Energy and entropy stable numerical methods with injected boundary conditions

Anita Gjesteland

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2024

UNIVERSITY OF BERGEN

Energy and entropy stable numerical methods with injected boundary conditions

Anita Gjesteland

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 26.01.2024

Year: 2024

Title: Energy and entropy stable numerical methods with injected boundary conditions

Name: Anita Gjesteland

Print: Skipnes Kommunikasjon / University of Bergen

Preface

This dissertation is submitted as a partial fulfilment of the requirements for the degree of Philosophiae Doctor (Ph.D.) at the University of Bergen. The advisory committee has consisted of Professor Magnus Svärd and Professor Henrik Kalisch.

Acknowledgements

I would like to show my gratitude towards my two advisors, Magnus Svärd and Henrik Kalisch. In particular, thank you, Magnus, for your encouragement and support, and for shaping me into a better researcher. I appreciate all your excellent explanations, impeccable manuscript feedbacks and that your door has always been open. I feel lucky to have had the chance to work with you, and it has been a lot of fun!

I would also like to thank David Del Rey Fernández for beeing a co-author on one of the papers, and for hosting me for one month at the University of Waterloo. I have greatly enjoyed working with you.

Finally, a big thank you to my colleagues at the Department of Mathematics, in particular Erlend, Maria, Anders and Stine, and to my good friends and supporting family.

Abstract

The compressible Navier-Stokes equation subject to both adiabatic wall boundary conditions and far-field boundary conditions are studied in this thesis. Although the well-posedness of these equations is generally unknown, they are of wide interest and are extensively used in computational fluid dynamics. A result by Strang (1964) states that if a non-linear problem is discretised using a difference method that is linearly stable, then this method is convergent for smooth solutions. That is, there exists theory we can use in the analysis of the Navier-Stokes equations. Thus, we study linear well-posedness and stability of numerical schemes both in the context of the compressible Navier-Stokes equations, but also linear partial differential equations as model problems. Furthermore, entropy estimates are derived for the fully non-linear Navier-Stokes equations, which pose as an admissibility criterion for the relevant weak solution we seek; it should additionally satisfy the second law of thermodynamics.

The main focus of this work is the stable imposition of the adiabatic wall and far-field boundary conditions for the Navier-Stokes equations. In particular, we prove that the no-slip condition can be imposed strongly and still yield an entropy estimate when used in combination with diagonal-norm summation-by-parts (SBP) operators with diagonal boundary operators. Furthermore, we introduce a new methodology for setting far-field boundary conditions, and prove that it leads to an entropy stable scheme for the compressible Navier-Stokes equations. The procedure is additionally linearly well-posed. Throughout, we employ SBP operators due to their remarkable stability properties. We also prove that a slightly modified version of the finite-volume SBP approximation of the second-derivative given by Chandrashekar (2016) is (weakly) consistent, thus making it suitable for discretising the viscous terms of the Navier-Stokes equations on unstructured grids.

Sammendrag

I denne avhandlingen studerer vi de kompressible Navier-Stokes-likningene formulert med både adiabatiske veggrandvilkår og fjernfeltvilkår. Selv om det er ukjent om disse likningene er velformulerte er de av stor interesse, og de er mye brukt innen numerisk fluiddynamikk. Et resultat av Strang (1964) sier at for ikke-lineære problem diskretisert ved hjelp av en differansemetode som er lineærstabil, er denne metoden konvergent for glatte løsninger. Altså finnes det teori vi kan bruke i analysen av Navier-Stokes-likningene. Derfor studerer vi her teori for velformulerte lineære problem, og stabilitet for numeriske metoder. Dette gjøres både for de kompressible Navier-Stokes-likningene, men også for lineære partielle differensiallikninger som modellproblem. Videre utleder vi entropiestimat for de ikke-lineære Navier-Stokes-likningene, et estimat som virker som et kriterium for den svake løsningen vi leter etter; den skal i tillegg til likningene tilfredsstille termodynamikkens andre lov.

Hovedfokuset ved dette arbeidet er stabil håndtering av de adiabatiske veggrandvilkårene og fjernfeltvilkår for Navier-Stokes-likningene. Vi beviser at heftelsesvilkåret (eng.: no-slip condition) kan bli implementert eksakt og fremdeles resultere i et entropiestimat når teknikken brukes i kombinasjon med delvissummasjonsoperatorer (SBP-operatorer) som har diagonale normmatriser og randmatriser. Vi introduserer også en ny metodikk for å sette fjernfeltvilkår, og beviser at den fører til et entropistabilt skjema for de kompressible Navier-Stokes-likningene. Teknikken er i tillegg lineært velformulert. Gjennom hele arbeidet bruker vi SBP-operatorer på grunn av deres gode stabilitetsegenskaper. Vi beviser også at en litt endret versjon av SBP-operatoren som tilnærmer den andrederiverte ved hjelp av endelig-volummetoden gitt av Chandrashekar (2016) er (svakt) konsistent, noe som gjør den egnet til å diskretisere de viskøse leddene i Navier-Stokes-likningene på ustrukturerte gitter.

List of papers

Part II, Scientific results, consists of the following four papers.

A. A. Gjesteland and M. Svärd. Entropy stability for the compressible Navier-Stokes equations with strong imposition of the no-slip boundary condition, *Journal of Computational Physics* **470**, 111572, 2022.

B. A. Gjesteland and M. Svärd. Convergence of Chandrashekar's Second-Derivative Finite-Volume Approximation, *Journal of Scientific Computing* **96**, 46, 2023.

C. A. Gjesteland, D. Del Rey Fernández and M. Svärd. Injected Dirichlet boundary conditions for general diagonal-norm SBP operators, *Under review*, 2023.

D. M. Svärd and A. Gjesteland. Entropy stable far-field boundary conditions for the compressible Navier-Stokes equations, *Submitted*, 2023.

Contents

Part I

Background

1. Introduction

This thesis concerns the numerical approximation of the compressible Navier-Stokes equations. These equations constitute a system of non-linear partial differential equations that model the flow of viscous and heat conducting compressible fluids. Herein, we are interested in the case where the fluid in question is air, and the problems are those arising in aerodynamics. In particular, the focus is on stable impositions of the adiabatic wall and far-field boundary conditions. These boundary conditions can be used in combination, for instance when modelling the airflow past an airplane wing.

When solving initial-boundary-value problems, they should ideally be *well-posed* such that the existence of a unique solution depending continuously on the problem data, is guaranteed. The well-posedness of the compressible Navier-Stokes equations is generally unknown, but the theory of linear well-posedness, which is well established in the literature, is often used in the analysis of these equations. Additionally, the theory of entropy can also be used, providing some form of non-linear stability estimates. Herein, we discuss the two concepts briefly.

We start by considering the linear well-posedness theory, which can be found in the books [16, 21]. We study the 1-D advection equation as a model problem to introduce the topic. We focus on the stability aspect in the definition of well-posedness, that is, we derive an a priori estimate for the solution. Throughout this thesis, a priori estimates obtained for linear problems are derived using the energy method, for which integration-by-parts (IBP) is fundamental. For many linear problems, the existence of a unique solution is closely related to the derived energy estimates (see [16]). The well-posedness of the linearised compressible Navier-Stokes equations has been studied in various papers (see e.g. [18, 28, 29, 37, 35]). However, the a priori estimates derived for the linearised equations are generally not sufficient to infer well-posedness of the original non-linear problem.

The theory of entropy is thus sometimes used to obtain non-linear estimates, and we proceed by discussing this topic. The idea is to introduce an additional entropy inequality to the problem. In order for a (weak) solution of the original non-linear problem to be deemed *physically relevant* (see [40]), it should additionally satisfy this entropy inequality. When deriving entropy estimates for the compressible Navier-Stokes equations, we

also rely on integration-by-parts.

Since the ultimate goal is to solve the Navier-Stokes equations numerically, we carry over the aforementioned concepts to the semi-discrete setting. As previously mentioned, the use of IBP is common to both linear stability estimates and entropy estimates. Thus, we utilise summation-by-parts (SBP) operators to approximate the spatial derivatives in the continuous problem. SBP is the discrete counterpart to IBP, and SBP operators thus allow us to mimic the continuous estimates. However, in the semi-discrete setting, difficulties may arise in the imposition of the boundary conditions. To obtain semi-discrete estimates analogous to the ones derived in the continuous setting, it is therefore important to use stable boundary procedures.

Throughout this thesis, we focus on *semi-discrete* schemes, meaning that we keep the temporal variable continuous. For linearly stable semi-discrete schemes, the fully discretised problem is stable if time is integrated using an appropriate Runge-Kutta method (see [23]). In Paper A, we have used a strong stability preserving Runge-Kutta method (see e.g. [13]) for time integration.

The main contribution of this thesis is threefold. First, we show in Paper A ([10]) and subsequently extend the results in Paper C ([12]), that the injection method for imposing Dirichlet boundary conditions strongly results in stable schemes when used in combination with diagonal-norm SBP operators with diagonal boundary operators. Specifically, we show in Paper C that the combination leads to linearly stable schemes for the advection and advection-diffusion equations with both homogeneous and in-homogeneous boundary data. For the compressible Navier-Stokes equations, when the methodology is used to impose the homogeneous no-slip boundary condition it results in entropy conservative/stable schemes. The proof does not impose any restrictions on the accuracy of the SBP operators. For two specific finite-difference SBP operators, the scheme for the 1-D equations is shown to be linearly stable in Paper A ([10]). Second, we propose a new methodology for imposing far-field boundary conditions for the compressible Navier-Stokes equations. Normally, characteristic far-field boundary conditions are used (see Section 3.2.1 for a discussion), and they are linearly well-posed. The new methodology leads to a boundary procedure that is also linearly well-posed, but additionally, we can prove that it leads to an entropy estimate. Finally, with the motivation of discretising the viscous terms of the Navier-Stokes equations using a finite-volume method on unstructured grids, we study an SBP operator approximating the second derivative derived in [5]. By showing convergence of a semi-discrete solution to a weak solution of the heat equation, we establish (weak) consistency of a slightly altered version of the approximation in Paper B ([11]).

2. The compressible Navier-Stokes equations

We start by introducing the compressible Navier-Stokes equations, the motivation behind the work of this thesis.

Let $\mathsf{u} = \Big[\rho, m, n, E\Big]^\top$ denote the vector of conserved variables; density, momentum (in x- and y-direction, respectively) and total energy. Furthermore, let Ω represent an open and bounded polygonal domain in two spatial dimensions, and suppose that a solid body occupies a subdomain $\Omega^{\mathsf{s}} \subset \Omega$ with boundary $\partial\Omega^{\mathsf{s}}$. Then the compressible Navier-Stokes equations on $\Omega \setminus \partial\Omega^{\mathsf{s}}$ can be stated as

$$\mathsf{u}_t + \mathsf{f}^{\mathsf{I}}(\mathsf{u})_x + \mathsf{g}^{\mathsf{I}}(\mathsf{u})_y = \mathsf{f}^{\mathsf{V}}(\mathsf{u}, \mathsf{u}_x, \mathsf{u}_y)_x + \mathsf{g}^{\mathsf{V}}(\mathsf{u}, \mathsf{u}_x, \mathsf{u}_y)_y. \tag{2.1}$$

The inviscid terms of (2.1) are given by

$$\mathsf{f}^{\mathsf{I}}(\mathsf{u}) = \begin{bmatrix} \rho u \\ \rho u^2 + p \\ \rho uv \\ u(E+p) \end{bmatrix}, \qquad \mathsf{g}^{\mathsf{I}}(\mathsf{u}) = \begin{bmatrix} \rho v \\ \rho uv \\ \rho v^2 + p \\ v(E+p) \end{bmatrix},$$

with $u = \frac{m}{\rho}$, $v = \frac{n}{\rho}$ denoting the velocity components in x- and y-direction, respectively, and $p = (\gamma - 1)(E - \frac{\rho u^2 + \rho v^2}{2})$ denoting the pressure, where $\gamma = \frac{c_p}{c_v}$ is the ratio of the specific heats at constant pressure and volume. Furthermore, the viscous fluxes take the forms

$$\mathsf{f}^{\mathsf{v}}(\mathsf{u}, \mathsf{u}_x, \mathsf{u}_y) = \begin{bmatrix} 0 \\ 2\mu u_x + \lambda(u_x + v_y) \\ \mu(u_y + v_x) \\ 2\mu u u_x + \lambda u(u_x + v_y) + \mu v(u_y + v_x) + \kappa T_x \end{bmatrix},$$

$$\mathsf{g}^{\mathsf{v}}(\mathsf{u}, \mathsf{u}_x, \mathsf{u}_y) = \begin{bmatrix} 0 \\ \mu(u_y + v_x) \\ 2\mu v_y + \lambda(u_x + v_y) \\ 2\mu v v_y + \lambda v(u_x + v_y) + \mu u(u_y + v_x) + \kappa T_y \end{bmatrix},$$

where $T = \frac{p}{\rho \mathcal{R}}$ denotes the temperature, and $\mathcal{R}$ represents the gas constant. Lastly, μ and λ are the viscosity parameters, and κ is the heat conductivity parameter. Throughout, we assume Stokes' hypothesis such that $\lambda = -\frac{2}{3}\mu$.

The system of equations (2.1) is augmented by proper boundary and initial conditions to form an initial-boundary-value problem. Here, we focus on the boundary conditions. In the field of aerodynamics, the adiabatic wall and far-field boundary conditions are common. These are used in combination, for example when simulating an external flow around an airfoil. In this case, the adiabatic wall conditions are used at the boundary of the airfoil. They take the forms

$$u, v = 0, \qquad \frac{\partial T}{\partial \vec{n}} = 0, \qquad \text{on } \partial\Omega^{\mathsf{s}}, \tag{2.2}$$

where $\frac{\partial T}{\partial \vec{n}}$ denotes the normal derivative of the temperature. They model the situation where the fluid velocity at the boundary of the airfoil is zero relative to that of the airfoil, and the zero heat flux through its boundary.

At a distance far away from the airfoil, we expect the flow to be unaffected by it, and that all variables approach their *free-stream* values. However, in practical simulations, the spatial domain cannot be arbitrarily large. Far-field boundary conditions are used at the artificial boundaries that occur when the domain is truncated. Ideally, the far-field boundary conditions would be stated such that the fluid flows through the artificial boundary as if it is not there . Common far-field boundary conditions for the compressible Navier-Stokes equations are characteristic boundary conditions that are derived from the linearised equations, see e.g. [18, 28, 37]. We further discuss the form of the far-field boundary conditions in Section 3.2.1.

3. Continuous theory

This chapter is devoted to a discussion on linear well-posedness and entropy theory. Moreover, we present an entropy estimate for the compressible Navier-Stokes equation (2.1).

3.1 Theory of linear well-posedness

When solving initial-boundary-value problems, we would like them to be well-posed. Well-posedness guarantees the existence of a unique solution that depends continuously on the problem data. Generally, we cannot establish well-posedness of the compressible Navier-Stokes equations. However, if they are linearised, we can prove that they are linearly well-posed when augmented by proper boundary conditions. It seems reasonable to require that non-linear problems are linearly well-posed. If a linearised problem is not stable against perturbations in the problem data, we may expect non-linearities to further amplify the unbounded effects. In fact, when discussing the linearisation principle, Gustafsson, Kreiss and Oliger argue that it is doubtful if the solution can be calculated if the linearised problem is not stable (see [16]).

Although the goal is to solve the compressible Navier-Stokes equations numerically, we study also their linearised version in this thesis (see Paper A) and other linear PDEs as they can provide some insights into the non-linear problem, and may be better suited to exemplify various concepts. In this thesis, the energy method can be used to prove well-posedness of the linear problems we consider.

We give a brief discussion of linear well-posedness, and refer the interested reader to the books [16, 21] for a comprehensive introduction.

Let $\Omega = (0, 1)$ denote a spatial domain. We consider the linear advection equation as a model problem, which is often used in the literature. In one spatial dimension, it is given by

$$u_t + au_x = 0, \qquad a > 0, \qquad x \in \Omega, \qquad t > 0, \tag{3.1a}$$
$$u(0,t) = g(t), \qquad t \geq 0, \tag{3.1b}$$
$$u(x,0) = f(x), \qquad x \in \Omega, \tag{3.1c}$$

where $f \in L^2(\Omega)$ is the initial data and g is the boundary data. Well-posedness of the problem (3.1a)-(3.1c) can be defined as follows.

Definition 3.1 (Def. 8.4.1 in [16])**.** The problem (3.1a)-(3.1c) with $g = 0$ is well-posed if there is a unique smooth solution that satisfies the stability estimate

$$\|u(\cdot, \mathcal{T})\| \leq Ke^{\alpha \mathcal{T}} \|f(\cdot)\|, \tag{3.2}$$

where K and α are constants that do not depend on f.

Herein, the norm in (3.2) is the usual L^2-norm defined by

$$\|u\|^2_{L^2(\Omega)} = \int_\Omega |u|^2 \, \mathrm{d}x.$$

We focus on stability proofs, i.e., on proving that the solution to the problems satisfy estimates like (3.2). For many problems, existence of solutions are related to these energy estimates. That is, we can define a difference approximation to the continuous problem that satisfies analogous discrete estimates. Then we can choose a smooth interpolant to interpolate the numerical solution that converges to the true solution upon grid refinement (see [16]). Moreover, uniqueness of the solution can also be inferred from the a priori estimates.

Returning to the linear advection equation (3.1a)-(3.1c), we may prove that it is stable by using the energy method. That is, we multiply the equation by u and integrate over Ω. This results in

$$\int_\Omega uu_t + \int_\Omega auu_x \, \mathrm{d}x = 0.$$

Integrating by parts yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(\cdot,t)\|^2_{L^2(\Omega)} + au^2|_0^1 = 0.$$

Using the boundary condition (3.1b) with $g = 0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(\cdot,t)\|^2_{L^2(\Omega)} + au^2(1,t) = 0,$$

or, alternatively

$$\frac{\mathrm{d}}{\mathrm{d}t}\|u(\cdot,t)\|^2_{L^2(\Omega)} \leq 0.$$

Integrating in time finally results in an estimate analogous to (3.2):

$$\|u(\cdot,\mathcal{T})\|^2_{L^2(\Omega)} \leq \|f(\cdot)\|^2_{L^2(\Omega)}.$$

That is, assuming the existence of a unique smooth solution, the problem (3.1a)-(3.1c) is well-posed in the sense of Definition 3.1. We will sometimes refer to the estimates (3.2) as *energy estimates*.

The above theory also applies to the compressible Navier-Stokes equations when they are linearised. The linearised Navier-Stokes equations are obtained by first restating the equations using the primitive variables, (ρ, u, v, p). Thereafter, the primitive variables are decomposed into an exact smooth solution and a small known perturbation, e.g. $\rho = \rho_{\text{ex}} + \rho'$, (see e.g. [16, 21]). This results in a variable-coefficient problem, which is next turned into a constant-coefficient problem by freezing the coefficients. Lastly, the equations are symmetrised using the matrices in [1]. A rigorous derivation of this procedure is given in Paper A, [10]. The linear well-posedness of the compressible Navier-Stokes equations (2.1) have been studied in various papers, see e.g. [30, 35] for wall boundary conditions and e.g. [18, 28, 37] for far-field boundary conditions.

3.2 Entropy theory

In addition to linear well-posedness, entropy estimates are often studied when analysing the compressible Navier-Stokes equations. Generally, entropy estimates can provide some a priori bounds on the solution of the continuous problem (see e.g. [34]). For the compressible Navier-Stokes equations, requiring an entropy estimate to hold, means that we additionally require the solution of the problem to satisfy the second law of thermodynamics in order for it to be considered *physically relevant* (see [40]).

We give here a brief overview of the entropy theory needed for the analysis of the compressible Navier-Stokes equations. We follow the presentation given in [40], and refer to that paper and the references therein for a more comprehensive introduction to the topic.

In both [17, 40], the theory of entropy is introduced by considering a system of hyperbolic conservation laws on the form:

$$\mathsf{u}_t + \mathsf{f}(\mathsf{u})_x = 0. \tag{3.3}$$

The equations (3.3) are said to be endowed with a scalar *entropy-entropy flux pair*, $(U(\mathsf{u}), F(\mathsf{u}))$ if $U(\mathsf{u})$ is strictly convex (i.e., the Hessian is positive definite, $U_{\mathsf{uu}}(\mathsf{u}) \succ 0$), and the relation $U_{\mathsf{u}}(\mathsf{u})\mathsf{f}_{\mathsf{u}}(\mathsf{u}) = F_{\mathsf{u}}(\mathsf{u})$ is satisfied (see [17]). Furthermore, the *entropy variables*, $\mathsf{w} = U_{\mathsf{u}}(\mathsf{u})$ symmetrise (3.3) (see e.g. [27]).

Since (3.3) may not have a unique weak solution, an entropy inequality can be added, posing as an admissibility criterion for *physically relevant* weak solutions (see [40]). The entropy inequality is found by considering the regularised version of (3.3):

$$\mathsf{u}^\varepsilon_t + \mathsf{f}(\mathsf{u}^\varepsilon)_x = \varepsilon(\mathsf{P}\mathsf{u}^\varepsilon_x)_x, \tag{3.4}$$

where P is some admissible viscosity matrix. When Equation (3.4) is multiplied by the entropy variables, $\mathsf{w}^\top$, and we pass to the limit $\varepsilon \to 0^+$, we find that

$$U(\mathsf{u})_t + F(\mathsf{u})_x \leq 0. \tag{3.5}$$

In the case of a smooth solution, the above is satisfied with equality (see [17]).

3.2.1 Entropy estimates for the compressible Navier-Stokes equations

For the compressible Navier-Stokes equations (2.1), affine functions of the specific entropy $\mathcal{S} = \ln(p/\rho^\gamma)$ are admissible entropy functions (see [20]). Usually, $U(\mathsf{u}) = -\rho\mathcal{S}$ is used. For this particular entropy function, the corresponding entropy flux functions (in the 2-D setting) are $F = -m\mathcal{S}$, $G = -n\mathcal{S}$, and the entropy variables take the form

$$\mathsf{w} \coloneqq U_{\mathsf{u}}(\mathsf{u}) = \frac{1}{c_v} \begin{bmatrix} -c_v(\mathcal{S} - \gamma) - \frac{u^2+v^2}{2T} \\ \frac{u}{T} \\ \frac{v}{T} \\ -\frac{1}{T} \end{bmatrix}. \tag{3.6}$$

It is well-known that the Navier-Stokes equations (2.1) augmented with the adiabatic wall boundary conditions (2.2) satisfy an entropy estimate on the form

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_\Omega U(\mathsf{u}) \, \mathrm{d}\Omega \leq 0. \tag{3.7}$$

The derivation can be found in for example [31, 39], but we include it here for the reader's convenience.

