Springer New York, 2008.

[14] Y. Liu, M. Vinokur, and Z. J. Wang, “Discontinuous spectral difference method for conservation laws on unstructured grids,” in *Computational Fluid Dynamics*. Springer Berlin Heidelberg, 2004, pp. 449–454.

[15] D. A. Kopriva and J. H. Kolias, “A conservative staggered-grid Chebyshev multidomain method for compressible flows,” *Journal of Computational Physics*, vol. 125, no. 1, pp. 244–261, Apr. 1996.

[16] H. T. Huynh, “A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods,” in *18th AIAA Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics, Jun. 2007.

[17] C. Cox, W. Trojak, T. Dzanic, F. Witherden, and A. Jameson, “Accuracy, stability, and performance comparison between the spectral difference and flux reconstruction schemes,” *Computers & Fluids*, vol. 221, p. 104922, May 2021.

[18] F. Ilinca and D. Pelletier, “Positivity preservation and adaptive solution of two-equation models of turbulence,” *International Journal of Thermal Sciences*, vol. 38, no. 7, pp. 560–571, Jul. 1999.

[19] F. Bassi, A. Crivellini, S. Rebay, and M. Savini, “Discontinuous Galerkin solution of the Reynolds-averaged Navier–Stokes and k–ω turbulence model equations,” *Computers &*

Fluids, vol. 34, no. 4-5, pp. 507–540, May 2005.

[20] M. Wurst, M. Keßler, and E. Krämer, "Detached eddy simulation using the discontinuous Galerkin method," in *Notes on Numerical Fluid Mechanics and Multidisciplinary Design*. Springer International Publishing, 2014, pp. 435–442.

[21] N. Nguyen, P.-O. Persson, and J. Peraire, "RANS solutions using high order discontinuous Galerkin methods," in *45th AIAA Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics, Jan. 2007.

[22] S. S. Girimaji, "Partially-averaged Navier-Stokes model for turbulence: A Reynolds-averaged Navier-Stokes to direct numerical simulation bridging method," *Journal of Applied Mechanics*, vol. 73, no. 3, pp. 413–421, Nov. 2005.

[23] F. S. Pereira, G. Vaz, and L. Eca, "An assessment of Scale-Resolving Simulation models for the flow around a circular cylinder," in *Proceeding of THMT-15. Proceedings of the Eighth International Symposium On Turbulence Heat and Mass Transfer*. Begellhouse, 2015.

[24] F. S. Pereira, L. Eça, G. Vaz, and S. S. Girimaji, "Challenges in Scale-Resolving Simulations of turbulent wake flows with coherent structures," *Journal of Computational Physics*, vol. 363, pp. 98–115, Jun. 2018.

[25] E. Hopf, "The partial differential equation $u_t + uu_x = \mu_{xx}$," *Communications on Pure and Applied Mathematics*, vol. 3, no. 3, pp. 201–230, Sep. 1950.

[26] E. F. Toro, "The Riemann problem for the Euler equations," in *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer Berlin Heidelberg, 1997, ch. 4, pp. 115–157.

[27] J.-L. Guermond, B. Popov, and I. Tomas, "Invariant domain preserving discretization-independent schemes and convex limiting for hyperbolic systems," *Computer Methods in Applied Mechanics and Engineering*, vol. 347, pp. 143–175, Apr. 2019.

[28] J. P. Boris and D. L. Book, "Flux-corrected transport," *Journal of Computational Physics*, vol. 135, no. 2, pp. 172–186, Aug. 1997.

[29] C.-W. Shu and S. Osher, "Efficient implementation of essentially non-oscillatory shock-capturing schemes," *Journal of Computational Physics*, vol. 77, no. 2, pp. 439–471, Aug.

1988.

[30] J. Glaubitz, P. Öffner, and T. Sonar, “Application of modal filtering to a spectral difference method,” *Mathematics of Computation*, vol. 87, no. 309, pp. 175–207, Aug. 2017.

[31] P.-O. Persson and J. Peraire, “Sub-cell shock capturing for discontinuous Galerkin methods,” in *44th AIAA Aerospace Sciences Meeting and Exhibit*. American Institute of Aeronautics and Astronautics, Jan. 2006.

[32] C. Cercignani, *The Boltzmann Equation and Its Applications*. Springer New York, 1988.

[33] T. Xiao, “A flux reconstruction kinetic scheme for the Boltzmann equation,” *Journal of Computational Physics*, vol. 447, p. 110689, Dec. 2021.

[34] S. Jaiswal, “An entropy stable scheme for the non-linear Boltzmann equation,” *Journal of Computational Physics*, vol. 463, p. 111289, Aug. 2022.

[35] F. Witherden, A. Farrington, and P. Vincent, “PyFR: An open source framework for solving advection–diffusion type problems on streaming architectures using the flux reconstruction approach,” *Computer Physics Communications*, vol. 185, no. 11, pp. 3028–3040, Nov. 2014.

[36] P. L. Bhatnagar, E. P. Gross, and M. Krook, “A model for collision processes in gases. I. Small amplitude processes in charged and neutral one-component systems,” *Physical Review*, vol. 94, no. 3, pp. 511–525, May 1954.

[37] P. A. Raviart and J. M. Thomas, “A mixed finite element method for 2-nd order elliptic problems,” in *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1977, pp. 292–315.

[38] W. Trojak and F. Witherden, “A new family of weighted one-parameter flux reconstruction schemes,” *Computers & Fluids*, vol. 222, p. 104918, May 2021.

[39] S. K. Godunov, “A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics,” *Matematicheskii Sbornik*, vol. 89, no. 3, pp. 271–306, 1959.

[40] V. Rusanov, “The calculation of the interaction of non-stationary shock waves and obstacles,” *USSR Computational Mathematics and Mathematical Physics*, vol. 1, no. 2, pp. 304–

320, Jan. 1962.

[41] P. Roe, “Approximate Riemann solvers, parameter vectors, and difference schemes,” *Journal of Computational Physics*, vol. 43, no. 2, pp. 357–372, Oct. 1981.

[42] S. F. Davis, “Simplified second-order Godunov-type methods,” *SIAM Journal on Scientific and Statistical Computing*, vol. 9, no. 3, pp. 445–473, May 1988.

[43] F. Bassi and S. Rebay, “A high order discontinuous Galerkin method for compressible turbulent flows,” in *Lecture Notes in Computational Science and Engineering*. Springer Berlin Heidelberg, 2000, pp. 77–88.

[44] P. Fernandez, N.-C. Nguyen, and J. Peraire, “Subgrid-scale modeling and implicit numerical dissipation in DG-based large-eddy simulation,” in *23rd AIAA Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics, Jun. 2017.

[45] R. Schiestel and A. Dejoan, “Towards a new partially integrated transport model for coarse grid and unsteady turbulent flow simulations,” *Theoretical and Computational Fluid Dynamics*, vol. 18, no. 6, pp. 443–468, Dec. 2004.

[46] C. G. Speziale, “Turbulence modeling for time-dependent RANS and VLES: A review,” *AIAA Journal*, vol. 36, no. 2, pp. 173–184, Feb. 1998.

[47] H. F. Fasel, J. Seidel, and S. Wernz, “A methodology for simulations of complex turbulent flows,” *Journal of Fluids Engineering*, vol. 124, no. 4, pp. 933–942, Dec. 2002.

[48] F. R. Menter, “Two-equation eddy-viscosity turbulence models for engineering applications,” *AIAA Journal*, vol. 32, no. 8, pp. 1598–1605, Aug. 1994.

[49] S. S. Girimaji and S. Wallin, “Closure modeling in bridging regions of variable-resolution (VR) turbulence computations,” *Journal of Turbulence*, vol. 14, no. 1, pp. 72–98, Jan. 2013.

[50] J. Fröhlich, C. P. Mellen, W. Rodi, L. Temmerman, and M. A. Leschziner, “Highly resolved large-eddy simulation of separated flow in a channel with streamwise periodic constrictions,” *Journal of Fluid Mechanics*, vol. 526, pp. 19–66, Mar. 2005.

[51] S. Benocci and A. Pinelli, “The role of the forcing term in the large eddy simulation of equilibrium channel flow,” in *Engineering Turbulence Modeling and Experiments 1*. Elsevier

BV, 1990, pp. 287–296.

[52] W. Wang, “A non-body conformal grid method for simulations of laminar and turbulent flows with a compressible large eddy simulation solver,” Ph.D. dissertation, Iowa State University, 2009.

[53] P. Razi, P. Tazraei, and S. Girimaji, “Partially-averaged Navier–Stokes (PANS) simulations of flow separation over smooth curved surfaces,” *International Journal of Heat and Fluid Flow*, vol. 66, pp. 157–171, Aug. 2017.

[54] P. Parnaudeau, J. Carlier, D. Heitz, and E. Lamballais, “Experimental and numerical studies of the flow over a circular cylinder at Reynolds number 3900,” *Physics of Fluids*, vol. 20, no. 8, p. 085101, Aug. 2008.