Let Ω be an open and bounded polygonal 2-D domain with boundary $\partial\Omega$. By multiplying the equations (2.1) by the entropy variables (3.6) and integrating over the spatial domain, we obtain

$$\int_\Omega \mathsf{w}^\top \mathsf{u}_t \, \mathrm{d}\Omega + \int_\Omega \mathsf{w}^\top \mathsf{f}^{\mathsf{I}}_x + \mathsf{w}^\top \mathsf{g}^{\mathsf{I}}_y \, \mathrm{d}\Omega = \int_\Omega \mathsf{w}^\top \mathsf{f}^{\mathsf{V}}_x + \mathsf{w}^\top \mathsf{g}^{\mathsf{V}}_y \, \mathrm{d}\Omega.$$

Since $\mathsf{w} = U(\mathsf{u})_{\mathsf{u}}$, we have that $\mathsf{w}^\top \mathsf{u}_t = U(\mathsf{u})_{\mathsf{u}} \mathsf{u}_t = U(\mathsf{u})_t$ in the first term above. The integrand of the second term is recognised as the spatial derivatives of the entropy flux functions, while applying integration-by-parts to the third term, yields

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega U(\mathsf{u})\,\mathrm{d}\Omega + \int_\Omega F_x + G_y\,\mathrm{d}\Omega \\
&= \int_{\partial\Omega}\left[\mathsf{w}^\top\mathsf{f}^\mathsf{v}, \mathsf{w}^\top\mathsf{g}^\mathsf{v}\right]\cdot\vec{n}\,\mathrm{d}s - \int_\Omega \mathsf{w}_x^\top\mathsf{f}^\mathsf{v} + \mathsf{w}_y^\top\mathsf{g}^\mathsf{v}\,\mathrm{d}\Omega, \\
\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega U(\mathsf{u})\,\mathrm{d}\Omega + \int_{\partial\Omega}\left[F,\, G\right]\cdot\vec{n}\,\mathrm{d}s \\
&= \int_{\partial\Omega}\left[\mathsf{w}^\top\mathsf{f}^\mathsf{v}, \mathsf{w}^\top\mathsf{g}^\mathsf{v}\right]\cdot\vec{n}\,\mathrm{d}s - \int_\Omega \mathsf{w}_x^\top\mathsf{f}^\mathsf{v} + \mathsf{w}_y^\top\mathsf{g}^\mathsf{v}\,\mathrm{d}\Omega. \qquad (3.8)
\end{aligned}$$

In the above, $\vec{n} = \left[n_x, n_y\right]$ denotes the outward pointing unit normal. The integrands of the boundary terms above read

$$\begin{aligned}
\left[F,\, G\right]\cdot\vec{n} &= \left[-m\mathcal{S}, -n\mathcal{S}\right]\cdot\vec{n}, \\
\left[\mathsf{w}^\top\mathsf{f}^\mathsf{v}, \mathsf{w}^\top\mathsf{g}^\mathsf{v}\right]\cdot\vec{n} &= \left[-\tfrac{\kappa}{c_v T}T_x, -\tfrac{\kappa}{c_v T}T_y\right]\cdot\vec{n}.
\end{aligned} \qquad (3.9)$$

With the adiabatic wall boundary conditions (2.2), they reduce to

$$\begin{aligned}
\left[F,\, G\right]\cdot\vec{n} &= \left[-m\mathcal{S}, -n\mathcal{S}\right]\cdot\vec{n} = 0, \\
\left[\mathsf{w}^\top\mathsf{f}^\mathsf{v}, \mathsf{w}^\top\mathsf{g}^\mathsf{v}\right]\cdot\vec{n} &= \left[-\tfrac{\kappa}{c_v T}T_x, -\tfrac{\kappa}{c_v T}T_y\right]\cdot\vec{n} = 0,
\end{aligned}$$

since $u, v = 0$ and $\frac{\partial T}{\partial\vec{n}} = 0$ on $\partial\Omega$. Thus, the entropy estimate (3.7) is guaranteed as long as the volume term on the right-hand side of (3.8) is negative semi-definite. Assuming positivity of the density and temperature, $\rho, T > 0$, this can be shown by simply writing out the terms.

Characteristic far-field boundary conditions

Characteristic far-field boundary conditions are derived from the linearised version of the compressible Navier-Stokes equations. They lead to a linearly stable problem, but do not necessarily bound the entropy. In order to show this, we follow the paper [37] in this section.

After the Navier-Stokes equations (2.1) are linearised and symmetrised, the energy method can be used to show well-posedness. In the energy analysis, boundary terms like

$$\int_{x=0} w^\top (A_{1w} w - 2\epsilon F_w^V)\, \mathrm{d}y\mathrm{d}z,$$

emerge. (Note that in [37], the equations are considered in 3-D). Above, w denotes the symmetrised variables obtained after the linearisation procedure is performed. We do not specify the form of the matrices A_{1w}, F_w^V, but refer to [37].

The matrix A_{1w} is rotated into a diagonal matrix $X^\top A_{1w} X = \Lambda_1$ holding the well-known eigenvalues $(u, u, u, u+c, u-c)$ (in 3-D), where c represents the speed of sound. Based on the ingoing/outgoing characteristics, the form of the boundary conditions bounding the term above can be specified. The diagonal matrix Λ_1 is decomposed into two matrices, one holding the positive eigenvalues and the other the negative eigenvalues, i.e., $\Lambda_1 = \Lambda_1^+ + \Lambda_1^-$. The boundary conditions that are studied in [37] are given as

$$\alpha A_{1w}^+ w - \epsilon F_w^V = g, \tag{3.10}$$

where $A_{1w}^+ = X\Lambda_1^+ X^\top$ and α is a scalar. We recognise that this is a Robin-type boundary condition since the viscous flux F_w^V includes a spatial derivative.

From this, the easiest way of seeing that the entropy cannot, in general, be bounded by the characteristic far-field boundary conditions is to consider a supersonic outflow boundary. At this boundary the number of positive eigenvalues is zero ($u < -c$) (see e.g. [15, 37]) and the boundary condition (3.10) is reduced to

$$-\epsilon F_w^V = g,$$

i.e., a Neumann type boundary condition. Consider the boundary terms emerging from the inviscid terms in (3.8) in the entropy analysis above:

$$\int_{\partial\Omega} \Big[F,\, G\Big] \cdot \vec{n}\, \mathrm{d}s.$$

We see that these cannot be bounded by a Neumann type boundary condition. The fact that the far-field boundary conditions do not bound the entropy was the motivation behind the work of Paper D. There, a new methodology for imposing far-field boundary conditions in an entropy stable manner for the compressible Navier-Stokes equations is

presented.

4. Semi-discrete theory

In this chapter, we introduce the analogous semi-discrete concepts to linear well-posedness and entropy theory as described in the previous chapter. Additionally, we introduce the summation-by-parts operators used to approximate spatial derivatives.

4.1 Semi-discrete linear stability

Once a problem is proven well-posedness, we can attempt to solve it numerically. In order to be confident that the numerical solution approximates the true solution, some requirements of the chosen numerical scheme must be met. For well-posed linear problems discretised using a consistent and stable difference scheme, the Lax-Richtmyer's equivalence theorem guarantees the convergence of the numerical solution to the true solution (see [24]). For linearly stable difference schemes approximating non-linear problems, a similar result applies if the solutions are smooth (see [33]). As in the previous chapter, we focus on stability proofs for the semi-discrete schemes.

Let $\bar{\Omega}_h$ be a discretisation of $\Omega \cup \partial\Omega = [0,1]$ into a set of $N+1$ grid points x_i, $i = 0, 1, \ldots, N$, and let h denote a typical grid spacing. Bold-face letters denote vector-valued semi-discrete representations of the corresponding continuous variables, e.g., $\boldsymbol{u} = \Big[\boldsymbol{u}_0(t), \boldsymbol{u}_1(t), \ldots, \boldsymbol{u}_N(t)\Big]^\top$, where $\boldsymbol{u}_i \approx u(x_i, t)$.

Consider the general semi-discrete approximation of (3.1a)-(3.1c).

$$\boldsymbol{u}_t + a\mathsf{D}\boldsymbol{u} = 0, \qquad t > 0, \tag{4.1a}$$

$$\boldsymbol{u}_0(t) = g(t), \qquad t \geq 0, \tag{4.1b}$$

$$\boldsymbol{u}(0) = \boldsymbol{f}. \tag{4.1c}$$

Here, D is a consistent approximation of $\frac{\partial}{\partial x}$ on the grid points, x_i, $i = 0, 1, \ldots, N$, and $\boldsymbol{f} = \Big[f(x_0), f(x_1), \ldots, f(x_N)\Big]^\top$. We use the following definition of stability for such a scheme.

Definition 4.1 (Def. 11.3.1 in [16])**.** The approximation (4.1a)-(4.1c) with $g = 0$ is stable if, for $h \leq h_0$, there are constants K, α such that

$$\|\boldsymbol{u}(\mathcal{T})\|_{\mathsf{H}} \leq Ke^{\alpha \mathcal{T}} \|\boldsymbol{f}\|_{\mathsf{H}}. \tag{4.2}$$

Generally, the constants, K and α above differs from the ones in the well-posedness defintion 3.1.

As discussed in Section 3.1, the well-posedness of the linearised compressible Navier-Stokes equations has been analysed with different boundary conditions in various papers. However, the linearisation of the schemes approximating the non-linear version has been studied to a lesser extent.

In Paper A ([10]), the semi-discrete scheme proposed to approximate the 1-D compressible Navier-Stokes equations was linearised, and its linear stability properties were analysed. In that paper, we found that the scheme for the non-linear equations is indeed linearly stable in the sense of Definition 4.1.

4.1.1 Summation-by-parts operators

We employ summation-by-parts (SBP) operators to approximate the spatial derivatives in the problems we study, i.e., the operator D in (4.1a) is assumed to be SBP. SBP operators are designed to discretely mimic integration-by-parts, which is used in the energy method. Thus, SBP operators can be used in the discrete energy method to derive stability estimates analogous to the ones obtained in the continuous setting.

SBP operators were first derived using the finite-difference method in [22] (and we refer to the review paper [8] for a historic description of the subsequent developments). In later years, several different numerical methods have been studied in the SBP context. As a result, also finite-volume methods, discontinuous Galerkin methods and spectral collocation methods have been shown to lead to SBP operators (see e.g. [2, 5, 9, 29]).

To prove stability for the semi-discrete scheme (4.1a)-(4.1c), we introduce a general definition of an SBP operator approximating the first derivative. To this end, define $\boldsymbol{x}^k = \left[x_0^k, x_1^k, \ldots, x_n^k\right]$.

Definition 4.2 (Def 1. in [8])**.** An operator D is a degree p SBP approximation of $\frac{\partial}{\partial x}$ if

1. $\mathsf{D}\boldsymbol{x}^k = k\boldsymbol{x}^{k-1}$ for all $k \in [0, p]$,

2. $\mathsf{D} = \mathsf{H}^{-1}\mathsf{Q}$, where H is symmetric, positive definite, i.e., $\mathsf{H} = \mathsf{H}^\top \succ 0$,

3. $\mathsf{Q} + \mathsf{Q}^\top = \mathsf{E} = \boldsymbol{t}_R\boldsymbol{t}_R^\top - \boldsymbol{t}_L\boldsymbol{t}_L^\top = \mathrm{diag}(-1, 0, \ldots, 0, 1)$, where $\boldsymbol{t}_R^\top = \begin{bmatrix}0, \ldots, 0, 1\end{bmatrix}$ and $\boldsymbol{t}_L^\top = \begin{bmatrix}1, 0, \ldots, 0\end{bmatrix}$.

The matrix, H, in the definition above, defines an L^2-equivalent discrete norm, $\boldsymbol{u}^\top\mathsf{H}\boldsymbol{u} := \|\boldsymbol{u}\|_{\mathsf{H}}^2$.

Together with proper boundary imposition techniques, SBP operators lead to provably stable numerical schemes. There are different ways of imposing boundary conditions stably. Here, we consider the simultaneous approximation term (SAT) and the injection method.

The simultaneous approximation term (SAT)

The simultaneous approximation term (SAT) was introduced by Carpenter, Gottlieb and Abarbanel in [3]. It imposes boundary (or interface) conditions in a weak manner by adding a penalty term to the scheme. The penalty term is designed such that the boundary conditions may not be satisfied exactly. That is, similar to the rest of the problem, the boundary conditions are approximated, but the SAT does not change the overall accuracy of the scheme (see [3]). The SAT technique is extensively used in the SBP community, and SBP-SAT schemes lead to stable approximations for many different problems (see e.g. [2, 3, 6, 7, 8, 36, 31]).

The following SBP-SAT scheme can be used to approximate the advection problem (3.1a)-(3.1c) with $g = 0$ in (3.1b).

$$\boldsymbol{u}_t + a\mathsf{D}\boldsymbol{u} = \mathsf{SAT}, \tag{4.3a}$$

$$\boldsymbol{u}(0) = \boldsymbol{f}, \tag{4.3b}$$

The SAT takes the form $\mathsf{SAT} = -\frac{1}{2}\mathsf{H}^{-1}\boldsymbol{t}_L\boldsymbol{t}_L^\top(a\boldsymbol{u} - 0)$, and $\boldsymbol{f} = \begin{bmatrix}f(x_0), f(x_1), \ldots f(x_n)\end{bmatrix}^\top$. The scheme can be proven stable by using the discrete energy method. That is, we multiply (4.3a) by $\boldsymbol{u}^\top\mathsf{H}$ to obtain

$$\boldsymbol{u}^\top \mathsf{H}\boldsymbol{u} + a\boldsymbol{u}^\top \mathsf{H}\mathsf{D}\boldsymbol{u} = \boldsymbol{u}^\top \mathsf{H}\mathsf{S}\mathsf{A}\mathsf{T}.$$

Recall from Definition 4.2 that $\mathsf{H}\mathsf{D} = \mathsf{Q}$. Thus, the above reads

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 + a\boldsymbol{u}^\top \mathsf{Q}\boldsymbol{u} = \boldsymbol{u}^\top \mathsf{H}\mathsf{S}\mathsf{A}\mathsf{T}.$$

Next, we split the second term into two equal parts and use Condition 3 of Definition 4.2 on one of these to obtain

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 + \frac{1}{2}a\boldsymbol{u}^\top \mathsf{Q}\boldsymbol{u} + \frac{1}{2}a\boldsymbol{u}^\top(\mathsf{E} - \mathsf{Q}^\top)\boldsymbol{u} &= \boldsymbol{u}^\top \mathsf{H}\mathsf{S}\mathsf{A}\mathsf{T},\\ \frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 + a\boldsymbol{u}^\top \mathsf{E}\boldsymbol{u} &= 2\boldsymbol{u}^\top \mathsf{H}\mathsf{S}\mathsf{A}\mathsf{T}.\end{aligned}$$

By inserting $\mathsf{S}\mathsf{A}\mathsf{T} = -\frac{1}{2}\mathsf{H}^{-1}\boldsymbol{t}_L\boldsymbol{t}_L^\top(a\boldsymbol{u} - 0)$, we arrive at

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 + a\boldsymbol{u}_N^2 - a\boldsymbol{u}_0^2 = -a\boldsymbol{u}_0^2,$$

or alternatively

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 \leq 0.$$

Integration in time finally yields a result analogous to the estimate (4.2):

$$\|\boldsymbol{u}(\mathcal{T})\|_{\mathsf{H}}^2 \leq \|\boldsymbol{f}\|_{\mathsf{H}}^2.$$

That is, the scheme (4.3a)-(4.3b) is stable in the sense of Definition 4.1.

The injection method

The injection method for imposing Dirichlet boundary conditions is straightforward to use. One simply sets the boundary nodes to the boundary data. It is therefore often considered a strong imposition technique. Contrary to the SAT, the boundary conditions are satisfied exactly. However, the technique does not necessarily lead to stable schemes (see [14, 32]). One of the contributions of this dissertation is the proof that injected Dirichlet boundary conditions lead to provably stable schemes for different problems using diagonal-norm, diagonal-E SBP operators (see Paper C, [12]). In the following example, we use the approach taken in Paper C.

Returning to the advection problem (3.1a)-(3.1c), a stable approximation using the injection method for imposing the boundary condition (3.1b) with $g = 0$ can be stated as

$$\boldsymbol{u}_t + a\tilde{\mathsf{D}}\boldsymbol{u} = 0, \tag{4.4a}$$

$$\boldsymbol{u}_0 = 0, \tag{4.4b}$$

$$\boldsymbol{u}(0) = \boldsymbol{f}. \tag{4.4c}$$

where $\tilde{\mathsf{D}}$ is any diagonal-norm, diagonal-E SBP operator, D, with the elements of the first row (the row acting on the boundary node) set to zero. Due to the form of $\tilde{\mathsf{D}}$, the equation at the boundary node, x_0, reads

$$(\boldsymbol{u}_0)_t = 0.$$

Thus, there is no equation updating $\boldsymbol{u}_0$, and it therefore remains zero for all time.

The stability of the scheme is again found by using the discrete energy method. That is, we multiply (4.4a) by $\boldsymbol{u}^\top\mathsf{H}$:

$$\boldsymbol{u}^\top\mathsf{H}\boldsymbol{u}_t + a\boldsymbol{u}^\top\mathsf{H}\tilde{\mathsf{D}}\boldsymbol{u} = 0.$$

Since $\boldsymbol{u}(t)_0 = 0$ for all t, and H is assumed to be diagonal, we recognise that $\boldsymbol{u}^\top\mathsf{H}\tilde{\mathsf{D}}\boldsymbol{u} = \boldsymbol{u}^\top\mathsf{H}\mathsf{D}\boldsymbol{u}$. That is, the SBP properties of Definition 4.2 apply, and we obtain (by following the analogous proof for the scheme in the previous subsection)

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 + a\boldsymbol{u}_N^2 - a\boldsymbol{u}_0^2 = 0.$$

Since $\boldsymbol{u}_0 = 0$, the above can be recast as

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{u}\|_{\mathsf{H}}^2 \leq 0,$$

which, after integration in time, results in the estimate

$$\|\boldsymbol{u}(\mathcal{T})\|_{\mathsf{H}}^2 \leq \|\boldsymbol{f}\|_{\mathsf{H}}^2.$$

That is, the scheme (4.4a)-(4.4c) is stable in the sense of Definition 4.1.

A comment on in-homogeneous Dirichlet boundary conditions

The definition of stability, Definition 4.1, is given for $g = 0$. For linear problems, this is no restriction, as we can transform a problem with in-homogeneous boundary data into one with homogeneous boundary data. This is done by introducing a smooth function, φ, that satisfies the in-homogeneous boundary condition. Then $\bar{u} = u - \varphi$ satisfies the original PDE with homogeneous boundary conditions (see [16]).

Thus, for linear problems, both SATs and the injection method can be used to stably approximate problems with in-homogeneous boundary data. For the advection problem considered above, the SBP-SAT scheme can be proven stable by simply restating the SAT as

$$\mathsf{SAT} = \sigma\mathsf{H}^{-1}\boldsymbol{t}_L\boldsymbol{t}_L^\top(a\boldsymbol{u} - g(t)\boldsymbol{t}_L),$$

where σ is some constant that is determined from the stability analysis (see [36]).

For the injection method, one could state the scheme as

$$\begin{aligned} \boldsymbol{u}_t + a\tilde{\mathsf{D}}\boldsymbol{u} &= 0, \\ \boldsymbol{u}_0 &= g(t), \\ \boldsymbol{u}(0) &= \boldsymbol{f}. \end{aligned}$$

Then, analogous to the continuous problem, we introduce a smooth function, φ, satisfying the boundary condition, such that the above scheme is transformed into one with homogeneous boundary data. Thereafter, the stability proof from the previous subsection applies.

4.1.2 Extensions

In the above, we have only considered SBP operators approximating the first derivative in one spatial dimension. Here, we briefly comment on some extensions to higher dimensions and to approximations of higher derivatives.

Higher dimensions

Finitie-difference SBP operators defined on a 1-D grid can be extended to several dimensions by using tensor products (see e.g. [36]). Let N_x, N_y be the number of nodes in the x- and y-direction, respectively, and denote by $S = \{(x_i, y_i)\}_{i=1}^{N}$ the set of all $N = N_x \cdot N_y$ nodes in the grid. Let D_x be a 1-D finite-difference SBP operator approximating $\frac{\partial}{\partial x}$. Then

$$\bar{\mathsf{D}}_x = \mathsf{I}_{N_y} \otimes \mathsf{D}_x,$$

where I_{N_y} denotes the $N_y \times N_y$ identity matrix, approximates $\frac{\partial}{\partial x}$ on the grid points S. Similarly, $\bar{\mathsf{D}}_y = \mathsf{D}_y \otimes \mathsf{I}_{N_x}$ approximates $\frac{\partial}{\partial y}$ on S. We refer the reader to the review papers [8, 36] for more information of the extension of 1-D finite-difference SBP operators to several dimensions.

Finite-difference SBP operators are well suited for structured grids. On the other hand, some SBP operators are defined directly on a 2-D (or 3-D) spatial domain, and may thus be applicable for unstructured grids. These include finite-volume SBP operators (see [5, 29]), the discontinuous Galerkin method (see e.g. [9]) and spectral collocation

methods (see e.g. [2]). In Paper C, the definition of a *multidimensional* SBP operator given in [19] is used to cover different types of operators. We will give the definition here for completeness.

We let now $S = \{(x_i, y_i)\}_{i=0}^{N}$ denote a set of $N+1$ grid points on a 2-D domain Ω, and denote by the bold-face letters $\boldsymbol{p}, \boldsymbol{q}$ vector-valued representations of the two polynomials $\mathcal{P}, \mathcal{Q}$ on S:

$$\boldsymbol{p} = \Big[\mathcal{P}(x_0, y_0), \mathcal{P}(x_1, y_1), \ldots, \mathcal{P}(x_N, y_N)\Big]^\top,$$
$$\boldsymbol{q} = \Big[\mathcal{Q}(x_0, y_0), \mathcal{Q}(x_1, y_1), \ldots, \mathcal{Q}(x_N, y_N)\Big]^\top.$$

Definition 4.3 (Def. 2.1 in [19])**.** Let $\mathcal{P}, \mathcal{Q}$ be two polynomials of degree less than or equal to p, i.e., $\mathcal{P}, \mathcal{Q} \in \mathbb{P}^p(\Omega)$. The matrix D_x is a degree p SBP approximation of $\frac{\partial}{\partial x}$ on the grid points S, if

1. $\mathsf{D}_x \boldsymbol{p} = \frac{\partial}{\partial x}\mathcal{P}$ on S for all $\mathcal{P} \in \mathbb{P}^p(\Omega)$.
2. $\mathsf{D}_x = \mathsf{H}^{-1}\mathsf{Q}_x$, where H is symmetric, positive definite, i.e., $\mathsf{H} = \mathsf{H}^\top \succ 0$.
3. $\mathsf{Q}_x + \mathsf{Q}_x^\top = \mathsf{E}_x$ where E_x satisfies

$$\boldsymbol{p}^\top \mathsf{E}_x \boldsymbol{q} = \int_{\partial\Omega} \mathcal{P}\mathcal{Q} n_x \, \mathrm{d}s,$$

for all $\mathcal{P}, \mathcal{Q} \in \mathbb{P}^\tau(\Omega)$, where $\tau \geq p$, and n_x denotes the x-component of the outward unit normal, $\vec{n} = \Big[n_x, n_y\Big]$.

The definition above covers tensor-product SBP operators as well as SBP operators defined directly on an unstrutured grid, see [19].

Higher derivatives

So far, we have only considered theory suited to approximate first derivatives. However, to discretise the viscous terms of the Navier-Stokes equations (2.1), we also need second-derivative approximations. The general form of SBP operators approximating the second derivative was given in [4], and was further developed in [26], where the following definition is found.

Definition 4.4 (Def. 1 in [26])**.** A difference operator $\mathsf{D}_2 = \mathsf{H}^{-1}(-\mathsf{A}+\mathsf{BS})$ approximating $\frac{\partial^2}{\partial x^2}$ is a second-derivative SBP operator if $\mathsf{A} + \mathsf{A}^\top \succeq 0$, S includes an approximation of the first-derivative operator at the boundary, and $\mathsf{B} = \text{diag}(-1, 0, \ldots, 0, 1)$.

There are essentially two ways of defining a second-derivative SBP operator. One approach is to use a first-derivative SBP approximation twice, i.e.,

$$\mathsf{D}_2 = \mathsf{D}_1\mathsf{D}_1 = \mathsf{H}^{-1}\mathsf{Q}\mathsf{H}^{-1}\mathsf{Q},$$

where D_1 denotes a first-derivative SBP operator. This results in a so-called *complete* second-derivative SBP operator (i.e., $\mathsf{A} = \mathsf{D}_1^\top\mathsf{H}\mathsf{D}_1$ in Definition 4.3) (see [26]). However, this approach results in a second-derivative operator with a wider stencil compared to a *minimal stencil* operator (see [26]). Minimal stencil operators are presented in [4, 26], and the extension to second-derivative approximations with variable-coefficients second derivatives is found in [25].