[55] O. Lehmkuhl, I. Rodríguez, R. Borrell, and A. Oliva, “Low-frequency unsteadiness in the vortex formation region of a circular cylinder,” *Physics of Fluids*, vol. 25, no. 8, p. 085109, Aug. 2013.

[56] F. Witherden, B. Vermeire, and P. Vincent, “Heterogeneous computing on mixed unstructured grids with PyFR,” *Computers & Fluids*, vol. 120, pp. 173–186, Oct. 2015.

[57] L. Ong and J. Wallace, “The velocity field of the turbulent very near wake of a circular cylinder,” *Experiments in Fluids*, vol. 20, no. 6, pp. 441–453, Apr. 1996.

[58] A. G. Kravchenko and P. Moin, “Numerical studies of flow over a circular cylinder at ReD=3900,” *Physics of Fluids*, vol. 12, no. 2, pp. 403–417, Feb. 2000.

[59] X. Ma, G.-S. Karamanos, and G. E. Karniadakis, “Dynamics and low-dimensionality of a turbulent near wake,” *Journal of Fluid Mechanics*, vol. 410, pp. 29–65, May 2000.

[60] T. Dzanic, S. Girimaji, and F. Witherden, “Partially-averaged Navier–Stokes simulations of turbulence within a high-order flux reconstruction framework,” *Journal of Computational Physics*, vol. 456, p. 110992, May 2022.

[61] J. von Neumann and R. D. Richtmyer, “A method for the numerical calculation of hydrodynamic shocks,” *Journal of Applied Physics*, vol. 21, no. 3, pp. 232–237, Mar. 1950.

[62] G. Barter and D. Darmofal, “Shock capturing with higher-order, PDE-based artificial vis-

cosity," in *18th AIAA Computational Fluid Dynamics Conference*. American Institute of Aeronautics and Astronautics, Jun. 2007.

[63] M. Nazarov and J. Hoffman, "Residual-based artificial viscosity for simulation of turbulent compressible flow using adaptive finite element methods," *International Journal for Numerical Methods in Fluids*, vol. 71, no. 3, pp. 339–357, Mar. 2012.

[64] T. Dzanic, W. Trojak, and F. Witherden, "Utilizing time-reversibility for shock capturing in nonlinear hyperbolic conservation laws," *Computers & Fluids*, vol. 247, p. 105652, Oct. 2022.

[65] G.-S. Jiang and C.-C. Wu, "A high-order WENO finite difference scheme for the equations of ideal magnetohydrodynamics," *Journal of Computational Physics*, vol. 150, no. 2, pp. 561–594, Apr. 1999.

[66] C.-W. Shu, "Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws," in *Lecture Notes in Mathematics*. Springer Berlin Heidelberg, 1998, pp. 325–432.

[67] A. Harten, B. Engquist, S. Osher, and S. R. Chakravarthy, "Uniformly high order accurate essentially non-oscillatory schemes, III," *Journal of Computational Physics*, vol. 71, no. 2, pp. 231–303, Aug. 1987.

[68] J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods*. Springer New York, 2008.

[69] C. M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics*, 4th ed. Springer Berlin Heidelberg, 2016, ch. 4.5, pp. 84–90.

[70] A. Harten, "On the symmetric form of systems of conservation laws with entropy," *Journal of Computational Physics*, vol. 49, no. 1, pp. 151–164, Jan. 1983.

[71] E. Tadmor, "A minimum entropy principle in the gas dynamics equations," *Applied Numerical Mathematics*, vol. 2, no. 3-5, pp. 211–219, Oct. 1986.

[72] C. M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics*, 4th ed. Springer Berlin Heidelberg, 2016, ch. 12.8, pp. 471–475.

[73] B. Khobalatte and B. Perthame, "Maximum principle on the entropy and second-order kinetic schemes," *Mathematics of Computation*, vol. 62, no. 205, pp. 119–131, 1994.

[74] P. E. Vincent, P. Castonguay, and A. Jameson, "A new class of high-order energy stable flux reconstruction schemes," *Journal of Scientific Computing*, vol. 47, no. 1, pp. 50–72, Sep. 2010.

[75] X. Zhang and C.-W. Shu, "On maximum-principle-satisfying high order schemes for scalar conservation laws," *Journal of Computational Physics*, vol. 229, no. 9, pp. 3091–3120, May 2010.

[76] ——, "On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes," *Journal of Computational Physics*, vol. 229, no. 23, pp. 8917–8934, Nov. 2011.

[77] X. Zhang, Y. Xia, and C.-W. Shu, "Maximum-principle-satisfying and positivity-preserving high order discontinuous Galerkin schemes for conservation laws on triangular meshes," *Journal of Scientific Computing*, vol. 50, no. 1, pp. 29–62, Feb. 2011.

[78] R. Courant, K. Friedrichs, and H. Lewy, "Uber die partiellen differenzengleichungen der mathematischen physik," *Mathematische Annalen*, vol. 100, no. 1, pp. 32–74, Dec. 1928.

[79] T. Chen and C.-W. Shu, "Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws," *Journal of Computational Physics*, vol. 345, pp. 427–461, Sep. 2017.

[80] P. D. Lax, "Weak solutions of nonlinear hyperbolic equations and their numerical computation," *Communications on Pure and Applied Mathematics*, vol. 7, no. 1, pp. 159–193, Feb. 1954.

[81] E. F. Toro, M. Spruce, and W. Speares, "Restoration of the contact surface in the HLL-Riemann solver," *Shock Waves*, vol. 4, no. 1, pp. 25–34, Jul. 1994.

[82] X. Zhang and C.-W. Shu, "Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms," *Journal of Computational Physics*, vol. 230, no. 4, pp. 1238–1248, Feb. 2011.

[83] L. Demkowicz, J. Oden, and W. Rachowicz, “A new finite element method for solving compressible Navier-Stokes equations based on an operator splitting method and h-p adaptivity,” *Computer Methods in Applied Mechanics and Engineering*, vol. 84, no. 3, pp. 275–326, Dec. 1990.

[84] S. Gottlieb, C.-W. Shu, and E. Tadmor, “Strong stability-preserving high-order time discretization methods,” *SIAM Review*, vol. 43, no. 1, pp. 89–112, Jan. 2001.

[85] G. A. Sod, “A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws,” *Journal of Computational Physics*, vol. 27, no. 1, pp. 1–31, Apr. 1978.

[86] P. Woodward and P. Colella, “The numerical simulation of two-dimensional fluid flow with strong shocks,” *Journal of Computational Physics*, vol. 54, no. 1, pp. 115–173, Apr. 1984.

[87] D. S. Balsara, “Self-adjusting, positivity preserving high order schemes for hydrodynamics and magnetohydrodynamics,” *Journal of Computational Physics*, vol. 231, no. 22, pp. 7504–7517, Sep. 2012.

[88] K. Wu and C.-W. Shu, “A provably positive discontinuous Galerkin method for multidimensional ideal magnetohydrodynamics,” *SIAM Journal on Scientific Computing*, vol. 40, no. 5, pp. B1302–B1329, Jan. 2018.

[89] G. Taylor and A. Green, “Mechanism of the production of small eddies from large ones,” *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, vol. 158, no. 895, pp. 499–521, Feb. 1937.

[90] S. C. Spiegel, J. R. DeBonis, and H. Huynh, “Overview of the NASA Glenn flux reconstruction based high-order unstructured grid code,” in *54th AIAA Aerospace Sciences Meeting*. American Institute of Aeronautics and Astronautics, Jan. 2016.

[91] W. Trojak, N. Vadlamani, J. Tyacke, F. Witherden, and A. Jameson, “Artificial compressibility approaches in flux reconstruction for incompressible viscous flow simulations,” *Computers & Fluids*, vol. 247, p. 105634, Oct. 2022.

[92] V. Daru and C. Tenaud, “Evaluation of TVD high resolution schemes for unsteady viscous shocked flows,” *Computers & Fluids*, vol. 30, no. 1, pp. 89–113, Sep. 2000.

[93] J.-L. Guermond, M. Maier, B. Popov, and I. Tomas, "Second-order invariant domain preserving approximation of the compressible Navier–Stokes equations," *Computer Methods in Applied Mechanics and Engineering*, vol. 375, p. 113608, Mar. 2021.

[94] J. Chu and J. M. Luckring, "Experimental surface pressure data obtained on 65-deg delta wing across Reynolds number and Mach number ranges," in *NASA Technical Memorandum 4645*. NASA Technical Reports, Feb. 1996.

[95] J. Luckring, "Reynolds number and leading-edge bluntness effects on a 65-deg delta wing," in *40th AIAA Aerospace Sciences Meeting & Exhibit*. American Institute of Aeronautics and Astronautics, Jan. 2002.