4.2 Semi-discrete entropy theory

Since the continuous Navier-Stokes equation (2.1) satisfy an entropy estimate, (3.7) (at least with adiabatic wall boundary conditions), we would like the schemes we propose to approximate the equations to satisfy an analogous estimate. Thus, we introduce the semi-discrete entropy theory. Again following [40], we consider the semi-discrete conservative scheme approximating the system of hyperbolic conservation laws (3.3):

$$(\boldsymbol{u}_i)_t + \frac{\boldsymbol{f}_{i+1/2} - \boldsymbol{f}_{i-1/2}}{\Delta x_i} = 0, \tag{4.6}$$

where $\boldsymbol{f}$ is a consistent and Lipschitz continuous numerical flux, and $\Delta x_i = \frac{x_{i+1}-x_{i-1}}{2}$. To obtain a scheme approximating the entropy inequality (3.5), the Equation (4.6) is multiplied by the entropy variables $\boldsymbol{w}_i^\top$.

$$\boldsymbol{w}_i^\top(\boldsymbol{u}_i)_t + \boldsymbol{w}_i^\top\frac{\boldsymbol{f}_{i+1/2} - \boldsymbol{f}_{i-1/2}}{\Delta x_i} = 0.$$

Introducing the approximation of the entropy flux function $\boldsymbol{F}_{i+1/2} = \frac{\boldsymbol{w}_i^\top + \boldsymbol{w}_{i+1}^\top}{2}\boldsymbol{f}_{i+1/2} - \frac{\boldsymbol{\psi}_i + \boldsymbol{\psi}_{i+1}}{2}$ (and correspondingly for $\boldsymbol{F}_{i-1/2}$) where $\boldsymbol{\psi}$ is the semi-discrete representation of the entropy flux potential, $\psi = wf - F$, the above can be rewritten into

$$\begin{aligned}&(\boldsymbol{U}_i)_t + \frac{\boldsymbol{F}_{i+1/2} - \boldsymbol{F}_{i-1/2}}{\Delta x_i}\\ &= \frac{\boldsymbol{w}_{i+1}^\top - \boldsymbol{w}_i^\top}{2\Delta x_i}\boldsymbol{f}_{i+1/2} - \frac{\boldsymbol{\psi}_{i+1} - \boldsymbol{\psi}_i}{2\Delta x_i} + \frac{\boldsymbol{w}_i^\top - \boldsymbol{w}_{i-1}^\top}{2\Delta x_i}\boldsymbol{f}_{i-1/2} - \frac{\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i-1}}{2\Delta x_i}\end{aligned}$$

We note that we arrive at the semi-discrete version of the entropy inequality (3.5) if the right-hand side above is non-positive. That is, we obtain

$$(\boldsymbol{U}_i)_t + \frac{\boldsymbol{F}_{i+1/2} - \boldsymbol{F}_{i-1/2}}{\Delta x_i} \leq 0. \tag{4.7}$$

Depending on whether the inequality is satisfied with equality or not, we distinguish between *entropy stable* and *entropy conservative* schemes.

Theorem 4.5 (Thm. 3.1 in [40])**.** *The scheme* (4.6) *is* entropy stable *if*

$$\frac{\boldsymbol{w}_{i+1}^\top - \boldsymbol{w}_i^\top}{2}\boldsymbol{f}_{i+1/2} - \frac{\boldsymbol{\psi}_{i+1} - \boldsymbol{\psi}_i}{2} \leq 0,$$

and entropy conservative *if*

$$\frac{\boldsymbol{w}_{i+1}^\top - \boldsymbol{w}_i^\top}{2}\boldsymbol{f}_{i+1/2} - \frac{\boldsymbol{\psi}_{i+1} - \boldsymbol{\psi}_i}{2} = 0.$$

In this introduction, we do not present a specific scheme for the compressible Navier-Stokes equations. In the papers that make up the scientific results of this thesis, both entropy conservative and entropy stable schemes have been studied.

5. Summary of papers

5.1 Paper A, [10]

Title: Entropy stability for the compressible Navier-Stokes equations with strong imposition of the no-slip boundary condition

Authors: A. Gjesteland and M. Svärd

Journal: *Journal of Computational Physics*, **470**, 111572, (2022)

DOI: 10.1016/j.jcp.2022.111572

In Paper A ([10]), we analyse the compressible Navier-Stokes equations subject to the adiabatic wall boundary conditions (2.2). First, we propose a semi-discrete scheme to approximate the equations. The scheme is defined using the finite-difference SBP operator that is second-order accurate in the interior, denoted by D in this section. The no-slip velocity condition is imposed strongly using the injection method, while the adiabatic temperature condition is imposed via a SAT. The injected no-slip condition is the focus of the paper.

For the sake of this summary, consider the 1-D Navier-stokes equations on $\Omega = (0, 1)$ where $x = 0$ models a wall.

$$\rho_t + (\rho u)_x = 0, \tag{5.1a}$$

$$m_t + (\rho u^2 + p)_x = (2\mu + \lambda)u_{xx}, \tag{5.1b}$$

$$E_t + (u(E + p))_x = (2\mu + \lambda)(uu_x)_x + \kappa T_{xx}. \tag{5.1c}$$

Furthermore, suppose that the domain $[0, 1]$ is discretised into $N + 1$ equidistant grid points, x_i, $i = 0, \ldots, N$ such that $0 = x_0 < x_1 < \ldots < x_N = 1$. The no-slip condition $u =$ is incorporated into the scheme by introducing the *Dirichlet-SBP* operator corresponding to D:

$$\tilde{D} = \frac{1}{2h} \begin{bmatrix} 0 & 0 & 0 & 0 & \dots & 0 \\ -1 & 0 & 1 & 0 & \dots & 0 \\ 0 & -1 & 0 & 1 & & 0 \\ \vdots & & & \ddots & & \vdots \\ 0 & \dots & 0 & -1 & 0 & 1 \\ 0 & \dots & 0 & 0 & -2 & 2 \end{bmatrix}.$$

The Dirichlet-SBP operator is used in the momentum equation (5.1b) where the no-slip condition applies. Due to the form of $\tilde{D}$, the equation for the node $i = 0$ reads

$$(m_0)_t = 0.$$

Thus, there is no scheme that updates m_0, and it remains zero for all time by setting it to zero initially. However, defining $\tilde{D}$ this way, ruins the structure of the original SBP operator at the left boundary. This results in a boundary operator, $\tilde{B}$ that is not diagonal in the upper left corner. (Note that the boundary operator B in this section is the same as E in Definition 4.2, but the former notation is used in Paper A).

The main contributions of Paper A is a proof that the semi-discrete scheme using $\tilde{D}$ for imposing the no-slip condition is entropy stable if the interior scheme is assumed entropy stable. Furthermore, a rigorous linearisation of the scheme approximating the 1-D equations with a local Lax-Friedrichs type artificial dissipation is performed. It is shown that the entropy stable scheme approximating the fully non-linear equations is also linearly stable in the sense of Definition 4.1. The linear well-posedness of the continuous Navier-Stokes equations is often studied, but the present proof also demonstrate that the scheme proposed to approximate the fully non-linear equations is linearly stable. A linearly stable third-order scheme is also proposed.

5.2 Paper B, [11]

Title: Convergence of Chandrashekar's Second-Derivative Finite-Volume Approximation

Authors: A. Gjesteland and M. Svärd

Journal: *Journal of Scientific Computing*, **96:46**, (2023)

DOI: 10.1007/s10915-023-02256-9

In Paper A, the spatial domain was discretised into a structured grid for which finite-difference methods are suitable. For domains with complex geometries, however, we may wish to define an unstructured grid. The finite-volume method is often used for this purpose.

In Paper B ([11]), we study the consistency of a finite-volume method approximating the second derivative. The motivation behind the work is the discretisation of the viscous terms of the compressible Navier-Stokes equations on unstructured grids. Although the finite-volume methods are robust, easy to derive and well suited for unstructured grids, the consistency of second-derivative approximations does not necessarily follow (see e.g. [38]).

The second-derivative finite-volume approximation that is considered in Paper B was derived by Chandrashekar in [5]. It is a local approximation in the sense that it only uses the nearest neighbouring nodes of a grid point i to approximate the second-derivative at this point. The approximation included Dirichlet boundary conditions that were imposed in a weak sense. The resulting operator was shown to satisfy the SBP property in [5]. In Paper B, this operator is slightly altered so as to not include any boundary conditions. The Laplacian approximation in [5] was built by first approximating the gradient on a triangle K_n by

$$\nabla_h u^n = \frac{1}{2|K_n|} \left(u_i \hat{\boldsymbol{n}}_i^n + u_j \hat{\boldsymbol{n}}_j^n + u_k \hat{\boldsymbol{n}}_k^n \right). \tag{5.2}$$

In the above, i, j, k denotes the three vertices of triangle K_n while $\hat{\boldsymbol{n}}_{i,j,k}^n$ represents the outward pointing normals on the edge opposite of node i, j, k, respectively. Finally, $|K_n|$ is the area of the triangle K_n. See Figure 5.1 for a visualisation of all the components.

Next, the Laplacian operator is approximated as (see [5])

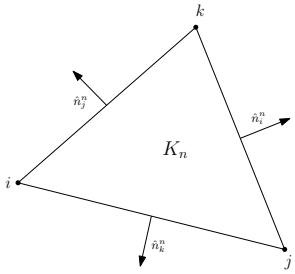


Figure 5.1: Figure depicting the elements of the gradient approximation (5.2). (Figure from [11].)

$$(\Delta_h u)_i = \frac{1}{2V_i}\left(\sum_{n\in N_i} \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \sum_{e\in E_i} \nabla_h u^{n(e)} \cdot \hat{\boldsymbol{b}}(e)\right). \tag{5.3}$$

Here, V_i denotes the dual volume of grid point i, N_i the set of triangles having grid point i as a vertex, E_i the set of boundary edges having vertex i as an endpoint and $\hat{\boldsymbol{b}}(e)$ the outward normal on a boundary edge e. The components are depicted in Figure 5.2.

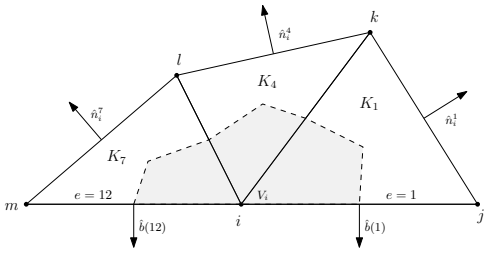


Figure 5.2: Figure depicting the components of the second-derivative approximation (5.3). (Figure from [11].)

Following the proof of [5], the altered operator is shown to be SBP in Paper B. Using the 2-D heat equation as a model problem, a priori estimates for the numerical solution corresponding to the ones obtained for the continuous solution, is found utilising the SBP property. From the a priori bounds, the solution is shown to converge weakly to a weak solution of the original problem. Subsequently, strong convergence to the weak solution is established by employing Aubin-Lions lemma. Thus, the finite-volume approximation is proven consistent (in a weak sense) on triangulated unstructured grids.

5.3 Paper C, [12]

Title: Injected Dirichlet boundary conditions for general diagonal-norm SBP operators

Authors: A. Gjesteland, D. Del Rey Fernández and M. Svärd

Preprint: 10.13140/RG.2.2.26067.55843, *Under review*, (2023)

In Paper C, the definition of the Dirichlet-SBP operator introduced in Paper A is extended to general diagonal-norm SBP operators with diagonal boundary operators, E. That is, we use the definition of multidimensional SBP operators (see Definition 4.3) given in [19] to cover a general class of operators that include tensor-product finite-difference SBP operators but also operators defined directly on unstructured grids.

The purpose of the Dirichlet-SBP operators is to study the stability properties of SBP schemes with injected Dirichlet boundary conditions. In Section 4.1 an example for the linear advection equation is provided. To further introduce the results of Paper C, consider the 1-D linear advection-diffusion equation.

$$\begin{aligned} u_t + au_x &= \varepsilon u_{xx}, \\ u(x,0) &= f(x). \end{aligned}$$

We consider the problem on $\Omega = (0,1)$, and neglect the right boundary. On the left boundary, we impose the homogeneous boundary condition $u(0,t) = 0$, and we let the initial function, f be bounded in L^2.

Remark 5.1. In Paper C, the results where shown for *multi*dimensional SBP operators. Herein, we consider the 1-D problem to reduce notation. However, the use of multidimensional SBP operators is not restrictive, but rather a generalisation. The important assumptions are that the norm matrices, H, and the boundary matrices, E, are diagonal.

The following semi-discrete scheme is a stable approximation of the above problem.

$$\begin{aligned} \boldsymbol{u}_t + a\tilde{\mathsf{D}}_x\boldsymbol{u} &= \varepsilon\tilde{\mathsf{D}}_{xx}\boldsymbol{u}, \\ \boldsymbol{u}(0) &= \boldsymbol{f}. \end{aligned}$$

In the above, $\tilde{\mathsf{D}}$ approximates the first derivative, and is any diagonal-norm, diagonal-E SBP operator with the elements of the first row set to zero. Analogously, $\tilde{\mathsf{D}}_{xx}$ is an approximation of the second derivative on the form given in Definition 4.4. For simplicity, we assume that the original second-derivative SBP operator that $\tilde{\mathsf{D}}_{xx}$ is defined from is *complete*, i.e., $\mathsf{D}_{xx} = \mathsf{D}_x\mathsf{D}_x$ (see Section 4.1).

To prove that the scheme is stable, we multiply by $\boldsymbol{u}^\top\mathsf{H}$. This results in

$$\boldsymbol{u}^\top\mathsf{H}\boldsymbol{u}_t + a\boldsymbol{u}^\top\mathsf{H}\tilde{\mathsf{D}}_x\boldsymbol{u} = \varepsilon\boldsymbol{u}^\top\mathsf{H}\tilde{\mathsf{D}}_{xx}\boldsymbol{u}.$$

Since $u_0 = 0$, and H is diagonal, the following relations hold, $\boldsymbol{u}^\top\mathsf{H}\tilde{\mathsf{D}} = \boldsymbol{u}^\top\mathsf{H}\mathsf{D}$ and $\boldsymbol{u}^\top\mathsf{H}\tilde{\mathsf{D}}_{xx} = \boldsymbol{u}^\top\mathsf{H}\mathsf{D}_{xx}$. From here, the stability proof is the same as for any SBP scheme, but the resulting boundary terms at $x = 0$ vanish due to $u_0 = 0$.

In Paper C, the above methodology was used to show that semi-discrete schemes defined using multidimensional SBP operators for the advection and advection-diffusion equations with both homogeneous and in-homogeneous Dirichlet boundary conditions are stable. Furthermore, by using the same methodology for imposing the no-slip boundary condition for the compressible Navier-Stokes equations, we prove that the resulting scheme satisfies an entropy estimate. Contrary to Paper A, there are no assumptions on the accuracy of the SBP operators used in Paper C.

5.4 Paper D

Title: Entropy stable far-field boundary conditions for the compressible Navier-Stokes equations

Authors: M. Svärd and A. Gjesteland, *Submitted*, (2023)

The compressible Navier-Stokes equations are commonly used for simulating external flows, for example air flow past an airplane wing. In this particular case, the wing is a solid body on which the adiabatic wall boundary conditions (2.2) applies. Sufficiently far away from the wing, we may assume that the flow is unaffected by the wing. However, in practical simulations, the spatial domain cannot be arbitrarily large, and it is often truncated. This introduces "artificial boundaries" for which we must give appropriate boundary conditions. As described in Section 3.2.1, characteristic far-field boundary conditions are often used for the Navier-Stokes equations. These are derived from the linearised equations to yield a linearly well-posed problem. However, as discussed in

Section 3.2.1, they do not necessarily lead to entropy estimates for the Navier-Stokes equations.

This is the motivation behind the work in Paper D. Here, we present a new methodology for setting far-field boundary conditions for the compressible Navier-Stokes equations that are entropy stable. To introduce the concept, consider the 1-D version of the Navier-Stokes equations on the infinitely large domain $\Omega = (0, \infty)$ where $x = 0$ models a wall and $x \to \infty$ models a far-field:

$$\begin{aligned} \rho_t + (\rho u)_x &= 0, \\ m_t + (\rho u^2 + p)_x &= (2\mu + \lambda) u_{xx}, \\ E_t + (u(E+p))_x &= (2\mu + \lambda)(u u_x)_x + \kappa T_{xx}. \end{aligned}$$

The above equations are augmented by the boundary conditions

$$\text{wall:} \qquad u = 0, \qquad T_x = 0, \tag{5.4}$$

$$\text{far-field:} \qquad \rho = \rho_\infty, \qquad u = u_\infty, \qquad T = T_\infty, \qquad T_x|_{x\to\infty} = 0, \tag{5.5}$$

where $\rho_\infty, u_\infty, T_\infty$ denote the constant free-stream values of the respective variables as $x \to \infty$.

Since affine functions of the specific entropy, $\mathcal{S} = \ln(p/\rho^\gamma)$ are admissible entropy functions for the Navier-Stokes equations (see [20]), we can introduce the renormalised entropy function and entropy flux function:

$$\tilde{U}(\mathsf{u}) = -\rho(\mathcal{S} - \mathcal{S}_\infty), \qquad \tilde{F}(\mathsf{u}) = -m(\mathcal{S} - \mathcal{S}_\infty), \tag{5.6}$$

where $\mathsf{u} = \left[\rho, m, E\right]^\top$, and $\mathcal{S}_\infty = \lim_{x\to\infty} S(x, 0)$ is a constant. The corresponding entropy variables are $\tilde{\mathsf{w}} = \mathsf{w} + \left[\mathcal{S}_\infty, 0, 0, 0\right]^\top$, where w denotes the entropy variables obtained from the commonly used entropy function, $U(\mathsf{u}) = -\rho \ln(p/\rho^\gamma)$.

The entropy estimate is found as in Section 3.2.1. With the new entropy variables, $\tilde{\mathsf{w}}$, the boundary terms corresponding to (3.9) that emerges read

$$\tilde{F}(\mathsf{u})|_0^\infty = \lim_{x\to\infty} -m(x,t)(\mathcal{S}(x,t)-\mathcal{S}_\infty) + m(0,t)(\mathcal{S}(0,t)-\mathcal{S}_\infty) = 0,$$
$$\tilde{\mathsf{w}}^\top \mathsf{f}^{\mathsf{v}}(\mathsf{u},\mathsf{u}_x)|_0^\infty = -\lim_{x\to\infty} \frac{\kappa}{c_v T(x,t)} T_x(x,t) + \frac{\kappa}{c_v T(0,t)} T_x(0,t) = 0.$$

They vanish due to the wall boundary conditions (5.4) and the fact that $\mathcal{S}(x,t) = \mathcal{S}_\infty$ and $T_x(x,t) = 0$ as $x \to \infty$.

In order to derive an entropy estimate in the semi-discrete setting, the infinitely large domain $\Omega = (0,\infty)$ is transformed to a finite one, where grid points, $i = 0,\dots,N$ are placed. Thereafter, the finitely large domain is transformed back to $\Omega = (0,\infty)$. Using a finite-volume method, the grid points in the physical domain is distributed such that the dual volume for the very last node, $V_N \to \infty$. By assuming that the resulting ODE system can be solved in time, the division by this infinitely large dual volume results in a scheme for the node $i = N$ that reads

$$(\mathbf{u}_i)_t = 0.$$

Thus, we use $\mathbf{S}_N = S_\infty$, and similarly for the other variables, to obtain estimate analogous to the one we derive for the continuous problem.

We note that this resembles the injection method, but the arguments that lead to the stability estimates differ. In Paper D, the far-field conditions (5.5) are set strongly in the computations.

Bibliography

[1] S. Abarbanel and D. Gottlieb. Optimal Time Splitting for Two- and Three-Dimensional Navier-Stokes Equations with Mixed Derivatives. *Journal of Computational Physics*, 41:1–33, 1981.

[2] M. H. Carpenter and D. Gottlieb. Spectral methods on arbitrary grids. *Journal of Computational Physics*, 129:74–86, 1996.

[3] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. Time-Stable Boundary Conditions for Finite-Difference Schemes Solving Hyperbolic Systems: Methodology and Application to High-Order Compact Schemes. *Journal of Computational Physics*, 111:220–236, 1994.

[4] M. H. Carpenter, J. Nordström, and D. Gottlieb. A Stable and Conservative Interface Treatment of Arbitrary Spatial Accuracy. *Journal of Computational Physics*, 148:341–365, 1999.

[5] P. Chandrashekar. Finite volume discretization of heat equation and compressible Navier-Stokes equations with weak Dirichlet boundary condition on triangular grids. *International Journal of Advances in Engineering Sciences and Applied Mathematics*, 8(3):174–193, 2016. doi: 10.1007/s12572-015-0160-z.

[6] J. Crean, J. E. Hicken, D. C. Del Rey Fernández, D. W. Zingg, and M. H. Carpenter. Entropy-stable summation-by-parts discretizations of the Euler equations on general curved elements. *Journal of Computational Physics*, 356:410–438, 2018. doi: 10.1016/j.jcp.2017.12.015.

[7] L. Dalcin, D. Rojas, S. Zampini, D. C. Del Rey Fernández, M. H. Carpenter, and M. Parsani. Conservative and entropy stable wall boundary conditions for the compressible Navier-Stokes equations: Adiabatic wall and heat entropy transfer. *Journal of Computational Physics*, 397:108775, 2019. doi: https://doi.org/10.1016/j.jcp.2019.06.051.

[8] D. C. Del Rey Fernández, J. E. Hicken, and D. W. Zingg. Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution

of partial differential equations. *Computers & Fluids*, 95:171–196, 2014. doi: 10.1016/j.compfluid.2014.02.016.

[9] G. J. Gassner. A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods. *SIAM Journal on Scientific Computing*, 35(3):A1233–A1253, 2013. doi: 10.1137/120890144.

[10] A. Gjesteland and M. Svärd. Entropy stability for the compressible Navier-Stokes equations with strong imposition of the no-slip boundary condition. *Journal of Computational Physics*, 470:111572, 2022. doi: 10.1016/j.jcp.2022.111572.

[11] A. Gjesteland and M. Svärd. Convergence of Chandrashekar's Second-Derivative Finite-Volume Approximation. *Journal of Scientific Computing*, 96(46), 2023. doi: 10.1007/s10915-023-02256-9.

[12] A. Gjesteland, D. Del Rey Fernández, and M. Svärd. Injected Dirichlet boundary conditions for general diagonal-norm SBP operators. DOI: 10.13140/RG.2.2.26067.55843, 2023.

[13] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong Stability-Preserving High-Order Time Discretization Methods. *SIAM Review*, 43(1):89–112, 2001.

[14] B. Gustafsson. *High Order Difference Methods for Time Dependent PDE*. Springer Series in Computational Mathematics. Springer, Berlin Heidelberg, 2008. ISBN 978-3-540-74993-6. doi: 10.1007/978-3-540-74993-6.

[15] B. Gustafsson and A. Sundström. Incompletely parabolic problems in fluid dynamics. *SIAM Journal on Applied Mathematics*, 35(2), 1978.

[16] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time-dependent problems and difference methods*. Wiley, New Jersey, 2 edition, 2013. ISBN 978-0-470-90056-7.

[17] A. Harten. On the Symmetric Form of Systems of Conservation Laws with Entropy. *Journal of Computational Physics*, 49:151–164, 1983.

[18] J. S. Hesthaven and D. Gottlieb. A stable penalty method for the compressible Navier-Stokes equations: I. open boundary conditions. *SIAM Journal on Scientific Computing*, 17(3):579–612, 1996.

[19] J. E. Hicken, D. C. Del Rey Fernández, and D. W. Zingg. Multidimensional summation-by-parts operators: general theory and application to simplex elements. *SIAM Journal on Scientific Computing*, 38(4):A1935–A1858, 2016. doi: 10.1137/15M1038360.

[20] T. J. R. Hughes, L. P. Franca, and M. Mallet. A new finite element formulation for computational fluid dynamics: I. symmetric form of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics. *Computer Methods in Applied Mechanics and Engineering*, 54:223–234, 1986.

[21] H.-O. Kreiss and J. Lorenz. *Initial-Boundary Value Problems and the Navier-Stokes Equations*. Classics in Applied Mathematics. SIAM, 1989. ISBN 0-89871-565-2.

[22] H.-O. Kreiss and G. Scherer. Finite Element and Finite Difference Methods for Hyperbolic Partial Differential Equations. In *Mathematical Aspects of Finite Elements in Partial Differential Equations*, pages 195–212. Academic Press, 1974.

[23] H. O. Kreiss and L. Wu. On the stability definition of difference approximations for the initial boundary value problem. *Applied Numerical Mathematics*, 12:213–227, 1993.

[24] P. D. Lax and R. D. Richtmyer. Survey of the Stability of Linear Finite Difference Equations. *Communications on Pure and Applied Mathematics*, 9:267–293, 1956.

[25] K. Mattsson. Summation by Parts Operators for Finite Difference Approximations of Second-Derivatives with Variable Coefficients. *Journal of Scientific Computing*, 51:650–682, 2012. doi: 10.1007/s10915-011-9525-z.

[26] K. Mattsson and J. Nordström. Summation by parts operators for finite difference approximations of second derivatives. *Journal of Computational Physics*, 199:503–540, 2004. doi: 10.1016/j.jcp.2004.03.001.

[27] M. S. Mock. System of Conservation Laws of Mixed Type. *Journal of Differential Equations*, 37:70–88, 1980.

[28] J. Nordstöm. The use of characteristic boundary conditions for the Navier-Stokes equations. *Computers & Fluids*, 24(5):609–623, 1995.

[29] J. Nordström, K. Forsberg, C. Adamsson, and P. Eliasson. Finite volume methods, unstructured meshes and strict stability for hyperbolic problems. *Applied Numerical Mathematics*, 45:453–473, 2003. doi: 10.1016/S0168-9274(02)00239-8.

[30] J. Nordström, S. Eriksson, and P. Eliasson. Weak and strong boundary procedures and convergence to steady-state of the Navier-Stokes equations. *Journal of Computational Physics*, 231:4867–4884, 2012. doi: 10.1016/j.jcp.2012.04.007.

[31] M. Parsani, M. H. Carpenter, and E. J. Nielsen. Entropy stable wall boundary conditions for the three-dimensional compressible Navier-Stokes equations. *Journal of Computational Physics*, 292:88–113, 2015. doi: 10.1016/j.jcp.2015.03.026.

[32] B. Strand. Numerical studies of hyperbolic IBVP with high-order finite difference operators satisfying a summation by parts rule. *Applied Numerical Mathematics*, 26:497–521, 1998.

[33] G. Strang. Accurate Partial Difference Methods II. Non-Linear Problems. *Numerische Mathematik*, 6:37–46, 1964.

[34] M. Svärd. Analysis of an alternative Navier-Stokes system: Weak entropy solutions and a convergent numerical scheme. *Mathematical odels and Methods in Applied Sciences*, 32(13):1601–2671, 2022.

[35] M. Svärd and J. Nordström. A stable high-order finite difference scheme for the compressible Navier-Stokes equations: no-slip wall boundary conditions. *Journal of Computational Physics*, 227:4805–4824, 2008. doi: 10.1016/j.jcp.2007.12.028.

[36] M. Svärd and J. Nordström. Review of summation-by-parts schemes for initial-boundary-value problems. *Journal of Computational Physics*, 268:17–38, 2014. doi: 10.1016/j.jcp.2014.02.031.

[37] M. Svärd, M. H. Carpenter, and J. Nordström. A stable high-order finite difference scheme for the compressible Navier-Stokes equations, far-field boundary conditions. *Journal of Computational Physics*, 225:1020–1038, 2007. doi: 10.1016/j.jcp.2007.01.023.

[38] M. Svärd, J. Gong, and J. Nordström. An accuracy evaluation of unstructured node-centred finite volume methods. *Applied Numerical Mathematics*, 58:1142–1158, 2008. doi: 10.1016/j.apnum.2007.05.002.

[39] M. Svärd, M. H. Carpenter, and M. Parsani. Entropy stability and the no-slip wall boundary condition. *SIAM Journal on Numerical Analysis*, 56(1):256–273, 2018. doi: 10.1137/16M1097225.

[40] E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, pages 451–512, 2003. doi: 10.1017/S0962492902000156.

Part II

Scientific results

Paper A

Entropy stability for the compressible Navier-Stokes equations with strong imposition of the no-slip boundary condition

A. Gjesteland and M. Svärd

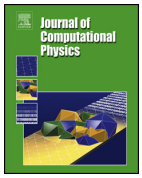
Journal of Computational Physics, **470**, 111572, (2022).

Journal of Computational Physics 470 (2022) 111572

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp



Entropy stability for the compressible Navier-Stokes equations with strong imposition of the no-slip boundary condition

Anita Gjesteland *, Magnus Svärd

Department of Mathematics, University of Bergen, Postbox 7800, 5020 Bergen, Norway

ARTICLE INFO

Article history:
Received 14 October 2021
Received in revised form 19 August 2022
Accepted 23 August 2022
Available online 29 August 2022

Keywords:
No-slip boundary condition
Injection method
Entropy stability
Linear stability

ABSTRACT

We consider the compressible Navier-Stokes equations subject to no-slip adiabatic wall boundary conditions. The main goal is to investigate stability properties of schemes imposing the no-slip condition strongly (injection) and the temperature condition weakly by a simultaneous approximation term. To this end, we propose a low-order summation-by-parts scheme. By verifying the complete linearisation procedure, we prove linear stability for the scheme. In addition, and assuming that the interior scheme is entropy stable, we also prove entropy stability for the full scheme including the boundary treatment. Furthermore, we propose a linearly stable 3rd-order scheme with the same imposition of the wall conditions. However, the 3rd-order scheme is not provably non-linearly stable. A number of simulations show that the boundary procedure is robust for both schemes.

1. Introduction

The compressible Navier-Stokes equations describe the motion of a compressible, viscous and heat conducting fluid. Together with appropriate boundary and initial conditions, they model e.g. aerodynamic problems. Here, we consider the case where the fluid is interacting with solid walls. At walls, the equations are augmented with the no-slip condition leading to the formation of boundary layers that may become unstable and even generate turbulence. These complex phenomena are often studied using computational fluid dynamics. To reliably obtain accurate numerical approximations, the problem must be well-posed and its discrete approximation scheme stable. Unfortunately, well-posedness is, by and large, unknown for the Navier-Stokes equations. However, for smooth solutions, [24] ensures that numerical solutions produced by linearly stable schemes converge.

Linear theory is well developed and one can readily employ the energy method to prove well-posedness of initial-boundary-value problems (IBVP) (see e.g. [8]). Since the continuous energy method relies heavily on the integration-by-parts rule, spatial operators that satisfy the corresponding discrete property, summation-by-parts (SBP), have been developed (see e.g. [14], [29], [5]). These are used to prove energy stability and convergence of linear schemes ([9]). The linear theory has successfully been used to design schemes appropriate for subsonic smooth flows.

In the non-linear regime, however, the linear theory is not sufficient to guarantee stability, let alone well-posedness. To obtain non-linear bounds on the solution, the second law of thermodynamics, stating that the entropy within a closed system cannot decrease, can be used. In mathematical terms, this takes the form of an additional inequality and solutions that satisfy this inequality are termed entropy solutions (see Harten [10] and Tadmor [32] for the Cauchy problem and

* Corresponding author.
E-mail addresses: anita.gjesteland@uib.no (A. Gjesteland), magnus.svard@uib.no (M. Svärd).

https://doi.org/10.1016/j.jcp.2022.111572

[23,27] for boundary treatments). Analogously, a numerical scheme is termed *entropy stable* if it satisfies a discrete equivalent of the continuous entropy inequality.

For both linear and non-linear problems, special attention must be paid to the boundaries, to ensure stability of the numerical scheme. SBP operators, together with the simultaneous approximation terms (SAT) to weakly impose boundary conditions, are applicable to a large class of problems, and are frequently used in the literature see [2,30,29,5,4,23,31]). In contrast to SAT, the injection method, which is the topic of this article, imposes the boundary conditions strongly. In practice, it does so by overwriting the boundary nodes with the boundary data after each time step (and/or Runge-Kutta stage). The injection method is appealing due to its simple nature, but may lead to unstable schemes (see e.g. [8,21]).

Here, we study SBP finite difference discretisations of the compressible Navier-Stokes equations augmented with the no-slip, i.e., homogeneous Dirichlet, boundary condition for the velocities and a homogeneous Neumann condition for the temperature. The no-slip condition is implemented strongly using the injection method, while the Neumann condition is implemented weakly with the SAT technique. Our primary objective is to demonstrate that this boundary procedure is entropy stable. Furthermore, this combination of boundary procedures has previously been considered in [22], where a stability proof for the symmetrised, constant-coefficient Navier-Stokes equations in two spatial dimensions was given. Our secondary objective is to investigate the nature of such linear stability proofs. Hence, we study the complete chain of arguments, from the linearisation of the full non-linear approximation scheme to a variable-coefficient problem and on to a symmetrisable frozen-coefficient problem. In particular, we focus on the validity of the last step.

The remaining article is organised as follows. First, we introduce linear well-posedness and stability, before we introduce the SBP operators and provide an example of the injection technique. Next, we review the linear well-posedness theory for the Navier-Stokes system. Thereafter, we prove stability for a scheme approximating the symmetric constant-coefficient version of the Navier-Stokes equations. (This is what is commonly referred to as linear stability analysis.) Next, we introduce the numerical scheme approximating the non-linear equations, and analyse its linear stability. In particular, we relate it to the constant-coefficient scheme. Next, we prove entropy stability of the scheme in one and two spatial dimensions. Lastly, we provide some numerical simulations that substantiate the findings of our stability proofs.

2. Preliminaries for the linear analysis

A general variable-coefficient initial-boundary-value problem (IBVP) can be written as

$$\begin{aligned} \frac{\partial u}{\partial t} &= P(\partial_x, x, t)u + F(x,t), \quad 0 < x < 1, \ \ t \geq 0, \\ Lu &= g(t), \\ u(x,0) &= f(x), \end{aligned} \tag{1}$$

where P is a spatial differential operator; F is a forcing function and L is an operator acting on the boundary. We will also need the standard L^2-norm defined by $\|u\|^2 = \int_0^1 |u|^2 \, dx$.

Definition 2.1 *(Well-posedness, [8]).* The initial-boundary-value problem (1) is **well-posed** if for $F = g = 0$ there exists a unique solution satisfying

$$\|u(\cdot,t)\| \leq K e^{\alpha t} \|f(\cdot)\|,$$

where K and α are constants independent of $f(x)$. ⌟

Next, define a computational grid with $N+1$ equidistant grid points on the domain $0 \leq x \leq 1$: $x_i = ih$, $h > 0$. Let $\boldsymbol{u}$, $\boldsymbol{f}$, $\boldsymbol{F}$ and $\boldsymbol{g}$ be grid functions corresponding to the continuous functions u, f, F and g, respectively. That is, $[\boldsymbol{u}(t)]_i$ is the approximation of $u(x_i,t)$ etc. Let

$$\begin{aligned} \frac{d\boldsymbol{u}}{dt} &= \mathcal{D}_h \boldsymbol{u} + \boldsymbol{F}, \\ \mathcal{B}\boldsymbol{u} &= \boldsymbol{g}(t), \\ \boldsymbol{u}(0) &= \boldsymbol{f}, \end{aligned} \tag{2}$$

be a semi-discrete approximation of the IBVP (1). $\mathcal{D}_h$ is an approximation of the differential operator, P, and $\mathcal{B}$ an approximation of the boundary operator, L. For the semi-discrete schemes, we use the discrete analog to the L^2-norm defined by $\|\boldsymbol{u}\|_H^2 = \boldsymbol{u}^T H \boldsymbol{u}$, where H is a symmetric positive-definite matrix with elements of size $\mathcal{O}(h)$. Herein, we only consider diagonal H matrices.

Definition 2.2 *(Stability, [8]).* The problem (2) is **stable** if for $\boldsymbol{F} = \boldsymbol{g} = 0$, the solution satisfies

$$\|\boldsymbol{u}(t)\|_H \leq K e^{\alpha t} \|\boldsymbol{f}\|_H,$$

where K and α are constants independent of $\boldsymbol{f}$ and h. ⌟

Stability of the semi-discrete scheme implies stability of the fully discrete scheme if the spatial scheme is advanced in time with an appropriate Runge-Kutta method (see [15] for a proof).

Remark. For many problems, stability in the sense of Definition 2.2 for the variable-coefficient problem follows from stability of the "frozen-coefficient" problem. This was stated as a Conjecture in [9] (page 82).

3. Spatial discretisation

An SBP operator approximating the first derivative takes the form $D = H^{-1}Q$, where the matrices have the following properties:

i) H is a symmetric positive-definite matrix with elements of $\mathcal{O}(h)$,
ii) Q is an *almost* skew-symmetric matrix, satisfying the relation $Q + Q^T = B = \text{diag}(-1, 0, \dots, 0, 1)$.

(For an introduction to SBP operators, see for example the review papers [5,29].)

For concreteness, we use the (2,1)-SBP operator that is second-order accurate in the interior and first-order accurate on the boundary. The operator takes the form

$$D = \frac{1}{2h}\begin{pmatrix} -2 & 2 & 0 & \dots & \\ -1 & 0 & 1 & \dots & \\ & & \ddots & & \\ & & -1 & 0 & 1 \\ & & 0 & -2 & 2 \end{pmatrix}, \quad Q = \frac{1}{2}\begin{pmatrix} -1 & 1 & 0 & \dots & \\ -1 & 0 & 1 & \dots & \\ & & \ddots & & \\ & & -1 & 0 & 1 \\ & & 0 & -1 & 1 \end{pmatrix}, \tag{3}$$

and $H = h \cdot \text{diag}(1/2, 1, \dots, 1, 1/2)$ (this operator can be found in e.g. [16]).

Example 3.1. To introduce the injection technique we consider the system of equations

$$\begin{aligned} u_t + v_x &= 0, \\ v_t + u_x - 2v_x &= 0, \end{aligned} \tag{4}$$

with the boundary condition $v = 0$ at $x = 1$ (neglecting the left boundary for simplicity), and the semi-discretisation

$$\boldsymbol{u}_t + D\boldsymbol{v} = 0, \tag{5}$$

$$\boldsymbol{v}_t + D\boldsymbol{u} - 2D\boldsymbol{v} = 0, \tag{6}$$

where $\boldsymbol{u}$ and $\boldsymbol{v}$ are the numerical solution vectors.

In the injection method, $v(1,t) = 0$ is enforced by $v_N = 0$. A common approach to enforce injection is to remove the equation for v_N from the scheme by removing the boundary element of the solution vector and the last row and column of the spatial differential operator, D (see e.g. [8]). However, for coupled systems such as (4) this may inadvertently introduce extra boundary conditions. Here, we enforce injection indirectly by approximating $(v_N)_t = 0$. To achieve this, we introduce a new operator by setting all elements in the last row of D in (3) to zero, i.e.,

$$\tilde{D} = \frac{1}{2h}\begin{pmatrix} -2 & 2 & 0 & \dots & 0 \\ -1 & 0 & 1 & \dots & 0 \\ & & \ddots & & \\ & & -1 & 0 & 1 \\ & & 0 & 0 & 0 \end{pmatrix}.$$

We term $\tilde{D}$ a *Dirichlet-SBP* operator. The Dirichlet-SBP operator satisfies a new SBP-type property replacing ii):

$$\tilde{Q} + \tilde{Q}^T = \tilde{B} = \begin{pmatrix} -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ & & & & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \dots & \frac{1}{2} & 0 \end{pmatrix}.$$

We alter the scheme (5)-(6), to take the boundary condition $v(1,t)=0$ into account, as follows,

$$\boldsymbol{u}_t + D\boldsymbol{v} = 0, \tag{7}$$

$$\boldsymbol{v}_t + \tilde{D}\boldsymbol{u} - 2\tilde{D}\boldsymbol{v} = 0. \tag{8}$$

Note that the last row of the $\boldsymbol{v}$-equation is, $(v_N)_t = 0$. Thus, since $v_N(0)=0$, it follows that $v_N(t)\equiv 0$. To prove that (7)-(8) is a stable scheme, we use the energy method (see e.g. [5,29]). For (7) we have

$$\frac{d}{dt}\|\boldsymbol{u}\|_H^2 = -2\boldsymbol{u}^T Q\,\boldsymbol{v},$$

and for (8), we obtain

$$\begin{aligned}\frac{d}{dt}\|\boldsymbol{v}\|_H^2 &= -\boldsymbol{v}^T H\tilde{D}\boldsymbol{u} - (\boldsymbol{v}^T H\tilde{D}\boldsymbol{u})^T + 2\boldsymbol{v}^T H\tilde{D}\boldsymbol{v} + 2(\boldsymbol{v}^T H\tilde{D}\boldsymbol{v})^T,\\ &= -2\boldsymbol{v}^T\tilde{Q}\,\boldsymbol{u} + 2\boldsymbol{v}^T(\tilde{Q}+\tilde{Q}^T)\boldsymbol{v} = -2\boldsymbol{v}^T\tilde{Q}\,\boldsymbol{u} + 2\boldsymbol{v}^T\tilde{B}\boldsymbol{v}.\end{aligned}$$

Adding the two estimates (and neglecting the left boundary terms emerging from (7)), we obtain

$$\frac{d}{dt}\left(\|\boldsymbol{u}\|_H^2 + \|\boldsymbol{v}\|_H^2\right) = -2\boldsymbol{u}^T Q\,\boldsymbol{v} - 2\boldsymbol{v}^T\tilde{Q}\,\boldsymbol{u} - 2v_0^2 + 2v_N v_{N-1}. \tag{9}$$

Since $v_N\equiv 0$ the following relations hold: $v_N v_{N-1}=0$, $\boldsymbol{v}^T\tilde{Q}\,\boldsymbol{u} = \boldsymbol{v}^T Q\,\boldsymbol{u}$ and $\boldsymbol{v}^T Q\,\boldsymbol{u} = \boldsymbol{v}^T(B-Q^T)\boldsymbol{u} = \boldsymbol{v}^T B\boldsymbol{u} - \boldsymbol{v}^T Q^T\boldsymbol{u} = -\boldsymbol{v}^T Q^T\boldsymbol{u}$. Furthermore, $-\boldsymbol{v}^T Q^T\boldsymbol{u} = -(\boldsymbol{u}^T Q\,\boldsymbol{v})^T$, and we conclude that $\boldsymbol{v}^T\tilde{Q}\,\boldsymbol{u} = -\boldsymbol{u}^T Q\,\boldsymbol{v}$. Hence, the two first terms on the right-hand side of (9) cancel, and our estimate reads

$$\frac{d}{dt}\left(\|\boldsymbol{u}\|_H^2 + \|\boldsymbol{v}\|_H^2\right) = -2v_0^2 \le 0,$$

which demonstrates that the semi-discrete scheme (7)-(8) is stable.

Remark. Note that the Dirichlet-SBP operator, $\tilde{D}$, need not be implemented. The same result is achieved by using D everywhere in (7)-(8) and setting $v_N=0$ after each Runge-Kutta stage. ⌟

4. The linearised compressible Navier-Stokes equations

Consider the compressible Navier-Stokes equations in one spatial dimension. These can be stated as

$$\mathrm{u}_t + \mathrm{f}^{\mathrm{I}}(\mathrm{u})_x = \mathrm{f}^{\mathrm{V}}(\mathrm{u},\mathrm{u}_x)_x, \qquad x\in\Omega=(0,1), \quad 0<t<\mathcal{T}, \tag{10}$$

where $\mathrm{u}=(\rho,m,E)^T$ are the conserved variables density, momentum ($m=\rho v$) and energy, and v denotes the velocity. $\mathrm{f}^{\mathrm{I}} = \left(m,\rho v^2+p, v(E+p)\right)^T$, is the inviscid flux where p denotes the pressure, which is related to the conserved quantities through $p=(\gamma-1)\left(E-\frac{1}{2}\rho v^2\right)$, where $\gamma=\frac{c_p}{c_v}$ is the ratio of the specific heats at constant pressure and volume. Furthermore, $\mathrm{f}^{\mathrm{V}} = (0,(2\mu+\lambda)v_x,(2\mu+\lambda)vv_x+\kappa\mathrm{T}_x)^T$ is the viscous flux, where T denotes the temperature, given by the ideal gas law $\mathrm{T}=\frac{p}{\mathcal{R}\rho}$, where $\mathcal{R}$ is the gas constant. Moreover, μ and λ denote the viscosity parameters, and we assume Stokes hypothesis, $\lambda=-\frac{2}{3}\mu$, with $\mu>0$. Lastly, κ denotes the thermal conductivity. (Below, we use c to denote the speed of sound.) The equations are augmented with the adiabatic wall boundary conditions,

$$v=0 \text{ (no-slip) and,} \quad \mathrm{T}_x=0. \tag{11}$$

To investigate linear well-posedness, the system (10) may be linearised and subsequently symmetrised with the symmetrising matrices found in [1]. (Since the details of the derivations are omitted in [1], we include them in Appendix A.1 for the reader's convenience.)

We repeat this procedure briefly. To linearise the equations, we decompose the primitive variables, $\mathrm{v}=(\rho,v,p)^T$, into their exact (known smooth and bounded) solution and a small smooth perturbation, e.g. $\rho=\rho_{\mathrm{ex}}+\rho'$, which yields a variable-coefficient problem. Then, we freeze the coefficients. Well-posedness of the variable-coefficient problem follows if all admissible frozen-coefficient problems are well-posed (see [9,13] for further information). The resulting linearised constant-coefficient problem is

$$\begin{aligned}\rho'_t + v^*\rho'_x + \rho^* v'_x &= 0,\\ v'_t + v^* v'_x + \tfrac{1}{\rho^*}p'_x &= \tfrac{2\mu+\lambda}{\rho^*}v'_{xx},\\ p'_t + \gamma p^* v'_x + v^* p'_x &= -\tfrac{\gamma\mu p^*}{\mathrm{Pr}\rho^{*2}}\rho'_{xx} + \tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}p'_{xx},\end{aligned} \tag{12}$$

where the star superscript, '$*$', indicates a frozen coefficient. Finally, we symmetrise the equations using the matrices S_p and S_p^{-1} from [1]. Using the linearised gas law (see Appendix A.2), we obtain

$$\mathrm{w}_t + \mathcal{A}\mathrm{w}_x = \mathcal{B}\mathrm{w}_{xx}, \tag{13}$$

where

$$\mathrm{w} = \left(\frac{c^*}{\sqrt{\gamma}\rho^*}\rho', v', \frac{\gamma\mathcal{R}}{c^*\sqrt{\gamma}\sqrt{\gamma-1}}\mathrm{T}' \right)^T$$

and

$$\mathcal{A} = \begin{pmatrix} v^* & \frac{c^*}{\sqrt{\gamma}} & 0 \\ \frac{c^*}{\sqrt{\gamma}} & v^* & \sqrt{\frac{\gamma-1}{\gamma}}c^* \\ 0 & \sqrt{\frac{\gamma-1}{\gamma}}c^* & v^* \end{pmatrix}, \quad \mathcal{B} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & \frac{2\mu+\lambda}{\rho^*} & 0 \\ 0 & 0 & \frac{\gamma\mu}{\mathrm{Pr}\rho^*} \end{pmatrix},$$

with $\mathrm{Pr} = \frac{c_p\mu}{\kappa}$ denoting the Prandtl number and $c_p = \frac{\gamma\mathcal{R}}{\gamma-1}$.

For completeness, we proceed by reviewing the well-posedness analysis found in e.g. [28]. Consider (13) on the spatial domain $\Omega = (0,1)$ with L^2-bounded initial data. The linearised boundary conditions (11), take the form

$$v'(\{0,1\},t) = 0, \quad v^*(\{0,1\},t) = 0, \quad \mathrm{T}'_x(\{0,1\},t) = 0. \tag{14}$$

(Note that admissible solutions satisfy the no-slip condition, whence $v^*(\{0,1\},t) = 0$.) The energy method and (14) lead to

$$\frac{d}{dt}\|\mathrm{w}\|^2 + 2\int_\Omega \mathrm{w}_x^T\mathcal{B}\mathrm{w}_x\,dx = 2\mathrm{w}^T\mathcal{B}\mathrm{w}_x|_0^1 - \mathrm{w}^T\mathcal{A}\mathrm{w}|_0^1 = 0.$$

Hence our problem is well-posed in the sense of Definition 2.1.

Remark. The Dirichlet condition $\mathrm{T}'(\{0,1\}) = 0$ would give the same result, but since the non-linear analysis later in this article requires $\mathrm{T}'_x(\{0,1\}) = 0$, we only consider the latter.

4.1. The semi-discrete scheme

Turning to the semi-discretisation of the problem (12) subject to the boundary conditions (14), we divide the spatial domain into $N+1$ equidistant grid points with grid spacing $h = 1/N$. Bold-face letters denote the numerical solution vectors.