[96] T. Dzanic and L. Martinelli, "Higher-order implicit large eddy simulations of a VFE-2 delta wing," in *AIAA Scitech 2019 Forum*. American Institute of Aeronautics and Astronautics, Jan. 2019.

[97] C. Baranger, Y. Dauvois, G. Marois, J. Mathé, J. Mathiaud, and L. Mieussens, "A BGK model for high temperature rarefied gas flows," *European Journal of Mechanics - B/Fluids*, vol. 80, pp. 1–12, Mar. 2020.

[98] R. Loek Van Heyningen, "Discontinuous Galerkin solutions of the Boltzmann equation: spectral collocation and moment methods," Master's thesis, Massachusetts Institute of Technologyy, 2021.

[99] L. Mieussens, "Discrete-velocity models and numerical schemes for the Boltzmann-BGK equation in plane and axisymmetric geometries," *Journal of Computational Physics*, vol. 162, no. 2, pp. 429–466, Aug. 2000.

[100] B. Evans, K. Morgan, and O. Hassan, "A discontinuous finite element solution of the Boltzmann kinetic equation in collisionless and BGK forms for macroscopic gas flows," *Applied Mathematical Modelling*, vol. 35, no. 3, pp. 996–1015, Mar. 2011.

[101] F. Bernard, A. Iollo, and G. Puppo, "A local velocity grid approach for BGK equation," *Communications in Computational Physics*, vol. 16, no. 4, pp. 956–982, Oct. 2014.

[102] E. M. Shakhov, "Generalization of the Krook kinetic relaxation equation," *Fluid Dynamics*,

vol. 3, no. 5, pp. 95–96, 1972.

[103] P. Wang, L. Wu, M. T. Ho, J. Li, Z.-H. Li, and Y. Zhang, "The kinetic Shakhov–Enskog model for non-equilibrium flow of dense gases," *Journal of Fluid Mechanics*, vol. 883, Nov. 2019.

[104] M. M. R. Williams, "A review of the rarefied gas dynamics theory associated with some classical problems in flow and heat transfer," *Zeitschrift für angewandte Mathematik und Physik*, vol. 52, no. 3, pp. 500–516, May 2001.

[105] B. Cockburn and C.-W. Shu, "Runge–Kutta discontinuous Galerkin methods for convection-dominated problems," *Journal of Scientific Computing*, vol. 16, no. 3, pp. 173–261, 2001.

[106] F. Bassi and S. Rebay, "A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations," *Journal of Computational Physics*, vol. 131, no. 2, pp. 267–279, Mar. 1997.

[107] M. Dumbser, O. Zanotti, R. Loubère, and S. Diot, "A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws," *Journal of Computational Physics*, vol. 278, pp. 47–75, Dec. 2014.

[108] V. Springel, "E pur si muove: Galilean-invariant cosmological hydrodynamical simulations on a moving mesh," *Monthly Notices of the Royal Astronomical Society*, vol. 401, no. 2, pp. 791–851, Jan. 2010.

[109] T. Dzanic and F. Witherden, "Positivity-preserving entropy-based adaptive filtering for discontinuous spectral element methods," *Journal of Computational Physics*, vol. 468, p. 111501, Nov. 2022.

[110] G. Cao, L. Pan, and K. Xu, "High-order gas-kinetic scheme with parallel computation for direct numerical simulation of turbulent flows," *Journal of Computational Physics*, vol. 448, p. 110739, Jan. 2022.

[111] T. Dzanic, F. Witherden, and L. Martinelli, "A positivity-preserving and conservative high-order flux reconstruction method for the polyatomic Boltzmann–BGK equation," *arXiv*, Oct. 2022.

[112] J. Weiss, "A tutorial on the proper orthogonal decomposition," in *AIAA Aviation 2019 Forum*. American Institute of Aeronautics and Astronautics, Jun. 2019.

APPENDIX A

TURBULENCE MODEL PARAMETERS

The constants for the SST model are as tabulated by Menter [48]. Certain constants in the SST model are defined with an inner (1) and outer (2) constant, such that its value is defined as

$$\phi = F_1\phi_1 + (1 - F_1)\phi_2$$

for the following inner and outer values

$$\begin{aligned} \alpha_1 &= 0.5532, & \alpha_2 &= 0.4403, \\ \beta_1 &= 0.075, & \beta_2 &= 0.0828, \\ \sigma_{k1} &= 0.85, & \sigma_{k2} &= 1.0, \\ \sigma_{\omega 1} &= 0.5, & \sigma_{\omega 2} &= 0.856. \end{aligned}$$