To enforce the no-slip condition at both boundaries, the Dirichlet-SBP operator is defined by

$$\tilde{D} = \frac{1}{2h}\begin{pmatrix} 0 & 0 & 0 & \dots \\ -1 & 0 & 1 & \dots \\ & \ddots & & \\ & & -1 & 0 & 1 \\ & & 0 & 0 & 0 \end{pmatrix}, \quad \tilde{B} = \tilde{\mathrm{Q}} + \tilde{\mathrm{Q}}^T = \begin{pmatrix} 0 & -\frac{1}{2} & 0 & \dots & 0 & 0 \\ -\frac{1}{2} & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & 0 \\ \vdots & & & \ddots & & \vdots \\ & & & & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \dots & \frac{1}{2} & 0 \end{pmatrix}. \tag{15}$$

We introduce $\hat{\boldsymbol{\rho}} = \frac{c^*}{\sqrt{\gamma}\rho^*}\boldsymbol{\rho}'$ and $\hat{\mathrm{T}} = -\frac{c^*}{\rho^*\sqrt{\gamma}\sqrt{\gamma-1}}\boldsymbol{\rho}' + \sqrt{\frac{\gamma}{\gamma-1}}\frac{1}{\rho^* c^*}\boldsymbol{p}' = \frac{\gamma\mathcal{R}}{c^*\sqrt{\gamma}\sqrt{\gamma-1}}\mathrm{T}'$ and consider the following semi-discrete numerical scheme to approximate the system (12).

$$\hat{\boldsymbol{\rho}}_t + v^* D\hat{\boldsymbol{\rho}} + \tfrac{c^*}{\sqrt{\gamma}} D\boldsymbol{v}' = 0, \tag{16}$$

$$\boldsymbol{v}'_t + \tfrac{c^*}{\sqrt{\gamma}}\tilde{D}\hat{\boldsymbol{\rho}} + v^*\tilde{D}\boldsymbol{v}' + \sqrt{\tfrac{\gamma-1}{\gamma}}c^*\tilde{D}\hat{\mathrm{T}} = \tfrac{2\mu+\lambda}{\rho^*}\tilde{D}D\boldsymbol{v}', \tag{17}$$

$$\hat{\mathrm{T}}_t + \sqrt{\tfrac{\gamma-1}{\gamma}}c^* D\boldsymbol{v}' + v^* D\hat{\mathrm{T}} = \tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}DD\hat{\mathrm{T}} + \mathsf{SAT}, \tag{18}$$

where

$$\mathsf{SAT} = -\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}H^{-1}B(D\hat{\mathrm{T}} - 0), \tag{19}$$

imposes the homogeneous Neumann condition for the temperature weakly. We observe the following

- Since $v' \equiv 0$ at the boundaries initially, the use of the Dirichlet-SBP operator, $\tilde{D}$, in (17) ensures that v'_0, v'_N remains zero for all $t \geq 0$.
- Note that in (16) and (18), $\hat{\rho}_{0,N}$ and $\hat{\mathrm{T}}_{0,N}$ are unknowns that are updated in time.
- Note that the Dirichlet-SBP operator is applied only once for the second-derivative approximation in the right-hand side of (17). Using the Dirichlet-SBP operator twice would inadvertently impose the improper boundary condition $v'_x = 0$.
- When implementing the scheme, the $\tilde{D}$ is not necessary. One can equivalently compute all derivatives using D and reset the velocity to zero after each Runge-Kutta stage.

Remark. A similar scheme for the non-dimensional linearised and symmetrised Navier-Stokes equations was demonstrated to be stable using SAT to impose the no-slip conditions in [28]. The modifications of the SBP operator described here only affect the wall boundaries. Far-field boundaries may be handled in a stable manner using SAT, see [30].

Proposition 4.1. *The semi-discrete scheme* (16) - (18) *is energy stable.*

Remark. Linear stability of a numerical scheme using the injection method for imposing the no-slip condition and SAT to impose the temperature condition for the linearised and symmetrised constant-coefficient problem (13) was proven in [22]. However, in the present analysis, we use a different methodology for the injection method.

Proof. We carry out the energy analysis for each equation separately, before adding the three preliminary results to obtain the final energy estimate. (We neglect the right boundary for the rest of this analysis to reduce notation. Its treatment resembles the left boundary.) For (16), we obtain

$$\frac{d}{dt}\|\hat{\boldsymbol{\rho}}\|_H^2 = -v^*\hat{\boldsymbol{\rho}}^T(Q+Q^T)\hat{\boldsymbol{\rho}} - 2\tfrac{c^*}{\sqrt{\gamma}}\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}'.$$

Utilising the SBP-properties and subsequently $v^* = 0$ yield

$$\frac{d}{dt}\|\hat{\boldsymbol{\rho}}\|_H^2 = v^*\hat{\rho}_0^2 - 2\tfrac{c^*}{\sqrt{\gamma}}\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}' = -2\tfrac{c^*}{\sqrt{\gamma}}\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}'. \tag{20}$$

Next, Equation (17) results in

$$\frac{d}{dt}\|\boldsymbol{v}'\|_H^2 = -2\tfrac{c^*}{\sqrt{\gamma}}\boldsymbol{v}'^T\tilde{Q}\,\hat{\boldsymbol{\rho}} - 2\sqrt{\tfrac{\gamma-1}{\gamma}}c^*\boldsymbol{v}'^T\tilde{Q}\,\hat{\mathrm{T}}\underbrace{-v^*\boldsymbol{v}'^T(\tilde{Q}+\tilde{Q}^T)\boldsymbol{v}' + 2\tfrac{2\mu+\lambda}{\rho^*}\boldsymbol{v}'^T\tilde{Q}\,D\boldsymbol{v}'}_{A_1}. \tag{21}$$

Using (15), we obtain

$$A_1 = -v^*\boldsymbol{v}'^T\tilde{B}\boldsymbol{v}' + 2\tfrac{2\mu+\lambda}{\rho^*}\boldsymbol{v}'^T(\tilde{B}-\tilde{Q}^T)D\boldsymbol{v}', = v^*v'_0v'_1 + \tfrac{2\mu+\lambda}{\rho^*}\left(-v'_1(D\boldsymbol{v}')_0 - v'_0(D\boldsymbol{v}')_1 - 2(\tilde{D}\boldsymbol{v}')^T HD\boldsymbol{v}'\right),$$

and, using $v'_0 = 0$,

$$A_1 = \tfrac{2\mu+\lambda}{\rho^*}\left(-v'_1\frac{v'_1 - v'_0}{h} - 2(\tilde{D}\boldsymbol{v}')^T HD\boldsymbol{v}'\right) \leq -\tfrac{2\mu+\lambda}{\rho^*}\left(2(\tilde{D}\boldsymbol{v}')^T HD\boldsymbol{v}'\right).$$

The estimate (21) therefore reduces to

$$\frac{d}{dt}\|\boldsymbol{v}'\|_H^2 \leq -2\tfrac{c^*}{\sqrt{\gamma}}\boldsymbol{v}'^T\tilde{Q}\,\hat{\boldsymbol{\rho}} - 2\sqrt{\tfrac{\gamma-1}{\gamma}}c^*\boldsymbol{v}'^T\tilde{Q}\,\hat{\mathrm{T}} - 2\tfrac{2\mu+\lambda}{\rho^*}(\tilde{D}\boldsymbol{v}')^T HD\boldsymbol{v}'. \tag{22}$$

The energy method for (18) with (19) gives

$$\frac{d}{dt}\|\hat{\mathrm{T}}\|_H^2 = -2\sqrt{\tfrac{\gamma-1}{\gamma}}c^*\hat{\mathrm{T}}Q\,\boldsymbol{v}'\underbrace{-v^*\hat{\mathrm{T}}^T(Q+Q^T)\hat{\mathrm{T}} + 2\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}\hat{\mathrm{T}}^T Q\,D\hat{\mathrm{T}} - 2\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}\hat{\mathrm{T}}^T BD\hat{\mathrm{T}}}_{A_2}. \tag{23}$$

Using $Q + Q^T = B$ and $v^* = 0$, yield

$$A_2 = v^*\hat{\mathrm{T}}_0^2 + 2\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}\left(\hat{\mathrm{T}}^T(B-Q^T)D\hat{\mathrm{T}} - \hat{\mathrm{T}}^T BD\hat{\mathrm{T}}\right) = v^*\hat{\mathrm{T}}_0^2 - 2\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}\hat{\mathrm{T}}^T Q^T D\hat{\mathrm{T}} = -2\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}(D\hat{\mathrm{T}})^T H(D\hat{\mathrm{T}}).$$

Hence, (23) results in

$$\frac{d}{dt}\|\hat{\mathrm{T}}\|_H^2 = -2\sqrt{\tfrac{\gamma-1}{\gamma}}c^*\hat{\mathrm{T}}^T Q\,\boldsymbol{v}' - 2\tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}(D\hat{\mathrm{T}})^T H(D\hat{\mathrm{T}}). \tag{24}$$

We add all the preliminary estimates (20), (22) and (24) to obtain

$$\frac{d}{dt}\left(\|\hat{\boldsymbol{\rho}}\|_H^2+\|\boldsymbol{v}'\|_H^2+\|\hat{\mathrm{T}}\|_H^2\right)\leq -2\frac{c^*}{\gamma}\left(\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}'+\boldsymbol{v}'^T\tilde{Q}\,\hat{\boldsymbol{\rho}}\right)-2\sqrt{\frac{\gamma-1}{\gamma}}c^*\left(\boldsymbol{v}'^T\tilde{Q}\,\hat{\mathrm{T}}+\hat{\mathrm{T}}^T Q\,\boldsymbol{v}'\right)$$
$$-2\frac{2\mu+\lambda}{\rho}(\tilde{D}\boldsymbol{v}')^T H\tilde{D}\boldsymbol{v}'-2\frac{\gamma\mu}{\mathrm{Pr}\rho}(D\hat{\mathrm{T}})^T HD\hat{\mathrm{T}}. \tag{25}$$

Consider the term $\boldsymbol{v}'^T\tilde{Q}\,\hat{\boldsymbol{\rho}}$. As in Example 3.1, it follows from $v_0'=v_N'\equiv 0$ that $\boldsymbol{v}'^T\tilde{Q}\,\hat{\boldsymbol{\rho}}\equiv\boldsymbol{v}'^T Q\,\hat{\boldsymbol{\rho}}=\boldsymbol{v}'^T B\hat{\boldsymbol{\rho}}-\boldsymbol{v}'^T Q^T\hat{\boldsymbol{\rho}}=-(\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}')^T$. Since $-(\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}')^T$ is a scalar we obtain $\boldsymbol{v}'^T\tilde{Q}\,\hat{\boldsymbol{\rho}}=-\hat{\boldsymbol{\rho}}^T Q\,\boldsymbol{v}'$. The same argument holds for the term $\boldsymbol{v}'^T\tilde{Q}\,\hat{\mathrm{T}}$, and the two first terms in (25) therefore vanish. Lastly, since $(\tilde{D}\boldsymbol{v}')_0=0$, we have $(\tilde{D}\boldsymbol{v}')^T HD\boldsymbol{v}'=(\tilde{D}\boldsymbol{v}')^T H\tilde{D}\boldsymbol{v}'$, and we obtain

$$\frac{d}{dt}\left(\|\hat{\boldsymbol{\rho}}\|_H^2+\|\boldsymbol{v}'\|_H^2+\|\hat{\mathrm{T}}\|_H^2\right)+2\frac{2\mu+\lambda}{\rho^*}\|\tilde{D}\boldsymbol{v}\|_H^2+2\frac{\gamma\mu}{\mathrm{Pr}\rho^*}\|D\hat{\mathrm{T}}\|_H^2\leq 0. \tag{26}$$

Hence, the scheme is stable in the sense of Definition 2.2. □

5. The non-linear Navier-Stokes equations

The semi-discrete scheme approximating (10) is given by

$$\mathbf{u}_t+\mathcal{D}^{\mathrm{I}}\mathbf{f}^{\mathrm{I}}=\mathcal{D}^{\mathrm{V}}\mathbf{f}^{\mathrm{V}}+\mathsf{SAT}, \tag{27}$$

where $\mathbf{u}=(\boldsymbol{\rho},\boldsymbol{v},\boldsymbol{E})^T$ is the numerical solution vector. The convective term of (10) is, once again ignoring the right boundary, approximated by

$$(\mathcal{D}^{\mathrm{I}}\mathbf{f}^{\mathrm{I}})_i=\begin{cases}\frac{\mathbf{f}^{\mathrm{I}}_{1/2}-\mathbf{f}^{\mathrm{I}}_0}{h/2}, & i=0,\\ \frac{\mathbf{f}^{\mathrm{I}}_{i+1/2}-\mathbf{f}^{\mathrm{I}}_{i-1/2}}{h}, & i=1,\dots,N-1,\end{cases} \tag{28}$$

where

$$\mathbf{f}^{\mathrm{I}}_{i+1/2}=\frac{\mathbf{f}^{\mathrm{I}}_{i+1}+\mathbf{f}^{\mathrm{I}}_i}{2}-\frac{\delta_{i+1/2}\,(\mathbf{u}_{i+1}-\mathbf{u}_i)}{2} \tag{29}$$

is the approximation of the inviscid flux and $\mathbf{f}^{\mathrm{I}}_i=\mathbf{f}^{\mathrm{I}}(\mathbf{u}_i)$. The second term is artificial diffusion and for δ sufficiently large, the approximation is entropy stable in the sense of (52) below (see also [32]). Furthermore, the components of $\mathbf{f}^{\mathrm{I}}_0$ are

$$\mathrm{f}^{\mathrm{I},\rho}_0=(\boldsymbol{\rho}.\boldsymbol{v})_0, \tag{30}$$
$$\mathrm{f}^{\mathrm{I},m}_0=\mathbf{f}^{\mathrm{I},m}_{1/2}, \tag{31}$$
$$\mathrm{f}^{\mathrm{I},E}_0=(\boldsymbol{v}.(\boldsymbol{E}+\boldsymbol{p}))_0, \tag{32}$$

where the superscripts ρ,m,E denote which equation the vector element corresponds to, and the dot symbolises element-wise vector multiplications.

Remark. Note that (31) implies that the flux difference (28) is identically equal to zero at grid point x_0 for the momentum equation.

Next, the diffusive term of (10) is conveniently approximated on matrix form by

$$\mathcal{D}^{\mathrm{V}}\mathbf{f}^{\mathrm{V}}=\left(0,\tilde{D}((2\mu+\lambda)D\boldsymbol{v}),D((2\mu+\lambda)\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v}+\kappa D\mathrm{T})\right)^T, \tag{33}$$

(the definition of $\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v}$ is given in (36)-(37)) and the SAT is given by

$$\mathsf{SAT}=\left(0,0,-\kappa H^{-1}B(D\mathrm{T}-0)\right)^T. \tag{34}$$

Remark. The scheme for the momentum equation is $m_t=0$ on the boundary, i.e., $m_0(t)\equiv 0$ such that $v_0\equiv 0$ for all $t\geq 0$.

Remark. It is also possible to handle a heat-entropy flow boundary condition, where the temperature condition in (11) is replaced with $\kappa\frac{\mathrm{T}_x}{\mathrm{T}}=g$. The corresponding SAT would then take the form $\mathsf{SAT}=H^{-1}(\kappa BDT-Tg)$, which would yield an entropy stable scheme for appropriately chosen g.

In (33) we use

$$(D(\boldsymbol{a}.\boldsymbol{b}))_i = \begin{cases} \frac{a_1b_1-a_0b_0}{h} = a_1(D\boldsymbol{b})_0 + b_0(D\boldsymbol{a})_0 = b_1(D\boldsymbol{a})_0 + a_0(D\boldsymbol{b})_0 & i=0, \\ \frac{a_{i+1}b_{i+1}-a_{i-1}b_{i-1}}{2h} = \frac{a_{i+1}+a_{i-1}}{2}(D\boldsymbol{b})_i + \frac{b_{i+1}+b_{i-1}}{2}(D\boldsymbol{a})_i & i=1,\dots,N-1, \\ \frac{a_Nb_N-a_{N-1}b_{N-1}}{h} = a_N(D\boldsymbol{b})_N + b_{N-1}(D\boldsymbol{a})_N = b_N(D\boldsymbol{a})_N + a_N(D\boldsymbol{b})_N & i=N. \end{cases} \tag{35}$$

To distinguish between the two boundary rules, we introduce the following notation

$$\begin{aligned} \overset{b,x}{\boldsymbol{a}} &= \left(a_0, \quad \tfrac{a_2+a_0}{2}, \quad \dots, \quad \tfrac{a_N+a_{N-2}}{2}, \quad a_N\right), \\ \overset{i,x}{\boldsymbol{a}} &= \left(a_1, \quad \tfrac{a_2+a_0}{2}, \quad \dots, \quad \tfrac{a_N+a_{N-2}}{2}, \quad a_{N-1}\right), \end{aligned} \tag{36}$$

and

$$\left(\overset{b,x}{\boldsymbol{a}}.(D\boldsymbol{b})\right)_i = \begin{cases} a_0(D\boldsymbol{b})_0, & i=0, \\ \frac{a_{i+1}+a_{i-1}}{2}(D\boldsymbol{b})_i, & i=1,\dots N-1, \\ a_N(D\boldsymbol{b})_N, & i=N, \end{cases} \tag{37}$$

$$\left(\overset{i,x}{\boldsymbol{a}}.(D\boldsymbol{b})\right)_i = \begin{cases} a_1(D\boldsymbol{b})_0, & i=0, \\ \frac{a_{i+1}+a_{i-1}}{2}(D\boldsymbol{b})_i, & i=1,\dots N-1, \\ a_{N-1}(D\boldsymbol{b})_N, & i=N. \end{cases} \tag{38}$$

Here, superscript b signifies that $\boldsymbol{a}$ is taken at the *boundary* node and superscript i signifies that $\boldsymbol{a}$ is taken at the first neighbouring *interior* node. A similar relation holds for the averages taken in the y-direction. Using (37) and (38), we can rewrite (35) as

$$\begin{aligned} D_x(\boldsymbol{a}.\boldsymbol{b}) &= \overset{b,x}{\boldsymbol{a}}.(D_x\boldsymbol{b}) + \overset{i,x}{\boldsymbol{b}}.(D_x\boldsymbol{a}) = \overset{b,x}{\boldsymbol{b}}.(D_x\boldsymbol{a}) + \overset{i,x}{\boldsymbol{a}}.(D_x\boldsymbol{b}), \\ D_y(\boldsymbol{a}.\boldsymbol{b}) &= \overset{b,y}{\boldsymbol{a}}.(D_y\boldsymbol{b}) + \overset{i,y}{\boldsymbol{b}}.(D_y\boldsymbol{a}) = \overset{b,y}{\boldsymbol{b}}.(D_y\boldsymbol{a}) + \overset{i,y}{\boldsymbol{a}}.(D_y\boldsymbol{b}), \end{aligned} \tag{39}$$

where D_x and D_y approximate the x- and y-derivative, respectively.

Lastly, a similar rule holds for quotients

$$\left(D\left(\frac{\boldsymbol{a}}{\boldsymbol{b}}\right)\right)_i = \begin{cases} \frac{\frac{a_1}{b_1}-\frac{a_0}{b_0}}{h} = \frac{b_0(D\boldsymbol{a})_0-a_0(D\boldsymbol{b})_0}{b_0b_1} = \frac{b_1(D\boldsymbol{a})_0-a_1(D\boldsymbol{b})_0}{b_0b_1} & i=0, \\ \frac{\frac{a_{i+1}}{b_{i+1}}-\frac{a_{i-1}}{b_{i-1}}}{2h} = \frac{1}{b_{i+1}b_{i-1}}\left(\frac{b_{i+1}+b_{i-1}}{2}(D\boldsymbol{a})_i - \frac{a_{i+1}+a_{i-1}}{2}(D\boldsymbol{b})_i\right), & i=1,\dots,N-1 \\ \frac{\frac{a_N}{b_N}-\frac{a_{N-1}}{b_{N-1}}}{h} = \frac{b_N(D\boldsymbol{a})_N-a_N(D\boldsymbol{b})_N}{b_Nb_{N-1}} = \frac{b_{N-1}(D\boldsymbol{a})_N-a_{N-1}(D\boldsymbol{b})_N}{b_nb_{N-1}}, & i=N. \end{cases} \tag{40}$$

The inviscid term in (27) can equivalently be recast on matrix form. To this end, the artificial diffusion can be recognised as a second-derivative SBP operator with variable coefficients (see [18] and [19]). Define $D_2^{(\delta)} = H^{-1}(-D_{(\delta)}^T\tilde{\Delta}D_{(\delta)} + \bar{\Delta}S)$, where $D_2^{(\delta)}$, $D_{(\delta)}$, $\tilde{\Delta} = \frac{h}{2}\,\mathrm{diag}(\delta_{1/2}, \delta_{3/2}, \dots, \delta_{N-1/2})$ and $\bar{\Delta} = \mathrm{diag}(-\delta_0, 0, \dots, 0, \delta_N)$ correspond to the matrices D_2, D, $\tilde{B}$ and B, respectively, given in [18]. Then the artificial diffusion (AD) terms in (29) can be recast as

$$\begin{aligned} AD^\rho &= -hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{\rho}, \\ AD^m &= -hH^{-1}\tilde{D}_{(\delta)}^T\tilde{\Delta}D_{(\delta)}(\boldsymbol{\rho}.\boldsymbol{v}), \\ AD^E &= -hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{E}, \end{aligned} \tag{41}$$

where $\tilde{D}_{(\delta)}^T$ is the Dirichlet-SBP operator corresponding to $D_{(\delta)}^T$, i.e., it is $D_{(\delta)}^T$ with the elements of first and last row set to zero. Then, using the SBP operators (3) and (15), (28) can be restated as,

$$\mathcal{D}^{\mathrm{I}}\mathbf{f}^{\mathrm{I}} = \begin{pmatrix} D\mathbf{f}^{\mathrm{I},\rho} - AD^\rho \\ \tilde{D}\mathbf{f}^{\mathrm{I},m} - AD^m \\ D\mathbf{f}^{\mathrm{I},E} - AD^E \end{pmatrix}, \tag{42}$$

where $\mathbf{f}^{\mathrm{I},\rho} = \boldsymbol{\rho}.\boldsymbol{v}$, $\mathbf{f}^{\mathrm{I},m} = \boldsymbol{\rho}.\boldsymbol{v}.^2 + \boldsymbol{p}$, and $\mathbf{f}^{\mathrm{I},E} = \boldsymbol{v}.(\boldsymbol{E}+\boldsymbol{p})$.

5.1. Linear stability of the non-linear scheme

To demonstrate the linear stability of the scheme (27), we consider its linearisation and relate it to the scheme in Proposition (4.1). Using (41) and (42), the scheme (27) is recast as

$$\begin{aligned}
&\boldsymbol{\rho}_t + D(\boldsymbol{\rho}.\boldsymbol{v}) + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{\rho} = 0,\\
&(\boldsymbol{\rho}.\boldsymbol{v})_t + \tilde{D}(\boldsymbol{\rho}.\boldsymbol{v}.^2 + \boldsymbol{p}) + hH^{-1}\tilde{D}_{(\delta)}^T\tilde{\Delta}D_{(\delta)}(\boldsymbol{\rho}.\boldsymbol{v}) = (2\mu+\lambda)\tilde{D}D\boldsymbol{v},\\
&\boldsymbol{E}_t + D(\boldsymbol{v}.(\boldsymbol{E}+\boldsymbol{p})) + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{E} = (2\mu+\lambda)D(\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v}) + \kappa(DD\mathrm{T} - H^{-1}BD\mathrm{T}).
\end{aligned} \tag{43}$$

In analogy with the continuous problem, we insert into the scheme the decomposition, $\mathbf{v}^T = (\boldsymbol{\rho}, \boldsymbol{v}, \boldsymbol{p}) = (\boldsymbol{\rho}_{\mathrm{ex}}, \boldsymbol{v}_{\mathrm{ex}}, \boldsymbol{p}_{\mathrm{ex}}) + (\boldsymbol{\rho}', \boldsymbol{v}', \boldsymbol{p}')$. (Smooth exact solution and a perturbation. Cf. Appendix A.1.) In the subsequent linearisation process, we neglect terms that are quadratically small in the perturbations and we omit zeroth-order terms since they do not affect stability (see [9]). The smooth exact solution satisfies the scheme up to a bounded truncation error, which is benign with respect to stability. Furthermore, linearisation of the variable-coefficient second-derivative approximations, yields terms on the form $D_{(\delta)}^T\tilde{\Delta}'D_{(\delta)}\boldsymbol{\beta}_{\mathrm{ex}}$ (where $\boldsymbol{\beta}_{\mathrm{ex}}$ is any of the independent variables). Due to the form of $\tilde{\Delta} \sim \boldsymbol{v} + \mathrm{T}$, these terms are bounded by the corresponding principal terms emanating from the momentum and energy equation. (All the terms that are assumed to be bounded or linear in the principal variable in the linearisation are denoted as $\mathcal{O}(1, \mathbf{v})$ in the derivations below.) The linearised equation scheme (43) becomes:

$$\begin{aligned}
&\boldsymbol{\rho}_t' + D(\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}}) + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{\rho}' = \mathcal{O}(1,\mathbf{v}),\\
&\boldsymbol{v}_t' + \tfrac{1}{\boldsymbol{\rho}_{\mathrm{ex}}}.\tilde{D}(2\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}}.^2 + \boldsymbol{p}') - \tfrac{\boldsymbol{v}_{\mathrm{ex}}.}{\boldsymbol{\rho}_{\mathrm{ex}}}.D(\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}})\\
&\quad - \tfrac{\boldsymbol{v}_{\mathrm{ex}}.}{\boldsymbol{\rho}_{\mathrm{ex}}}.hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{\rho}' + \tfrac{1}{\boldsymbol{\rho}_{\mathrm{ex}}}.hH^{-1}\tilde{D}_{(\delta)}^T\tilde{\Delta}D_{(\delta)}(\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}}) = \tfrac{2\mu+\lambda}{\boldsymbol{\rho}_{\mathrm{ex}}}.\tilde{D}D\boldsymbol{v}' + \mathcal{O}(1,\mathbf{v}),\\
&\frac{\boldsymbol{p}_t'}{\gamma-1} + \tfrac{1}{2}\boldsymbol{v}_{\mathrm{ex}}.^2D(\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}}) - \boldsymbol{v}_{\mathrm{ex}}.\tilde{D}(2\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'\boldsymbol{v}_{\mathrm{ex}}.^2 + \boldsymbol{p}')\\
&\qquad + D\left(\tfrac{\gamma}{\gamma-1}(\boldsymbol{v}'\boldsymbol{p}_{\mathrm{ex}} + \boldsymbol{v}_{\mathrm{ex}}\boldsymbol{p}') + \tfrac{3}{2}\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}_{\mathrm{ex}}.\boldsymbol{v}' + \tfrac{1}{2}\boldsymbol{\rho}'\boldsymbol{v}_{\mathrm{ex}}.^3\right)\\
&\quad + \tfrac{1}{2}\boldsymbol{v}_{\mathrm{ex}}.^2hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{\rho}' - \boldsymbol{v}_{\mathrm{ex}}.hH^{-1}\tilde{D}_{(\delta)}^T\tilde{\Delta}D_{(\delta)}(\boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}' + \boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}})\\
&\qquad + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}(\tfrac{\boldsymbol{p}'}{\gamma-1} + \boldsymbol{\rho}_{\mathrm{ex}}.\boldsymbol{v}_{\mathrm{ex}}.\boldsymbol{v}' + \tfrac{1}{2}\boldsymbol{\rho}'.\boldsymbol{v}_{\mathrm{ex}}.^2) = (2\mu+\lambda)\left(D(\overset{b,x}{\boldsymbol{v}}_{\mathrm{ex}}.D\boldsymbol{v}') - \boldsymbol{v}_{\mathrm{ex}}.\tilde{D}D\boldsymbol{v}'\right)\\
&\qquad\qquad + \tfrac{\kappa}{\mathcal{R}}\left(DD\left(\tfrac{\boldsymbol{p}'.}{\boldsymbol{\rho}_{\mathrm{ex}}} - \tfrac{\boldsymbol{p}_{\mathrm{ex}}.\boldsymbol{\rho}'.}{\boldsymbol{\rho}_{\mathrm{ex}}.^2}\right)\right.\\
&\qquad\qquad \left. - H^{-1}BD\left(\tfrac{\boldsymbol{p}'.}{\boldsymbol{\rho}_{\mathrm{ex}}} - \tfrac{\boldsymbol{p}_{\mathrm{ex}}.\boldsymbol{\rho}'.}{\boldsymbol{\rho}_{\mathrm{ex}}.^2}\right)\right) + \mathcal{O}(1,\mathbf{v}).
\end{aligned}$$

Next, we use a result found in the proof of Lemma 2.2 of [20]: For a known continuously differentiable function, $a_{\mathrm{ex}}(x,t)$,

$$D(\boldsymbol{a}_{\mathrm{ex}}.\boldsymbol{b}') = \boldsymbol{a}_{\mathrm{ex}}.D\boldsymbol{b}' + \text{zeroth-order terms of } \boldsymbol{b}'.$$

The zeroth-order terms can then be included in the $\mathcal{O}(1,\mathbf{v})$ terms. Furthermore, we obtain first-derivative approximations of $\boldsymbol{v}'$, $\boldsymbol{p}'$ and $\boldsymbol{\rho}'$ in the pressure equation, but as in the continuous analysis, they are bounded by the corresponding principal terms, see Appendix A.1. Hence, the part of the scheme that affects the linear stability reduces to

$$\begin{aligned}
&\boldsymbol{\rho}_t' + \boldsymbol{\rho}_{\mathrm{ex}}.D\boldsymbol{v}' + \boldsymbol{v}_{\mathrm{ex}}.D\boldsymbol{\rho}' + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{\rho}' = \mathcal{O}(1,\mathbf{v}),\\
&\boldsymbol{v}_t' + \boldsymbol{v}_{\mathrm{ex}}.\tilde{D}\boldsymbol{v}' + \tfrac{1}{\boldsymbol{\rho}_{\mathrm{ex}}}.\tilde{D}\boldsymbol{p}' + hH^{-1}\tilde{D}_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{v}' = \tfrac{2\mu+\lambda}{\boldsymbol{\rho}_{\mathrm{ex}}}.\tilde{D}D\boldsymbol{v}' + \mathcal{O}(1,\mathbf{v}),\\
&\boldsymbol{p}_t' + \boldsymbol{v}_{\mathrm{ex}}.D\boldsymbol{p}' + \gamma\,\boldsymbol{p}_{\mathrm{ex}}.D\boldsymbol{v}' + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{p}' = (2\mu+\lambda)(\gamma-1)(\overset{b,x}{\boldsymbol{v}}_{\mathrm{ex}} - \boldsymbol{v}_{\mathrm{ex}}).\tilde{D}D\boldsymbol{v}'\\
&\qquad\qquad + \tfrac{\kappa(\gamma-1)}{\mathcal{R}}\left(\tfrac{1}{\boldsymbol{\rho}_{\mathrm{ex}}}.DD\boldsymbol{p} - \tfrac{\boldsymbol{p}_{\mathrm{ex}}.}{\boldsymbol{\rho}_{\mathrm{ex}}.^2}.DD\boldsymbol{\rho}' - \tfrac{1}{\boldsymbol{\rho}_{\mathrm{ex}}}.H^{-1}BD\boldsymbol{p}'\right.\\
&\qquad\qquad \left. + \tfrac{\boldsymbol{p}_{\mathrm{ex}}.}{\boldsymbol{\rho}_{\mathrm{ex}}.^2}.H^{-1}BD\boldsymbol{\rho}'\right) + \mathcal{O}(1,\mathbf{v}).
\end{aligned}$$

In order to symmetrise the system, we must freeze the coefficients. That is, we assume that the variable coefficients are constants. Specifically, with $\boldsymbol{v}_{\mathrm{ex}} = \textit{constant}$, the difference $(\overset{b,x}{\boldsymbol{v}}_{\mathrm{ex}} - \boldsymbol{v}_{\mathrm{ex}}) = 0$, since $\overset{b,x}{\boldsymbol{v}}_{\mathrm{ex}}$ is an arithmetic mean of $\boldsymbol{v}_{\mathrm{ex}}$. This would immediately take us to (16)-(18) (plus some benign terms), which we already know is stable. However, *the method of freezing the coefficients is only allowed if it implies stability of the variable-coefficient problems.*

Remark. Consider the advection equation, $u_t + a(x,t)u_x = 0$ whose estimate is $\partial_t \|u\|^2 + au^2|_0^1 + \int_0^1 a_x u^2\, dx = 0$, and analogously for any SBP semi-discretisation. Clearly, if a bound is obtained for any constant a within the range of a, and $|a_x|$ is bounded, the variable-coefficient problem is also bounded.

Here, it is straightforward that the above principle applies to most terms. However, we can not immediately omit the term with $(\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}})$ in the "freezing" process, since it is a part of a second-derivative term of velocity in the temperature equation. Hence, we keep it as a variable coefficient while all other coefficients are frozen (signified with the superscript star). By applying the symmetrising matrices from [1] and introducing $\hat{\boldsymbol{\rho}} = \frac{c^*}{\sqrt{\gamma}\rho^*}\boldsymbol{\rho}'$ and $\hat{\mathrm{T}} = -\frac{c^*}{\rho^*\sqrt{\gamma}\sqrt{\gamma-1}}\boldsymbol{\rho}' + \sqrt{\frac{\gamma}{\gamma-1}}\frac{1}{\rho^* c^*}\boldsymbol{p}' = \frac{\gamma\mathcal{R}}{c^*\sqrt{\gamma}\sqrt{\gamma-1}}\mathrm{T}'$ to reduce notation, we arrive at

$$\hat{\boldsymbol{\rho}}_t + v^* D\hat{\boldsymbol{\rho}} + \tfrac{c^*}{\sqrt{\gamma}} D\boldsymbol{v}' + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\hat{\boldsymbol{\rho}} = \mathcal{O}(1,\mathbf{v}), \tag{44}$$

$$\boldsymbol{v}_t' + \tfrac{c^*}{\sqrt{\gamma}}\tilde{D}\hat{\boldsymbol{\rho}} + v^*\tilde{D}\boldsymbol{v}' + \sqrt{\tfrac{\gamma-1}{\gamma}}c^*\tilde{D}\hat{\mathrm{T}} + hH^{-1}\tilde{D}_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\boldsymbol{v}' = \tfrac{2\mu+\lambda}{\rho^*}\tilde{D}D\boldsymbol{v}' + \mathcal{O}(1,\mathbf{v}), \tag{45}$$

$$\begin{aligned}\hat{\mathrm{T}}_t + \sqrt{\tfrac{\gamma-1}{\gamma}}c^* D\boldsymbol{v}' + v^* D\hat{\mathrm{T}} + hH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\hat{\mathrm{T}} &= \tfrac{\sqrt{\gamma}\sqrt{\gamma-1}}{\rho^* c^*}(2\mu+\lambda)(\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}}).\tilde{D}D\boldsymbol{v}' \\ &\quad + \tfrac{\gamma\mu}{\mathrm{Pr}\rho^*}\left(DD\hat{\mathrm{T}} - H^{-1}BD\hat{\mathrm{T}}\right) + \mathcal{O}(1,\mathbf{v}).\end{aligned} \tag{46}$$

Note the resemblance to (16)-(18).

Proposition 5.1. *The non-linear scheme (43) is linearly stable.*

Proof. Linearising and symmetrising the non-linear scheme (43) leads to (44)-(46). Linear stability can then be established by employing the discrete energy method. In this process, the terms that differ from the analysis in Proposition 4.1 are the artificial diffusion terms, the $\mathcal{O}(1,\mathbf{v})$-terms and the additional (velocity dependent) diffusive term in (46). The last one is the only non-trivial term. Hence, we only consider the temperature equation (the artificial diffusion terms and the $\mathcal{O}(1,\mathbf{v})$-terms are handled in the same way in (44) - (45)).

In the energy analysis for equation (46) we multiply by $\hat{\mathrm{T}}^T H$ and add the transpose. We focus on the terms that differ from the scheme (18),

$$\begin{aligned}&2h\hat{\mathrm{T}}^T HH^{-1}D_{(\delta)}^T\tilde{\Delta}D_{(\delta)}\hat{\mathrm{T}} - 2\tfrac{\sqrt{\gamma}\sqrt{\gamma-1}}{\rho^* c^*}(2\mu+\lambda)\hat{\mathrm{T}}^T H\left((\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}}).DD\boldsymbol{v}'\right) + 2\hat{\mathrm{T}}^T H\mathcal{O}(1,\mathbf{v}) \\ &= 2h(D_{(\delta)}\hat{\mathrm{T}})^T\tilde{\Delta}D_{(\delta)}\hat{\mathrm{T}} + 2\tfrac{\sqrt{\gamma}\sqrt{\gamma-1}}{\rho^* c^*}(2\mu+\lambda)(D(\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}}).\hat{\mathrm{T}})^T H\left(D\boldsymbol{v}'\right) + 2\hat{\mathrm{T}}^T H\mathcal{O}(1,\mathbf{v}).\end{aligned} \tag{47}$$

The first term is quadratic with positive sign since $\tilde{\Delta}$ is diagonal and $\tilde{\Delta}_{ii} = \delta_{i+1/2} \geq 0$. Moreover, the last term will at most contribute with a finite growth in the final estimate. The only term that requires attention is: $(D(\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}}).\hat{\mathrm{T}})^T H(D\boldsymbol{v}')$. Since $\overset{b,x}{\boldsymbol{v}}_{\text{ex}}$ is an average of the smooth function $\boldsymbol{v}_{\text{ex}}$, $(\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}}) \sim \mathcal{O}(h)$. Hence, $(D(\overset{b,x}{\boldsymbol{v}}_{\text{ex}} - \boldsymbol{v}_{\text{ex}}).\hat{\mathrm{T}})^T H(D\boldsymbol{v}') \leq C\|D\hat{\mathrm{T}}\|\|L(\boldsymbol{v}')\|$ where $L(\boldsymbol{v}')$ represents a vector whose entries are linear combinations of the elements of $\boldsymbol{v}'$. Hence, these terms do not cause an unbounded growth in the final estimate for all components. (They enter the estimate corresponding to (26) as $\|D\hat{\mathrm{T}}\|^2$ and $\|\boldsymbol{v}'\|^2$ terms.) □

It is common to demonstrate linear stability of non-linear schemes by directly considering (16)-(18). Here, we have rigorously proven that the non-linear scheme (43) indeed reduces to (16)-(18).

Remark. A 3rd-order version of (43) is obtained as follows: Replace the difference operators D, with the diagonal-norm (4,2)-scheme that is 4th-order in the interior and 2nd-order near the boundaries, and $\tilde{D}$ with its counterpart obtained by zeroing the first and last row in the (4,2)-operator. (See [8] for information on the (4,2)-operator and [17] for high-order version of artificial diffusion.) We have verified the linearisation process for this scheme leading to (16)-(18) (now with the (4,2)-operators) and proven stability for the symmetrised scheme. (Note that the high-order scheme produce a different set of boundary terms in the energy estimate.)

5.2. Entropy

For the non-linear analysis, we give a brief summary of the theory of entropy and refer the reader to the papers of Harten ([10]) and Tadmor ([32]) for a more comprehensive introduction.

Hyperbolic conservation laws take the form

$$u_t + f(u)_x = 0, \qquad x \in \mathbb{R}. \tag{48}$$

A strictly convex scalar function, $U(u)$, is said to be an *entropy function* of the problem (48), if it satisfies the relation $U_u^T f_u = F_u$, where F is the entropy flux function. $U_u = w$ are the *entropy variables*, which symmetrises the problem (48) (see e.g. [12]). Furthermore, the scalar function $\psi = w^T f - F$ is called the entropy potential. By multiplying by the entropy variables, the equation (48) can be recast as

$$U(u)_t + F(u)_x = 0,$$

which is satisfied for smooth solutions of the problem (48). However, it is well-known that solutions of (48) can develop shocks (even from continuous data), and we therefore have to consider weak solutions that satisfy

$$\int_{\mathbb{R}} \phi u_t \, dx - \int_{\mathbb{R}} \phi_x f(u) \, dx = 0,$$

for any $\phi \in \mathcal{C}^\infty$ with compact support. Weak solutions are generally not unique, however, the physically relevant solution satisfies the Second Law of Thermodynamics, which states that the entropy within a closed system cannot decrease. In mathematical terms, this can be stated as the entropy inequality

$$U(u)_t + F(u)_x \leq 0. \tag{49}$$

Solutions that satisfy the *entropy inequality* (49) are called *entropy solutions* (see e.g. [32]).

5.3. Entropy stability

In order to carry the concept of entropy over to the semi-discrete setting, we consider a scheme of the form

$$(u_i)_t + \frac{g_{i+1/2} - g_{i-1/2}}{h_i} = 0, \tag{50}$$

where h_i is the distance between grid node $i+1$ and i. Furthermore, $g_{i+1/2} = \frac{f_{i+1} + f_i - \delta_{i+1/2}(u_{i+1} - u_i)}{2}$ is the approximation of the flux $f(u)$, where $\delta_{i+1/2}(u_{i+1} - u_i)$, with $\delta_{i+1/2} \geq 0$, is an artificial dissipation term. Schemes such as (50) are termed entropy stable, if they satisfy the discrete entropy inequality

$$(U_i)_t + \frac{F_{i+1/2} - F_{i-1/2}}{h_i} \leq 0, \quad \text{for all } i, \tag{51}$$

where $F_{i+1/2} = \frac{1}{2}\left((\boldsymbol{w}_{i+1}^T + \boldsymbol{w}_i^T)\boldsymbol{g}_{i+1/2}\right) - \frac{1}{2}(\psi_{i+1} + \psi_i)$, and $\psi_i = \psi(u_i)$. This holds true for schemes where δ is chosen such that the flux approximation satisfies *Tadmor's shuffle condition*,

$$\langle \Delta w_{i+1/2}, g_{i+1/2} \rangle \leq \Delta \psi_{i+1/2} = \psi_{i+1} - \psi_i, \tag{52}$$

where $\Delta w_{i+1/2} = w_{i+1} - w_i$. See [32] for more details.

5.4. Entropy analysis for the 1-D Navier-Stokes equations

Consider the continuous problem (10) augmented with the no-slip wall boundary condition $v(0,t) = 0$, and a Neumann condition on the temperature; $\mathrm{T}_x(0,t) = 0$ (neglecting the right boundary), and L^2-bounded initial data. (The entropy estimate for this problem is derived in e.g. [23] and also [31], but we repeat it here for completeness.)

For the compressible Navier-Stokes equations, there is only one entropy function ([11]); $U(u) = -\rho\mathcal{S}$ with $F(u) = -m\mathcal{S}$ and $\psi = (\gamma - 1)m$, where $\mathcal{S} = \ln\left(\frac{p}{\rho^\gamma}\right)$, and $\mathcal{S}$ is the specific entropy. For this entropy function, the entropy variables are given by

$$w^T = -\frac{1}{c_v \mathrm{T}}\left(\frac{v^2}{2} + c_v \mathrm{T}(\mathcal{S} - \gamma), -v, 1\right).$$

To obtain an entropy estimate, multiply Equation (10) by the entropy variables, w^T, and integrate over the spatial domain $\Omega = (0,1)$,

$$\int_\Omega \mathrm{U}(\mathrm{u})_t \, dx + \int_\Omega \mathrm{F}(\mathrm{u})_x \, dx = \int_\Omega w^T \mathrm{f}^\mathrm{v}(\mathrm{u}, \mathrm{u}_x)_x \, dx,$$

which leads to

$$\int_\Omega \mathrm{U}(\mathrm{u})_t \, dx - \mathrm{F}(\mathrm{u})|_0 = -w^T \mathrm{f}^\mathrm{v}(\mathrm{u}, \mathrm{u}_x)|_0 - \int_\Omega w_x^T \mathrm{f}^\mathrm{v}(\mathrm{u}, \mathrm{u}_x) \, dx.$$

Since $F = -m\mathcal{S}$, we have that $F(\mathrm{u})|_0 = 0$, due to the no-slip boundary condition. Furthermore, the term $w^T \mathrm{f}^{\mathrm{v}}(\mathrm{u}, \mathrm{u}_x)|_0$ reduces to

$$w^T \mathrm{f}^{\mathrm{v}}(\mathrm{u}, \mathrm{u}_x)\big|_0 = -\frac{(\gamma-1)\kappa}{\mathcal{R}} \frac{\mathrm{T}_x}{\mathrm{T}}\bigg|_0 = 0,$$

(see e.g. [23]). The last equality is due to the Neumann condition at $x = 0$. Hence, the estimate reads

$$\int_\Omega U(\mathrm{u})_t \, dx = -\int_\Omega \frac{1}{c_v \mathrm{T}^2} \left((2\mu + \lambda) v_x^2 \mathrm{T} + \kappa \mathrm{T}_x^2 \right) dx. \tag{53}$$

Since admissible solutions satisfy $\mathrm{T} > 0$, the entropy is bounded from above.

5.4.1. Non-linear stability

We now turn to the non-linear analysis of the scheme (27), and show that it is entropy stable.

Proposition 5.2. *If* $\mathbf{f}^{\mathrm{I}}$ *satisfies* (52) *for* $i = 1, \dots, N-1$, *then the semi-discrete scheme* (27)

$$\mathbf{u}_t + \mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}} = \mathcal{D}^{\mathrm{V}} \mathbf{f}^{\mathrm{V}} + \mathsf{SAT}, \tag{27}$$

with (28)-(34), *approximating system* (10) *is entropy-stable in the sense of* (51).

Remark. The scheme (27) is inspired by the one proposed in [31] and [25].

Remark. Possible entropy stable choices of $\mathbf{f}^{\mathrm{I}}$ are for instance the local- and global Lax-Friedrichs schemes and entropy-fixed Roe schemes. An entropy conservative flux can be recast into the form of (29). For linear stability, it is evident that we need $\delta_{i+1/2} \geq 0$ for all i, (cf. (47)) which is not necessarily true for entropy *conservative* fluxes (see [6]). However, the non-linear analysis presented below holds also for entropy conservative fluxes.

Proof. For each grid point, multiply the scheme (27) by the corresponding entropy variable $\boldsymbol{w}_i^T = -\frac{1}{c_v \mathrm{T}_i}\left(\frac{v_i^2}{2} + c_v \mathrm{T}_i (\mathcal{S}_i - \gamma), -v_i, 1 \right)$ and the norm element H_{ii} (the i-th diagonal element of H), and sum over all nodes (and neglect the right boundary for brevity)

$$\sum_{i=0}^{N-1} \boldsymbol{w}_i^T H_{ii} (\mathbf{u}_i)_t + \sum_{i=0}^{N-1} \boldsymbol{w}_i^T H_{ii} (\mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}})_i = \sum_{i=0}^{N-1} \boldsymbol{w}_i^T H_{ii} (\mathcal{D}^{\mathrm{V}} \mathbf{f}^{\mathrm{V}})_i + \sum_{i=0}^{N-1} \boldsymbol{w}_i^T H_{ii} \mathsf{SAT}_i. \tag{54}$$

For the convective flux approximation, we perform the analysis using index notation in order to use the entropy stability results in [32]. The left-hand side of (54) is recast as

$$\sum_{i=0}^{N-1} \boldsymbol{w}_i^T H_{ii} (\mathbf{u}_i)_t + \sum_{i=0}^{N-1} \boldsymbol{w}_i^T H_{ii} (\mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}})_i = \sum_{i=0}^{N-1} H_{ii} (U_i)_t + \underbrace{\boldsymbol{w}_0^T H_{00} (\mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}})_0 + \sum_{i=1}^{N-1} \boldsymbol{w}_i^T H_{ii} \left(\mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}} \right)_i}_{A}. \tag{55}$$

Utilising (28) and the theory of [32], we manipulate A as

$$\begin{aligned} A = {} & \boldsymbol{w}_0^T H_{00} (\mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}})_0 - \sum_{i=1}^{N-1} F_{i-1/2} + \sum_{i=1}^{N-1} F_{i+1/2} \\ & - \sum_{i=1}^{N-1} \left(\frac{1}{2} (\boldsymbol{w}_i - \boldsymbol{w}_{i-1})^T \mathbf{f}^{\mathrm{I}}_{i-1/2} - \frac{1}{2} (\psi_i - \psi_{i-1}) \right) - \sum_{i=1}^{N-1} \left(\frac{1}{2} (\boldsymbol{w}_{i+1} - \boldsymbol{w}_i)^T \mathbf{f}^{\mathrm{I}}_{i+1/2} - \frac{1}{2} (\psi_{i+1} - \psi_i) \right), \end{aligned}$$

where $F_{i+1/2} = \frac{\boldsymbol{w}_{i+1}^T + \boldsymbol{w}_i^T}{2} \mathbf{f}^{\mathrm{I}}_{i+1/2} - \frac{\psi_{i+1} + \psi_i}{2}$. All F's except $F_{1/2}$ cancel due to the series' telescoping nature. Assuming that Tadmor's shuffle condition (52) is fulfilled, A reduces to

$$A \geq \boldsymbol{w}_0^T H_{00} (\mathcal{D}^{\mathrm{I}} \mathbf{f}^{\mathrm{I}})_0 - F_{1/2} = \boldsymbol{w}_0^T H_{00} \frac{\mathbf{f}^{\mathrm{I}}_{1/2} - \mathbf{f}^{\mathrm{I}}_0}{h/2} - F_{1/2} = \boldsymbol{w}_0^T (\mathbf{f}^{\mathrm{I}}_{1/2} - \mathbf{f}^{\mathrm{I}}_0) - F_{1/2} \geq -\boldsymbol{w}_0^T \mathbf{f}^{\mathrm{I}}_0 + \psi_0,$$

where we in the last step have used similar manipulations as for the interior nodes. Thanks to $v_0 = 0$, the entropy variable corresponding to the momentum equation is $w_0^m = \frac{v_0}{c_v \mathrm{T}_0} = 0$, $\psi = (\gamma - 1)(\boldsymbol{\rho}.\boldsymbol{v})_0 = 0$ and by (30) and (32), $\mathbf{f}^{\mathrm{I},\rho}_0 = (\boldsymbol{\rho}.\boldsymbol{v})_0 = 0$ and $\mathbf{f}^{\mathrm{I},E}_0 = (\boldsymbol{v}.(\boldsymbol{E} + \boldsymbol{p}))_0 = 0$ respectively, such that we obtain $A \geq 0$. Equation (54) therefore reduces to

$$\sum_{i=0}^{N} H_{ii}(U_i)_t \leq \sum_{i=0}^{N} \boldsymbol{w}_i^T H_{ii}(\mathcal{D}^{\mathrm{v}}\mathbf{f}^{\mathrm{v}})_i + \sum_{i=0}^{N} \boldsymbol{w}_i^T H_{ii}\mathsf{SAT}_i. \tag{56}$$