The remaining constants are explicitly defined as

$$a_1 = 0.31, \qquad \beta^* = 0.09, \qquad Pr_t = 0.9.$$

APPENDIX B

PROPER ORTHOGONAL DECOMPOSITION

The POD method, as described by Weiss [112], can be used to decompose a time-dependent velocity fluctuation field, $u'(\mathbf{x}, t)$, into a set of orthonormal spatial modes, $\Phi_k(\mathbf{x})$, such that

$$u'(\mathbf{x}, t) = \sum_{k=1}^{\infty} a_k(t)\Phi_k(\mathbf{x}), \tag{B.1}$$

for some temporal coefficient $a_k(t)$. In the discrete form, this is performed by first forming a matrix of "snapshots" of the solution,

$$\mathbf{U} = \begin{pmatrix} u_{11} & \dots & u_{1n} \\ u_{21} & \dots & u_{2n} \\ \vdots & \ddots & \vdots \\ u_{m1} & \dots & u_{mn} \end{pmatrix} = \begin{pmatrix} u'(\mathbf{x}_1, t_1) & \dots & u'(\mathbf{x}_n, t_1) \\ u'(\mathbf{x}_1, t_2) & \dots & u'(\mathbf{x}_n, t_2) \\ \vdots & \ddots & \vdots \\ u'(\mathbf{x}_1, t_m) & \dots & u'(\mathbf{x}_n, t_m) \end{pmatrix}, \tag{B.2}$$

where n denotes the number of sample points in space and m denotes the number of sample points in time. A correlation matrix $\mathbf{C}$ can then be defined as

$$\mathbf{C} = \frac{1}{m-1}\mathbf{U}\mathbf{U}^T. \tag{B.3}$$

By performing an eigendecomposition of $\mathbf{C}$, a set of eigenvectors $\mathbf{\Psi}_k$ and their associated eigenvalues λ_k can be extracted and sorted such that $|\lambda_k| > |\lambda_{k+1}|$. The spatial modes are then defined as

$$\mathbf{\Phi}_k = \mathbf{U}^T\mathbf{\Psi}_k \tag{B.4}$$

and normalized such that $\|\mathbf{\Phi}_k\|_2 = 1$.

In this work, the snapshots were formed by resampling the flow field to a two-dimensional grid on the spanwise periodic boundary. 800 and 200 points were sampled in the streamwise and normal directions, respectively, and 20,000 snapshots were obtained over the time-averaging horizon, corresponding to values of $n = 160,000$ and $m = 20,000$.

APPENDIX C

ALGORITHMIC IMPLEMENTATION OF ENTROPY FILTERING

An in-depth description of the computational details of the proposed filtering approach is presented below. In Algorithm C.0.1, the computation of the entropy constraints is presented. This method computes the discrete local minima of the entropy function σ_*^k at the solution nodes for each element Ω_k. The entropy constraint $\sigma_{\min}^k$ is then taken as the minima of σ_*^k across the face-neighbors of Ω_k, including Ω_k itself.

Algorithm C.0.1 Calculate entropy constraints

```
INPUT: $\mathbf{u}$ // Global solution
OUTPUT: $\sigma_{\min}$ // Global entropy constraints

GetEntropyConstraints($\mathbf{u}$) :
    for $\Omega_k \in \Omega$ do
        $\sigma_*^k = \min\limits_{i \in S(\Omega_k)} \sigma(\mathbf{u}_i)$
    end for
    for $\Omega_k \in \Omega$ do
        $\sigma_{\min}^k = \min\limits_{j \in \mathcal{A}(k)} \sigma_*^j$
    end for
    return $\sigma_{\min}$
```

In Algorithm C.0.2, the element-wise filtering operation is outlined. This method takes in the unfiltered element-wise solution $\mathbf{u}$ and its associated entropy constraint $\sigma_{\min}$ as the input in addition to a precomputed Vandermonde matrix $\mathbf{V}$ (from the nodal basis to a modal basis). If the discrete solution is within bounds, it returns the unfiltered solution. If not, it performs n_{iters} iterations of the bisection approach. The initial guess for the upper bound of the bisection method is taken as $-\log(\varepsilon)$ for some value of machine precision ε as this effectively approximates an infinite filter strength down to machine precision.

Algorithm C.0.2 Filter solution

```
INPUT: $\mathbf{u}$, $\sigma_{\min}$ // Element-wise solution, element-wise entropy constraint
OUTPUT: $\widetilde{\mathbf{u}}$ // Element-wise filtered solution

FilterSolution($\mathbf{u}$, $\sigma_{\min}$) :
  // PARAMETERS
  $\rho_{\min} = 10^{-8}$ // Minimum density
  $P_{\min} = 10^{-8}$ // Minimum pressure
  $\epsilon_{\sigma} = 10^{-4}$ // Entropy tolerance
  $\varepsilon = 10^{-8}$ // Machine precision (FP32)
  $n_{\text{iters}} = 20$ // Number of filter iterations

  // Compute minima within element
  $[\rho_*, P_*, \sigma_*] = \min_{i \in S} [\rho(\mathbf{u}_i), P(\mathbf{u}_i), \sigma(\mathbf{u}_i)]$

  // Return unfiltered solution if within bounds
  if ($\rho_* \geq \rho_{\min}$) & ($P_* \geq P_{\min}$) & ($\sigma_* \geq \sigma_{\min} - \epsilon_{\sigma}$) then
    return $\mathbf{u}$
  else
    $\zeta_1$ = 0.0
    $\zeta_2$ = $-\log(\varepsilon)$

    // Compute modal basis
    $\hat{\mathbf{u}} = \mathbf{V}^{-1}\mathbf{u}$

    for $i \in \{0, .., n_{\text{iters}}\}$ do
      $\zeta_3 = 0.5(\zeta_1 + \zeta_2)$
      $\widetilde{\mathbf{u}} = \mathbf{V}\left(\hat{\mathbf{u}} \odot \exp(-\zeta_3 \mathbf{p}^2)\right)$

      // Compute minima within element
      $[\rho_*, P_*, \sigma_*] = \min_{i \in S} [\rho(\widetilde{\mathbf{u}}_i), P(\widetilde{\mathbf{u}}_i), \sigma(\widetilde{\mathbf{u}}_i)]$

      // Choose new midpoint
      if ($\rho_* \geq \rho_{\min}$) & ($P_* \geq P_{\min}$) & ($\sigma_* \geq \sigma_{\min} - \epsilon_{\sigma}$) then
        $\zeta_2 = \zeta_3$
      else
        $\zeta_1 = \zeta_3$
      end if
    end for

    // Compute filtered solution with bounds-preserving filter
    $\zeta = \zeta_2$
    $\widetilde{\mathbf{u}} = \mathbf{V}\left(\hat{\mathbf{u}} \odot \exp(-\zeta \mathbf{p}^2)\right)$
    return $\widetilde{\mathbf{u}}$

  end if
```

In Algorithm C.0.3, an outline of the temporal integration approach utilizing an SSP-RK3 scheme *for purely hyperbolic systems* is given. For a given time step Δt and number of time steps nt, the method computes each intermediate stage of the temporal integration scheme and filters the solution based on the constraints computed on the previous stage. A similar outline is shown in Algorithm C.0.4 for mixed hyperbolic-parabolic systems that details the operator splitting approach.

Algorithm C.0.3 SSP-RK3 stepper for hyperbolic systems

```
// PARAMETERS
$\Delta t$ = ... // Time step
$n_t$ = ... // Number of time steps

// Step in time
  for $n \in \{0, .., n_t\}$ do
    // Calculate entropy constraints for each element
    $\sigma_{\min}$ = GetEntropyConstraints($\mathbf{u}^n$)
    // Compute first stage
    $\mathbf{u}^{(1)} = \mathbf{u}^n + \Delta t \left(-\boldsymbol{\nabla}\cdot\mathbf{F}(\mathbf{u}^n)\right)$
    // Check if constraints are satisfied and filter if not
    for $\Omega_k \in \Omega$ do
      $\mathbf{u}_k^{(1)}$ = FilterSolution($\mathbf{u}_k^{(1)}$, $\sigma_{\min}^k$)
    end for

    // Repeat procedure for second stage
    $\sigma_{\min}$ = GetEntropyConstraints($\mathbf{u}^{(1)}$)
    $\mathbf{u}^{(2)} = \frac{3}{4}\mathbf{u}^n + \frac{1}{4}\mathbf{u}^{(1)} + \Delta t \left(-\boldsymbol{\nabla}\cdot\mathbf{F}(\mathbf{u}^{(1)})\right)$
    for $\Omega_k \in \Omega$ do
      $\mathbf{u}_k^{(2)}$ = FilterSolution($\mathbf{u}_k^{(2)}$, $\sigma_{\min}^k$)
    end for

    // Repeat procedure for third stage and compute next temporal step
    $\sigma_{\min}$ = GetEntropyConstraints($\mathbf{u}^{(2)}$)
    $\mathbf{u}^{n+1} = \frac{1}{3}\mathbf{u}^n + \frac{2}{3}\mathbf{u}^{(2)} + \Delta t \left(-\boldsymbol{\nabla}\cdot\mathbf{F}(\mathbf{u}^{(2)})\right)$
    for $\Omega_k \in \Omega$ do
      $\mathbf{u}_k^{n+1}$ = FilterSolution($\mathbf{u}_k^{n+1}$, $\sigma_{\min}^k$)
    end for
  end for
```