For the analysis of the diffusive term, we introduce the following vectors

$$\begin{aligned}
\boldsymbol{w}^m &= \left(w_0^m, \quad w_1^m, \quad w_2^m, \quad \ldots, \quad w_N^m\right)^T = \left(\tfrac{v_0}{c_v \mathrm{T}_0}, \quad \tfrac{v_1}{c_v \mathrm{T}_1}, \quad \tfrac{v_2}{c_v \mathrm{T}_2}, \quad \ldots, \quad \tfrac{v_N}{c_v \mathrm{T}_N}\right)^T, \\
\boldsymbol{w}^E &= \left(w_0^E, \quad w_1^E, \quad w_2^E, \quad \ldots, \quad w_N^E\right)^T = \left(-\tfrac{1}{c_v \mathrm{T}_0}, \quad -\tfrac{1}{c_v \mathrm{T}_1}, \quad -\tfrac{1}{c_v \mathrm{T}_2}, \quad \ldots, \quad -\tfrac{1}{c_v \mathrm{T}_N}\right)^T,
\end{aligned} \tag{57}$$

where the superscript denotes which equation the vector acts on. Using (33) and (34) the right-hand side of (56) can be restated using matrix notation as

$$\begin{aligned}
\sum_{i=0}^{N-1} \left(\boldsymbol{w}_i^T H_{ii}(\mathcal{D}^{\mathrm{v}}\mathbf{f}^{\mathrm{v}})_i + \boldsymbol{w}_i^T H_{ii}\mathsf{SAT}\right) = \underbrace{(\boldsymbol{w}^m)^T H \tilde{D}\left((2\mu+\lambda) D\boldsymbol{v}\right)}_{A_1} \\
\underbrace{+ (\boldsymbol{w}^E)^T H \left(D\left((2\mu+\lambda)\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v} + \kappa D\mathrm{T}\right) - \kappa H^{-1} B D\mathrm{T}\right)}_{A_2}.
\end{aligned} \tag{58}$$

Utilising that $H\tilde{D} = \tilde{Q} = \tilde{B} - \tilde{Q}^T = \tilde{B} - (H\tilde{D})^T$ and (15) we obtain

$$\begin{aligned}
A_1 &= (2\mu+\lambda)(\boldsymbol{w}^m)^T \tilde{B} D\boldsymbol{v} - (2\mu+\lambda)(\tilde{D}\boldsymbol{w}^m)^T H D\boldsymbol{v} \\
&= -\frac{(2\mu+\lambda)}{2}\left(w_1^m (D\boldsymbol{v})_0 + w_0^m (D\boldsymbol{v})_1\right) - (2\mu+\lambda)(\tilde{D}\boldsymbol{w}^m)^T H D\boldsymbol{v}.
\end{aligned}$$

Insert $w_1^m = \frac{v_1}{c_v \mathrm{T}_1}$ and $v_0 = 0$ to obtain

$$\begin{aligned}
A_1 &= -\frac{(2\mu+\lambda)}{2h}\frac{1}{c_v \mathrm{T}_1} v_1 (v_1 - v_0) - (2\mu+\lambda)(\tilde{D}\boldsymbol{w}^m)^T H D\boldsymbol{v}, \\
&= -\frac{(2\mu+\lambda)}{2h}\frac{1}{c_v \mathrm{T}_1} v_1^2 - (2\mu+\lambda)(\tilde{D}\boldsymbol{w}^m)^T H D\boldsymbol{v} \leq -(2\mu+\lambda)(\tilde{D}\boldsymbol{w}^m)^T H D\boldsymbol{v}.
\end{aligned} \tag{59}$$

Next, we turn to A_2 on the right-hand side of equation (58). Utilising the SBP properties, $HD = Q$ and $Q = B - Q^T$ yields

$$\begin{aligned}
A_2 &= (\boldsymbol{w}^E)^T B\left((2\mu+\lambda)\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v} + \kappa D\mathrm{T}\right) - \kappa(\boldsymbol{w}^E)^T B D\mathrm{T} - (\boldsymbol{w}^E)^T Q^T \left((2\mu+\lambda)\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v} + \kappa D\mathrm{T}\right), \\
&= -(D\boldsymbol{w}^E)^T H\left((2\mu+\lambda)\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v} + \kappa D\mathrm{T}\right),
\end{aligned} \tag{60}$$

where we used $\overset{b,x}{v}_0 = v_0 = 0$ in the last step.

Combining the preliminary results (56), (59) and (60) leads to

$$\sum_{i=0}^{N} H_{ii}(U_i)_t \leq -(2\mu+\lambda)\underbrace{\left((\tilde{D}\boldsymbol{w}^m)^T H D\boldsymbol{v} + (D\boldsymbol{w}^E)^T H(\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v})\right)}_{A_3} \underbrace{-\kappa(D\boldsymbol{w}^E)^T H D\mathrm{T}}_{A_4}.$$

Using (57) and the discrete product rule (39) result in

$$\begin{aligned}
A_3 &= \frac{1}{c_v}\left(\tilde{D}(\boldsymbol{v}.\mathrm{T}^{-1})\right)^T H D\boldsymbol{v} - \frac{1}{c_v}\left(D\mathrm{T}^{-1}\right)^T H(\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v}), \\
&= \frac{1}{c_v}\left(\left(\overset{i,x}{\mathrm{T}^{-1}}.\tilde{D}\boldsymbol{v}\right)^T H D\boldsymbol{v} + \left(\overset{b,x}{\boldsymbol{v}}.\tilde{D}\mathrm{T}^{-1}\right)^T H D\boldsymbol{v} - \left(D\mathrm{T}^{-1}\right)^T H(\overset{b,x}{\boldsymbol{v}}.D\boldsymbol{v})\right).
\end{aligned}$$

The first term in the last row is a discrete equivalent of the L^2-norm, $\left(\overset{i,x}{\mathrm{T}^{-1}}.\tilde{D}\boldsymbol{v}\right)^T H D\boldsymbol{v} = \sum_{i=0}^{N} \overset{i,x}{(\mathrm{T}^{-1})}_i (\tilde{D}\boldsymbol{v})_i H_{ii}(D\boldsymbol{v})_i = \|\sqrt{\overset{i,x}{\mathrm{T}^{-1}}}.\tilde{D}\boldsymbol{v}\|_H^2 \geq 0$, $(\mathrm{T} > 0)$. (Note that $(\tilde{D}\boldsymbol{v})_0 (D\boldsymbol{v})_0 = (\tilde{D}\boldsymbol{v})_0^2 = 0$.) Furthermore, it is easily verified that the two last terms cancel.

Lastly, by the discrete quotient rule (40), we have

$$A_4 = \kappa (D\boldsymbol{w}^{\mathrm{E}})^T H D\mathrm{T} = -\frac{1}{c_v}\kappa \left(D\mathrm{T}^{-1}\right)^T H D\mathrm{T} = \frac{\kappa \overset{x}{(\mathrm{T}^2)}^{-1}}{c_v}.(D\mathrm{T})^T H D\mathrm{T} = \frac{\kappa}{c_v}\|\sqrt{\overset{x}{(\mathrm{T}^2)^{-1}}}.D\mathrm{T}\|_H^2,$$

where $\overset{x}{(\mathrm{T})}_0^2 = \mathrm{T}_0\mathrm{T}_1$ and $\overset{x}{(\mathrm{T})}_i^2 = \mathrm{T}_{i-1}\mathrm{T}_{i+1}$, $i = 1, \ldots, N-1$. This term is non-negative as long as all T_i's > 0.

Finally, our entropy estimate (54) reads

$$\sum_{i=0}^{N} H_{ii}(U_i)_t \leq -(2\mu+\lambda)\|\sqrt{\overset{i,x}{\mathrm{T}^{-1}}}.\tilde{D}\boldsymbol{v}\|_H^2 - \frac{\kappa}{c_v}\|\sqrt{\overset{x}{(\mathrm{T}^2)^{-1}}}.D\mathrm{T}\|_H^2 \leq 0.$$

Hence, we conclude that our scheme is entropy stable. □

5.5. Non-linear analysis for the 2-D Navier-Stokes equations

Let $\Omega = (0,1)\times(0,1)$ be the spatial domain with boundary $\partial\Omega$. The compressible Navier-Stokes equations in two space dimensions are stated as

$$\mathsf{u}_t + \mathsf{f}^{\mathsf{I}}_x + \mathsf{g}^{\mathsf{I}}_y = \mathsf{f}^{\mathsf{V}}(\mathsf{u}, \mathsf{u}_x, \mathsf{u}_y)_x + \mathsf{g}^{\mathsf{V}}(\mathsf{u}, \mathsf{u}_x, \mathsf{u}_y)_y, \quad (x,y)\in\Omega = (0,1)^2, \quad 0 < t < \mathcal{T}, \tag{61}$$

where $\mathsf{u} = (\rho, m, n, E)^T$ are the conserved variables and

$$\begin{aligned}
\mathsf{f}^{\mathsf{I}} &= \left(m, \rho v_1^2 + p, \rho v_1 v_2, v_1(E+p)\right)^T,\\
\mathsf{g}^{\mathsf{I}} &= \left(n, \rho v_1 v_2, \rho v_2^2 + p, v_2(E+p)\right)^T,\\
\mathsf{f}^{\mathsf{V}} &= \left(0, 2\mu v_{1_x} + \lambda\left(v_{1_x} + v_{2_y}\right), \mu\left(v_{1_y} + v_{2_x}\right), v_1\left(2\mu v_{1_x} + \lambda\left(v_{1_x} + v_{2_y}\right)\right) + \mu v\left(v_{1_y} + v_{2_x}\right) + \kappa \mathrm{T}_x\right)^T,\\
\mathsf{g}^{\mathsf{V}} &= \left(0, \mu\left(v_{1_y} + v_{2_x}\right), 2\mu v_{2_y} + \lambda\left(v_{1_x} + v_{2_y}\right), v_2\left(2\mu v_{2_y} + \lambda\left(v_{1_x} + v_{2_y}\right)\right) + \mu v_1\left(v_{1_y} + v_{2_x}\right) + \kappa \mathrm{T}_y\right)^T,
\end{aligned}$$

are the inviscid and viscous fluxes; $n = \rho v_2$ is the momentum in the y-direction and v_1, v_2 denote the velocity components in the x- and y-directions, respectively. Equation (61) is augmented with no-slip boundary conditions and homogeneous Neumann conditions for the temperature, i.e.

$$v_1|_{\partial\Omega} = 0, \quad v_2|_{\partial\Omega} = 0, \quad \frac{\partial \mathrm{T}}{\partial n}|_{\partial\Omega} = 0, \tag{62}$$

and appropriate initial conditions. In 2-D, the entropy fluxes are $F_x = w^T \mathsf{f}^{\mathsf{I}}_x$ and $G_y = w^T \mathsf{g}^{\mathsf{I}}_y$, where $F = -m\mathcal{S}$ and $G = -n\mathcal{S}$. Following the same procedure as for the one-dimensional case, we can demonstrate that this problem satisfies the entropy inequality (49) (see again [23] or [31] for the derivation in 3-D). That is, multiply equation (61) by the entropy variables $w^T = -\frac{1}{c_v \mathrm{T}}\left(\frac{v_1^2+v_2^2}{2} + c_v\mathrm{T}(\mathcal{S}-\gamma), -v_1, -v_2, 1\right)$ and integrate over the spatial domain. Apply integration-by-parts to the entropy flux function and the viscous flux. Ignoring the boundaries at $x = 1$ and $y = 1$, this results in

$$\int_\Omega U_t\, d\Omega - \int_{\partial\Omega, x=0} F\, dy - \int_{\partial\Omega, y=0} G\, dx = -\int_{\partial\Omega, x=0} w^T \mathsf{f}^{\mathsf{V}}\, dy - \int_{\partial\Omega, y=0} w^T \mathsf{g}^{\mathsf{V}}\, dx - \int_\Omega w_x^T \mathsf{f}^{\mathsf{V}} + w_y^T \mathsf{g}^{\mathsf{V}}\, d\Omega.$$

In view of (62), $F = G = 0$ at the boundaries. Using the temperature condition in (62), the boundary integrals take the form

$$\int_{\partial\Omega, x=0} w^T \mathsf{f}^{\mathsf{V}}\, dy = -\frac{\kappa}{c_v}\int_{\partial\Omega, x=0} \frac{\mathrm{T}_x}{\mathrm{T}}\, dy = 0, \qquad \int_{\partial\Omega, y=0} w^T \mathsf{g}^{\mathsf{V}}\, dx = -\frac{\kappa}{c_v}\int_{\partial\Omega, y=0} \frac{\mathrm{T}_y}{\mathrm{T}}\, dx = 0.$$

Furthermore, by contracting the derivatives of the entropy variables with f^{V} and g^{V}, and using $\lambda = -\frac{2}{3}\mu$ we obtain

$$w_x^T \mathsf{f}^{\mathsf{V}} + w_y^T \mathsf{g}^{\mathsf{V}} = \frac{1}{c_v \mathrm{T}}\left(\tfrac{2}{3}\mu\left(v_{1_x} - v_{2_y}\right)^2 + \tfrac{2}{3}\mu v_{1_x}^2 + \tfrac{2}{3}\mu v_{2_y}^2 + \mu\left(v_{1_y} + v_{2_x}\right)^2\right) + \frac{\kappa}{c_v}\frac{\mathrm{T}_x^2 + \mathrm{T}_y^2}{\mathrm{T}^2} \geq 0,$$

as long as $\mathrm{T} > 0$. Hence, we have proved that $\int_\Omega U_t\, d\Omega \leq 0$.

5.6. Entropy stability for the semi-discrete scheme

For the discretisation in two spatial dimensions, we use the formalism found in e.g. [29]. We divide the spatial domain into $(N+1)(M+1)$ grid points, such that $x_i = ih_x$, $i = 0, 1, \dots N$, where $h_x = 1/N$ and $y_i = ih_y$, $i = 0, 1, \dots, M$ where $h_y = 1/M$. We denote $u_{i,j}(t)$ as the approximation of $u(x_i, y_j, t)$, and the solution vectors are ordered in the following way

$$\boldsymbol{u}^T = (u_{0,0}, u_{1,0}, \dots, u_{N,0}, u_{0,1}, u_{1,1}, \dots, u_{N,1}, \dots, u_{0,M}, u_{1,M}, \dots, u_{N,M}).$$

The 2-D differential operators are defined by Kronecker products as $D_x = I_M \otimes D_N$, where I_M is the $(M+1) \times (M+1)$ identity matrix, and D_N is the $(N+1) \times (N+1)$ (2,1)-SBP operator. Similarly, we have $D_y = D_M \otimes I_N$.

To impose the no-slip boundary conditions by injection, we introduce the initial velocity solution vectors as

$$(\boldsymbol{v}^1)^T = (0, 0, 0, \dots, 0,\ 0, v^1_{1,1}, v^1_{2,1}, \dots, 0,\ 0, v^1_{1,M-1}, v^1_{2,M-1}, \dots, v^1_{N,M-1},\ 0, 0, 0, \dots, 0),$$
$$(\boldsymbol{v}^2)^T = (0, 0, 0, \dots, 0,\ 0, v^2_{1,1}, v^2_{2,1}, \dots, 0,\ 0, v^2_{1,M-1}, v^2_{2,M-1}, \dots, v^2_{N,M-1},\ 0, 0, 0, \dots, 0),$$

i.e., $\boldsymbol{v}^1$ and $\boldsymbol{v}^2$ have all elements along $x = 0$, $x = 1$, $y = 0$ and $y = 1$ set to zero. In addition, we define the SBP-operators for the momentum equations so that they do not act on the boundary nodes.

$$\tilde{D}_x = (\tilde{I}_M \otimes \tilde{D}_N), \quad \tilde{D}_y = (\tilde{D}_M \otimes \tilde{I}_N),$$

where $\tilde{D}_N$ and $\tilde{D}_M$ are Dirichlet-SBP operators corresponding to D_N and D_M, respectively. Moreover, $\tilde{I}_N$ and $\tilde{I}_M$ are almost the identity matrices, but with the upper left and lower right elements set to zero. The norm matrix for the two-dimensional grid is given by $H = H_y \otimes H_x$, where H_y and H_x are equal to the 1-D norms defined in Section 3, with elements of size h_y and h_x and matrix sizes $(M+1) \times (M+1)$ and $(N+1) \times (N+1)$, respectively.

Similarly as for the one-dimensional case, the SBP operators satisfy a discrete product - (39) and quotient (40) rule.

5.7. Entropy stability for the semi-discrete scheme

The 2-D inviscid fluxes are approximated by

$$\mathbf{f}^{\mathrm{I}}_{i+1/2,j} = \frac{\mathbf{f}^{\mathrm{I}}_{i+1,j} + \mathbf{f}^{\mathrm{I}}_{i,j}}{2} - \frac{\delta_{i+1/2,j}\left(\mathbf{u}_{i+1,j} - \mathbf{u}_{i,j}\right)}{2},$$
$$\mathbf{g}^{\mathrm{I}}_{i,j+1/2} = \frac{\mathbf{g}^{\mathrm{I}}_{i,j+1} + \mathbf{g}^{\mathrm{I}}_{i,j}}{2} - \frac{\delta_{i,j+1/2}\left(\mathbf{u}_{i,j+1} - \mathbf{u}_{i,j}\right)}{2},$$

and the convective terms by the flux differences

$$\begin{aligned}(\mathcal{D}^{\mathrm{I}}_x \mathbf{f}^{\mathrm{I}})_{i,j} &= \frac{\mathbf{f}^{\mathrm{I}}_{i+1/2,j} - \mathbf{f}^{\mathrm{I}}_{i-1/2,j}}{h_x}, \quad i = 1, \dots, N-1, \quad j = 0, \dots, M,\\ (\mathcal{D}^{\mathrm{I}}_y \mathbf{g}^{\mathrm{I}})_{i,j} &= \frac{\mathbf{g}^{\mathrm{I}}_{i,j+1/2} - \mathbf{g}^{\mathrm{I}}_{i,j-1/2}}{h_y}, \quad i = 0, \dots, N, \quad j = 1, \dots, M-1,\end{aligned} \tag{63}$$

in the interior and by

$$\begin{aligned}(\mathcal{D}^{\mathrm{I}}_x \mathbf{f}^{\mathrm{I}})_{0,j} &= \frac{\mathbf{f}^{\mathrm{I}}_{1/2,j} - \mathbf{f}^{\mathrm{I}}_{0,j}}{h_x/2}, \quad i = 0, \quad j = 0, \dots, M,\\ (\mathcal{D}^{\mathrm{I}}_y \mathbf{g}^{\mathrm{I}})_{i,0} &= \frac{\mathbf{g}^{\mathrm{I}}_{i,1/2} - \mathbf{g}^{\mathrm{I}}_{i,0}}{h_y/2}, \quad i = 0, \dots, N, \quad j = 0,\end{aligned} \tag{64}$$

at the boundaries $x = 0$, $y = 0$ (once again, we neglect the right and upper boundaries to reduce notation). As in the 1-D case, we have

$$\mathbf{f}^{\mathrm{I},\rho}_{0,j} = (\boldsymbol{\rho}.\boldsymbol{v}^1)_{0,j}, \quad \mathbf{f}^{\mathrm{I},m}_{0,j} = \mathbf{f}^{\mathrm{I},m}_{1/2,j}, \quad \mathbf{f}^{\mathrm{I},n}_{0,j} = \mathbf{f}^{\mathrm{I},n}_{1/2,j}, \quad \mathbf{f}^{\mathrm{I},E}_{0,j} = (\boldsymbol{v}^1.(\boldsymbol{E} + \boldsymbol{p}))_{0,j},$$

and similarly for $\mathbf{g}^{\mathrm{I}}_{i,0}$.

Next, we approximate the viscous terms by

$$\mathcal{D}^{\mathrm{V}}_x \mathbf{f}^{\mu} = \begin{pmatrix} 0 \\ \tilde{D}_x\left(\mathbf{1}_{\boldsymbol{x}}.\left(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\right)\right) \\ \tilde{D}_x\left(\mathbf{1}_{\boldsymbol{x}}.\left(\mu(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\right)\right) \\ D_x\Big(\mathbf{1}_{\boldsymbol{x}}.\Big(\overset{b,x}{\boldsymbol{v}^1}\left(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\right) + \mu \overset{b,x}{\boldsymbol{v}^2}(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\Big)\Big) \end{pmatrix}, \tag{65}$$

$$\mathcal{D}_y^{\mathrm{v}}\mathbf{g}^{\mu} = \begin{pmatrix} 0 \\ \tilde{D}_y\left(\mathbf{1}_{\boldsymbol{y}}.\left(\mu(D_y\boldsymbol{v}^1 + D_x\boldsymbol{v}^2)\right)\right) \\ \tilde{D}_y\left(\mathbf{1}_{\boldsymbol{y}}.\left(2\mu D_y\boldsymbol{v}^2 + \lambda(D_x\boldsymbol{v}^1 + D_y\boldsymbol{v}^2)\right)\right) \\ D_y\left(\mathbf{1}_{\boldsymbol{y}}.\left(\overset{b,y}{\boldsymbol{v}^2}\left(2\mu D_y\boldsymbol{v}^2 + \lambda(D_x\boldsymbol{v}^1 + D_y\boldsymbol{v}^2)\right) + \mu\overset{b,y}{\boldsymbol{v}^1}(D_y\boldsymbol{v}^1 + D_x\boldsymbol{v}^2)\right)\right) \end{pmatrix}, \tag{66}$$

where we have used the approximation of 1 from [31]:

$$[\mathbf{1}_{\boldsymbol{x}}]_{i,j} = \frac{\mathrm{T}^{-1}_{i,j}}{\underset{i,x}{\mathrm{T}^{-1}_{i,j}}}, \qquad [\mathbf{1}_{\boldsymbol{y}}]_{i,j} = \frac{\mathrm{T}^{-1}_{i,j}}{\underset{i,y}{\mathrm{T}^{-1}_{i,j}}}. \tag{67}$$

Note that the operator $\mathcal{D}_{x,y}$ uses $\tilde{D}_{x,y}$ for the momentum equations where the no-slip condition is imposed by injection, and uses $D_{x,y}$ for the continuity equation and the equation for total energy.

Lastly, the approximations of the heat diffusive fluxes are given by

$$\mathcal{D}_x^{\mathrm{v}}\mathbf{f}^{\kappa} = \left(0, 0, 0, \kappa D_x D_x \mathrm{T}\right)^T, \quad \mathcal{D}_y^{\mathrm{v}}\mathbf{g}^{\kappa} = \left(0, 0, 0, \kappa D_y D_y \mathrm{T}\right)^T. \tag{68}$$

Then, a semi-discretisation of (61) is given by

$$\mathbf{u}_t + \mathcal{D}_x^{\mathrm{I}}\mathbf{f}^{\mathrm{I}} + \mathcal{D}_y^{\mathrm{I}}\mathbf{g}^{\mathrm{I}} = \mathcal{D}_x^{\mathrm{v}}\mathbf{f}^{\mu} + \mathcal{D}_x^{\mathrm{v}}\mathbf{f}^{\kappa} + \mathcal{D}_y^{\mathrm{v}}\mathbf{g}^{\mu} + \mathcal{D}_y^{\mathrm{v}}\mathbf{g}^{\kappa} + \mathsf{SAT}, \tag{69}$$

with $\mathsf{SAT} = \left(0, 0, 0, -\kappa\left((I_M \otimes H_x^{-1}B)(D_x\mathrm{T} - 0) + (H_y^{-1}B \otimes I_N)(D_y\mathrm{T} - 0)\right)\right)^T$.

Remark. Our scheme resembles the ones proposed in [31], where the no-slip condition was imposed using SAT, and [25].