Algorithm C.0.4 SSP-RK3 stepper for mixed hyperbolic-parabolic systems

```
// PARAMETERS
$\Delta t$ = ... // Time step
$n_t$ = ... // Number of time steps

// Step in time
  for $n \in \{0, .., nt\}$ do
    // Calculate entropy constraints for each element
    $\sigma_{\min} =$ GetEntropyConstraints($\mathbf{u}^n$)
    // Compute inviscid and viscous first stages
    $\mathbf{u}^{(1,I)} = \mathbf{u}^n + \Delta t \left(-\boldsymbol{\nabla} \cdot \mathbf{F}_I(\mathbf{u}^n)\right)$
    $\Delta \mathbf{u}^{(1,V)} = \Delta t \left(-\boldsymbol{\nabla} \cdot \mathbf{F}_V(\mathbf{u}^n)\right)$
    // Check if constraints are satisfied and filter if not
    for $\Omega_k \in \Omega$ do
      // Enforce positivity/entropy constraints for inviscid component
      $\mathbf{u}_k^{(1,I)}$ = FilterSolution($\mathbf{u}_k^{(1,I)}$, $\sigma_{\min}^k$)
      // Add viscous component to filtered inviscid component
      $\mathbf{u}_k^{(1)} = \mathbf{u}_k^{(1,I)} + \Delta \mathbf{u}_k^{(1,V)}$
      // Enforce only positivity constraints for inviscid + viscous component
      $\mathbf{u}_k^{(1)}$ = FilterSolution($\mathbf{u}_k^{(1)}$, $-\infty$)
    end for

    // Repeat procedure for second stage
    $\sigma_{\min} =$ GetEntropyConstraints($\mathbf{u}^{(1)}$)
    $\mathbf{u}^{(2,I)} = \frac{3}{4}\mathbf{u}^n + \frac{1}{4}\mathbf{u}^{(1)} + \Delta t \left(-\boldsymbol{\nabla} \cdot \mathbf{F}_I(\mathbf{u}^{(1)})\right)$
    $\Delta \mathbf{u}^{(2,V)} = \Delta t \left(-\boldsymbol{\nabla} \cdot \mathbf{F}_V(\mathbf{u}^{(1)})\right)$
    for $\Omega_k \in \Omega$ do
      $\mathbf{u}_k^{(2,I)}$ = FilterSolution($\mathbf{u}_k^{(2,I)}$, $\sigma_{\min}^k$)
      $\mathbf{u}_k^{(2)} = \mathbf{u}_k^{(2,I)} + \Delta \mathbf{u}_k^{(2,V)}$
      $\mathbf{u}_k^{(2)}$ = FilterSolution($\mathbf{u}_k^{(2)}$, $-\infty$)
    end for

    // Repeat procedure for third stage and compute next temporal step
    $\sigma_{\min} =$ GetEntropyConstraints($\mathbf{u}^{(2)}$)
    $\mathbf{u}^{n+1,I} = \frac{1}{3}\mathbf{u}^n + \frac{2}{3}\mathbf{u}^{(2)} + \Delta t \left(-\boldsymbol{\nabla} \cdot \mathbf{F}_I(\mathbf{u}^{(2)})\right)$
    $\Delta \mathbf{u}^{n+1,V} = \Delta t \left(-\boldsymbol{\nabla} \cdot \mathbf{F}_V(\mathbf{u}^{(2)})\right)$
    for $\Omega_k \in \Omega$ do
      $\mathbf{u}_k^{n+1,I}$ = FilterSolution($\mathbf{u}_k^{n+1,I}$, $\sigma_{\min}^k$)
      $\mathbf{u}_k^{n+1} = \mathbf{u}_k^{n+1} + \Delta \mathbf{u}_k^{n+1,V}$
      $\mathbf{u}_k^{n+1}$ = FilterSolution($\mathbf{u}_k^{n+1}$, $-\infty$)
    end for
  end for
```