Proposition 5.3. *The 2D semi-discrete scheme (69) approximating the problem (61) is entropy stable.*

Before stating the proof, we prove several lemmas to simplify the presentation. Similarly as in the proof of Proposition (5.2), we perform the calculations for the convective terms using index notation. To this end, we define $\boldsymbol{w}_{i,j}^T = \frac{1}{c_v \mathrm{T}_{i,j}}\left(\frac{(v^1_{i,j})^2 + (v^2_{i,j})^2}{2} + c_v\mathrm{T}_{i,j}(\mathcal{S}_{i,j} - \gamma), -v^1_{i,j}, -v^2_{i,j}, 1\right)$.

Lemma 5.4. *The convective flux approximations (63) and (64) satisfy*

$$\sum_{i,j=0}^{N,M} \boldsymbol{w}_{i,j}^T H_k(\mathcal{D}_x^{\mathrm{I}}\mathbf{f}^{\mathrm{I}})_{i,j} + \sum_{i,j=0}^{N,M} \boldsymbol{w}_{i,j}^T H_k(\mathcal{D}_y^{\mathrm{I}}\mathbf{g}^{\mathrm{I}})_{i,j} \geq 0, \tag{70}$$

(where $k = (j(N+1) + i)$ and H_k denotes the diagonal elements of H).

Proof. (70) follows by applying the same technique as for A in (55) to all j's in the x-direction for $\mathbf{f}^{\mathrm{I}}$ and to all i's in the y-direction for $\mathbf{g}^{\mathrm{I}}$. □

For the diffusive terms, we define $\mathbf{H} = \mathrm{diag}(H, H, H, H)^T$ and $\boldsymbol{w}^T = \left((\boldsymbol{w}^{\rho})^T, (\boldsymbol{w}^m)^T, (\boldsymbol{w}^n)^T, (\boldsymbol{w}^E)^T\right)^T$ where

$$\begin{aligned} \boldsymbol{w}^{\rho} &= \tfrac{1}{c_v}\mathrm{T}^{-1}.\left(\frac{\boldsymbol{v}^1.\boldsymbol{v}^1 + \boldsymbol{v}^2.\boldsymbol{v}^2}{2} + c_v\mathrm{T}^{-1}.(\mathcal{S} - \gamma)\right), \\ \boldsymbol{w}^m &= -\tfrac{1}{c_v}\mathrm{T}^{-1}.\boldsymbol{v}^1, \\ \boldsymbol{w}^n &= -\tfrac{1}{c_v}\mathrm{T}^{-1}.\boldsymbol{v}^2, \\ \boldsymbol{w}^E &= \tfrac{1}{c_v}\mathrm{T}^{-1}, \end{aligned} \tag{71}$$

and $[\mathrm{T}^{-1}]_{i,j} = \frac{1}{\mathrm{T}_{i,j}}$. (Recall that the dot product is the component wise vector multiplication.)

Lemma 5.5. *Contracting the entropy variables with the viscous fluxes, we obtain*

$$\boldsymbol{w}^T\mathbf{H}(\mathcal{D}_x^{\mathrm{v}}\mathbf{f}^{\mu} + \mathcal{D}_y^{\mathrm{v}}\mathbf{g}^{\mu}) \leq 0. \tag{72}$$

Proof. Consider the viscous flux in the x-direction given by (65). Contracting the vector (65) by the entropy variables in (71) and $\mathbf{H}$, we obtain

$$\begin{aligned}
A_1 &= \boldsymbol{w}^T \mathbf{H} \mathcal{D}_x^{\mathrm{v}} \mathbf{f}^{\mu}, \\
&= (\boldsymbol{w}^m)^T H \tilde{D}_x \big(\mathbf{1}_x.\big(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\big)\big) + (\boldsymbol{w}^n)^T H \tilde{D}_x \big(\mathbf{1}_x.\big(\mu(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\big)\big) \\
&\quad + (\boldsymbol{w}^E)^T H D_x \big(\mathbf{1}_x.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\big) + \mu \overset{b,x}{\boldsymbol{v}^2}.(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\big)\big).
\end{aligned}$$

Utilise that $H\tilde{D}_x = \tilde{H}_y \otimes \tilde{B}_N - \tilde{H}_y \otimes (H_x \tilde{D}_N)^T = \tilde{H}_y \otimes \tilde{B}_N - \tilde{D}_x^T H$, and the analogous properties of D_x, to obtain

$$\begin{aligned}
A_1 = \underbrace{(\boldsymbol{w}^m)^T (\tilde{H}_y \otimes \tilde{B}_N)\big(\mathbf{1}_x.\big(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\big)\big) + (\boldsymbol{w}^n)^T (\tilde{H}_y \otimes \tilde{B}_N)\big(\mathbf{1}_x.\big(\mu(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\big)\big)}_{\mathcal{B}_1} \\
-(\tilde{D}_x \boldsymbol{w}^m)^T H \mathbf{1}_x.\big(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\big) - (\tilde{D}_x \boldsymbol{w}^n)^T H \big(\mu(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\big) \\
+(\boldsymbol{w}^E)^T (H_y \otimes B_N)\big(\mathbf{1}_x.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\big) + \mu \overset{b,x}{\boldsymbol{v}^2}.(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\big)\big) \\
-(D_x \boldsymbol{w}^E)^T H \mathbf{1}_x.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)\big) + \mu \overset{b,x}{\boldsymbol{v}^2}.(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)\big).
\end{aligned}$$

Consider the boundary terms, $\mathcal{B}_1$. Using the result (B.1) obtained in Appendix B, we find that

$$\begin{aligned}
\mathcal{B}_1 = &\underbrace{(2\mu+\lambda) h_y (\boldsymbol{w}^m)^T \big(\underline{\mathbf{0}},\quad \tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^1)_{i,1},\quad \tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^1)_{i,2},\quad \ldots\quad \tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^1)_{i,M-1},\quad \underline{\mathbf{0}}\big)^T}_{\mathcal{B}_{1,1}} \\
&\underbrace{+\lambda h_y (\boldsymbol{w}^m)^T \big(\underline{\mathbf{0}},\quad \tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^2)_{i,1},\quad \tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^2)_{i,2},\quad \ldots\quad \tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^2)_{i,M-1},\quad \underline{\mathbf{0}}\big)^T}_{\mathcal{B}_{1,2}} \\
&\underbrace{+\mu h_y (\boldsymbol{w}^n)^T \big(\underline{\mathbf{0}},\quad \tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^1)_{i,1},\quad \tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^1)_{i,2},\quad \ldots\quad \tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^1)_{i,M-1},\quad \underline{\mathbf{0}}\big)^T}_{\mathcal{B}_{1,3}} \\
&\underbrace{+\mu h_y (\boldsymbol{w}^n)^T \big(\underline{\mathbf{0}},\quad \tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^2)_{i,1},\quad \tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^2)_{i,2},\quad \ldots\quad \tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^2)_{i,M-1},\quad \underline{\mathbf{0}}\big)^T}_{\mathcal{B}_{1,4}}.
\end{aligned}$$

Consider $\mathcal{B}_{1,1} + \mathcal{B}_{1,2}$. As they depend on the same component of the entropy variables, the terms can be rewritten as

$$\mathcal{B}_{1,1} + \mathcal{B}_{1,2} = (2\mu+\lambda) h_y \sum_{j=1}^{M-1} (\boldsymbol{w}^m)_{i,j}^T (\tilde{B}_N(\mathbf{1}_x.D_x \boldsymbol{v}^1)_{i,j}) + \lambda h_y \sum_{j=1}^{M-1} (\boldsymbol{w}^m)_{i,j}^T (\tilde{B}_N(\mathbf{1}_x.D_y \boldsymbol{v}^2)_{i,j}).$$

Consider an arbitrary node $j \neq \{0, M\}$, and neglect the parameters. Then we have

$$\begin{aligned}
\mathcal{B}_{1,1} + \mathcal{B}_{1,2} = &\begin{pmatrix} w_{0,j}^m \\ w_{1,j}^m \\ \vdots \\ w_{N-1,j}^m \\ w_{N,j}^m \end{pmatrix}^T \begin{pmatrix} 0 & -\frac{1}{2} & 0 & 0 & 0 & \ldots & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & \ldots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \ldots & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \ldots & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{0,j} \\ (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{1,j} \\ \vdots \\ (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{N-1,j} \\ (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{N,j} \end{pmatrix} \\
&+ \begin{pmatrix} w_{0,j}^m \\ w_{1,j}^m \\ \vdots \\ w_{N-1,j}^m \\ w_{N,j}^m \end{pmatrix}^T \begin{pmatrix} 0 & -\frac{1}{2} & 0 & 0 & 0 & \ldots & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & \ldots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \ldots & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \ldots & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{0,j} \\ (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{1,j} \\ \vdots \\ (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{N-1,j} \\ (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{N,j} \end{pmatrix}, \\
= &\frac{1}{2}\Big(-w_{1,j}^m (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{0,j} - w_{0,j}^m (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{1,j} + w_{N,j}^m (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{N-1,j} + w_{N-1,j}^m (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{N,j}\Big) \\
&+\frac{1}{2}\Big(-w_{1,j}^m (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{0,j} - w_{0,j}^m (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{1,j} + w_{N,j}^m (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{N-1,j} + w_{N-1,j}^m (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{N,j}\Big).
\end{aligned}$$

Since $w_{0,j}^m = v_{0,j}^1 = 0$ and $w_{N,j}^m = v_{N,j}^1 = 0$, due to the no-slip condition, this reduces to

$$\mathcal{B}_{1,1} + \mathcal{B}_{1,2} = \frac{1}{2}\Big(-w_{1,j}^m (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{0,j} + w_{N-1,j}^m (\mathbf{1}_x.D_x \boldsymbol{v}^1)_{N,j} - w_{1,j}^m (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{0,j} + w_{N-1,j}^m (\mathbf{1}_x.D_y \boldsymbol{v}^2)_{N,j}\Big).$$

Next, we insert the specific form of the derivatives, which gives us

$$\begin{aligned}\mathcal{B}_{1,1}+\mathcal{B}_{1,2}=&\frac{1}{2}\Bigg(-w^m_{1,j}1_{x_{0,j}}\frac{v^1_{1,j}-v^1_{0,j}}{h_x}+w^m_{N-1,j}1_{x_{N,j}}\frac{v^1_{N,j}-v^1_{N-1,j}}{h_x}\\&-w^m_{1,j}1_{x_{0,j}}\frac{v^1_{0,j+1}-v^2_{0,j-1}}{2h_y}+w^m_{N-1,j}1_{x_{N,j}}\frac{v^2_{N,j+1}-v^2_{N,j-1}}{2h_y}\Bigg).\end{aligned}$$

Using the no-slip condition yet again, and inserting the specific expression of the entropy variable, we have

$$\mathcal{B}_{1,1}+\mathcal{B}_{1,2}=-\frac{1}{2h_x}\left(1_{x_{0,j}}\frac{(v^1_{1,j})^2}{\mathrm{T}_{1,j}}+1_{x_{N,j}}\frac{(v^1_{N-1,j})^2}{\mathrm{T}_{N-1,j}}\right)\leq 0,\qquad (\mathrm{T}_{i,j}>0).$$

By analogous manipulations to $\mathcal{B}_{1,3}+\mathcal{B}_{1,4}$, A_1 reduces to

$$\begin{aligned}A_1\leq&-(\tilde{D}_x\boldsymbol{w}^m)^TH\mathbf{1}_{\boldsymbol{x}}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)-(\tilde{D}_x\boldsymbol{w}^n)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\\&\underbrace{+(\boldsymbol{w}^E)^T(H_y\otimes B_N)\big(\mathbf{1}_{\boldsymbol{x}}.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)+\mu\overset{b,x}{\boldsymbol{v}^2}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\big)}_{\mathcal{B}_2}\\&-(D_x\boldsymbol{w}^E)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)+\mu\overset{b,x}{\boldsymbol{v}^2}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big).\end{aligned}$$

The boundary term, $\mathcal{B}_2$, is produced by the (2,1)-SBP operator, and from the SBP-properties in Section 3, we know it will extract boundary terms (in contrast to the $\tilde{B}$s, which extract terms along the boundaries *and* the neighbouring nodes). Since $\overset{b,x}{\boldsymbol{v}^1}=\overset{b,x}{\boldsymbol{v}^2}=0$ at the boundaries (see (37)), it follows that all boundary terms vanish. The resulting estimate is therefore

$$\begin{aligned}A_1\leq&-(\tilde{D}_x\boldsymbol{w}^m)^TH\mathbf{1}_{\boldsymbol{x}}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)-(\tilde{D}_x\boldsymbol{w}^n)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\\&-(D_x\boldsymbol{w}^E)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)+\mu\overset{b,x}{\boldsymbol{v}^2}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big).\end{aligned}\tag{73}$$

Similarly for the viscous flux in the y-direction, we multiply (66) by the entropy variables and the norm matrix, H, to end up with

$$\begin{aligned}A_2\leq&-(\tilde{D}_y\boldsymbol{w}^m)^TH\mathbf{1}_{\boldsymbol{y}}.(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2))-(\tilde{D}_y\boldsymbol{w}^n)^TH\mathbf{1}_{\boldsymbol{y}}.(2\mu D_y\boldsymbol{v}^2)+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\\&-(D_y\boldsymbol{w}^E)^TH\mathbf{1}_{\boldsymbol{y}}.\big(\overset{b,y}{\boldsymbol{v}^2}.(2\mu D_y\boldsymbol{v}^2+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2))+\mu\overset{b,y}{\boldsymbol{v}^1}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\end{aligned}\tag{74}$$

Combining (73) and (74), we obtain

$$\begin{aligned}\boldsymbol{w}^T\mathbf{H}\mathcal{D}^{\mathrm{v}}_x\mathbf{f}^\mu+\boldsymbol{w}^T\mathbf{H}\mathcal{D}^{\mathrm{v}}_y\mathbf{g}^\mu\leq&-(\tilde{D}_x\boldsymbol{w}^m)^TH\mathbf{1}_{\boldsymbol{x}}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)-(\tilde{D}_x\boldsymbol{w}^n)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\\&-(D_x\boldsymbol{w}^E)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)+\mu\overset{b,x}{\boldsymbol{v}^2}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\\&-(\tilde{D}_y\boldsymbol{w}^m)^TH\mathbf{1}_{\boldsymbol{y}}.(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2))-(\tilde{D}_y w^n)^TH\mathbf{1}_{\boldsymbol{y}}.(2\mu D_y\boldsymbol{v}^2)+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\\&-(D_y\boldsymbol{w}^E)^TH\mathbf{1}_{\boldsymbol{y}}.\big(\overset{b,y}{\boldsymbol{v}^2}.(2\mu D_y\boldsymbol{v}^2+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2))+\mu\overset{b,y}{\boldsymbol{v}^1}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big).\end{aligned}\tag{75}$$

To recast (75) as a quadratic form, we use the entropy variables (71) and utilise that the derivative operators satisfy the discrete product rule (39). Then,

$$\begin{aligned}\boldsymbol{w}^T\mathbf{H}(\mathcal{D}^{\mathrm{v}}_x\mathbf{f}^\mu+\mathcal{D}^{\mathrm{v}}_y\mathbf{g}^\mu)\leq&-\frac{1}{c_v}\left(\underline{\big(\overset{b,x}{\boldsymbol{v}^1}.\tilde{D}_x\mathrm{T}^{-1}}+\overset{i,x}{\mathrm{T}^{-1}}.\tilde{D}_x\boldsymbol{v}^1\big)^TH\mathbf{1}_{\boldsymbol{x}}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)\right)\\&-\frac{1}{c_v}\left(\underline{\big(\overset{b,x}{\boldsymbol{v}^2}.\tilde{D}_x\mathrm{T}^{-1}}+\overset{i,x}{\mathrm{T}^{-1}}.\tilde{D}_x\boldsymbol{v}^2\big)^TH\mathbf{1}_{\boldsymbol{x}}.\big(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\right)\\&+\frac{1}{c_v}\left((D_x\mathrm{T}^{-1})^TH\mathbf{1}_{\boldsymbol{x}}.\big(\underline{\overset{b,x}{\boldsymbol{v}^1}.\big(2\mu D_x\boldsymbol{v}^1+\lambda(D_x\boldsymbol{v}^1+D_y\boldsymbol{v}^2)\big)}+\underline{\mu\overset{b,x}{\boldsymbol{v}^2}.(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)}\big)\right)\\&-\frac{1}{c_v}\left(\underline{\big(\overset{b,y}{\boldsymbol{v}^1}.\tilde{D}_y\mathrm{T}^{-1}}+\overset{i,y}{\mathrm{T}^{-1}}.\tilde{D}_y\boldsymbol{v}^1\big)^TH\mathbf{1}_{\boldsymbol{y}}.\big(\mu(D_y\boldsymbol{v}^1+D_x\boldsymbol{v}^2)\big)\right)\end{aligned}$$

$$
\begin{aligned}
&- \frac{1}{c_v}\left(\underline{(\overset{b,y}{\boldsymbol{v}^2}.\tilde{D}_y \mathrm{T}^{-1}} + \overset{i,y}{\mathrm{T}^{-1}}.\tilde{D}_y \boldsymbol{v}^2)^T H \mathbf{1}_{\boldsymbol{y}}.(2\mu D_y \boldsymbol{v}^2 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)) \right) \\
&+ \frac{1}{c_v}\left((D_y \mathrm{T}^{-1})^T H \mathbf{1}_{\boldsymbol{y}}.(\underline{\overset{b,y}{\boldsymbol{v}^2}.(2\mu D_y \boldsymbol{v}^2 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2))} + \underline{\mu \overset{b,y}{\boldsymbol{v}^1}.(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)}) \right).
\end{aligned}
$$

A number of terms cancel (see the colour code above), and we end up with

$$
\begin{aligned}
\boldsymbol{w}^T \mathbf{H}(\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu + \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu) \leq &- \frac{1}{c_v}\left((\overset{i,x}{\mathrm{T}^{-1}}.\tilde{D}_x \boldsymbol{v}^1)^T H \mathbf{1}_{\boldsymbol{x}}.(2\mu D_x \boldsymbol{v}^1 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)) + (\overset{i,x}{\mathrm{T}^{-1}}.\tilde{D}_x \boldsymbol{v}^2)^T H \mathbf{1}_{\boldsymbol{x}}.(\mu(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^2)) \right) \\
&- \frac{1}{c_v}\left((\overset{i,y}{\mathrm{T}^{-1}}.\tilde{D}_y \boldsymbol{v}^1)^T H \mathbf{1}_{\boldsymbol{y}}.(\mu(D_y \boldsymbol{v}^1 + D_x \boldsymbol{v}^1)) + (\overset{i,y}{\mathrm{T}^{-1}}.\tilde{D}_y \boldsymbol{v}^2)^T H \mathbf{1}_{\boldsymbol{y}}.(2\mu D_y \boldsymbol{v}^2 + \lambda(D_x \boldsymbol{v}^1 + D_y \boldsymbol{v}^2)) \right).
\end{aligned}
$$

Use the form of $\mathbf{1}_{\boldsymbol{x}}$ and $\mathbf{1}_{\boldsymbol{y}}$ from (67) and Stokes' hypothesis, $\lambda = -\frac{2}{3}\mu$, to obtain

$$
\begin{aligned}
\boldsymbol{w}^T \mathbf{H}(\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu + \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu) \leq - \frac{\mu}{c_v}\Big(&\tfrac{4}{3}(\tilde{D}_x \boldsymbol{v}^1)^T H(\mathrm{T}^{-1}.D_x \boldsymbol{v}^1) - \tfrac{2}{3}(\tilde{D}_x \boldsymbol{v}^1)^T H(\mathrm{T}^{-1}.D_y \boldsymbol{v}^2) \\
&+ (\tilde{D}_x \boldsymbol{v}^2)^T H(\mathrm{T}^{-1}.D_y \boldsymbol{v}^1) + (\tilde{D}_x \boldsymbol{v}^2)^T H(\mathrm{T}^{-1}.D_x \boldsymbol{v}^2) \\
&+ (\tilde{D}_y \boldsymbol{v}^1)^T H(\mathrm{T}^{-1}.D_y \boldsymbol{v}^1) + (\mathrm{T}^{-1}.\tilde{D}_y \boldsymbol{v}^1)^T H(D_x \boldsymbol{v}^2) \\
&- \tfrac{2}{3}(\tilde{D}_y \boldsymbol{v}^2)^T H(\mathrm{T}^{-1}.D_x \boldsymbol{v}^1) + \tfrac{4}{3}(\tilde{D}_y \boldsymbol{v}^2)^T H(\mathrm{T}^{-1}.D_y \boldsymbol{v}^2)\Big),
\end{aligned}
$$

which can be further rearranged into

$$
\begin{aligned}
\boldsymbol{w}^T \mathbf{H}(\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu + \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu) \leq - \frac{\mu}{c_v}\Bigg(&\tfrac{2}{3}\begin{pmatrix} (\tilde{D}_x \boldsymbol{v}^1)^T \\ (\tilde{D}_y \boldsymbol{v}^2)^T \end{pmatrix}^T \begin{pmatrix} H & -H \\ -H & H \end{pmatrix} \begin{pmatrix} \mathrm{T}^{-1}.D_x \boldsymbol{v}^1 \\ \mathrm{T}^{-1}.D_y \boldsymbol{v}^2 \end{pmatrix} + \begin{pmatrix} (\tilde{D}_x \boldsymbol{v}^2)^T \\ (\tilde{D}_y \boldsymbol{v}^1)^T \end{pmatrix}^T \begin{pmatrix} H & H \\ H & H \end{pmatrix} \begin{pmatrix} \mathrm{T}^{-1}.D_x \boldsymbol{v}^2 \\ \mathrm{T}^{-1}.D_y \boldsymbol{v}^1 \end{pmatrix} \\
&+ \tfrac{2}{3}\|\sqrt{\mathrm{T}^{-1}}.\tilde{D}_x \boldsymbol{v}^1\|_H^2 + \tfrac{2}{3}\|\sqrt{\mathrm{T}^{-1}}.\tilde{D}_y \boldsymbol{v}^2\|_H^2\Bigg) \leq 0, \qquad (\mathrm{T} > 0)
\end{aligned}
$$

i.e., (72) holds true. □

Lemma 5.6. *The diffusive heat fluxes* (68) *satisfy*

$$
\boldsymbol{w}^T \mathbf{H} \mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\kappa + \boldsymbol{w}^T \mathbf{H} \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\kappa + \boldsymbol{w}^T \mathbf{H} \mathbf{SAT} \leq 0. \tag{76}
$$

Proof. Denote the left-hand side of (76) by A, then

$$
\begin{aligned}
A &= \kappa\left((\boldsymbol{w}^E)^T H D_x D_x \mathrm{T} + (\boldsymbol{w}^E)^T H D_y D_y \mathrm{T} \right) + (\boldsymbol{w}^E)^T \mathbf{H} \mathbf{SAT}, \\
&= \kappa\left((\boldsymbol{w}^E)^T (H_y \otimes H_x)(I_y \otimes D_N) D_x \mathrm{T} + (\boldsymbol{w}^E)^T (H_x \otimes H_y)(D_M \otimes I_x) D_y \mathrm{T} \right) + (\boldsymbol{w}^E)^T \mathbf{H} \mathbf{SAT}.
\end{aligned}
$$

This can be stated more compactly as

$$
A = \kappa\left((\boldsymbol{w}^E)^T (H_y \otimes Q_N) D_x \mathrm{T} + (\boldsymbol{w}^E)^T (Q_M \otimes H_x) D_y \mathrm{T} \right) + (\boldsymbol{w}^E)^T \mathbf{H} \mathbf{SAT}.
$$

Utilising that $Q = B - Q^T$, we obtain

$$
\begin{aligned}
A = \kappa\Bigg(&\underbrace{-(\boldsymbol{w}^E)^T (H_y \otimes Q_N^T) D_x \mathrm{T} - (\boldsymbol{w}^E)^T (Q_M^T \otimes H_x) D_y \mathrm{T}}_{A_1} \\
&\underbrace{+ (\boldsymbol{w}^E)^T (H_y \otimes B_N) D_x \mathrm{T} + (\boldsymbol{w}^E)^T (B_M \otimes H_x) D_y \mathrm{T}\Bigg)}_{A_2} + (\boldsymbol{w}^E)^T \mathbf{H} \mathbf{SAT}.
\end{aligned}
$$

Manipulations of A_1 give us

$$A_1 = -\kappa \left((\boldsymbol{w}^E)^T (I_y \otimes D_N)^T (H_y \otimes H_x) D_x \mathrm{T} + (\boldsymbol{w}^E)^T (D_M \otimes I_x)^T (H_y \otimes H_x) D_y \mathrm{T} \right),$$
$$= -\kappa \left((D_x \boldsymbol{w}^E)^T H D_x \mathrm{T} + (D_y \boldsymbol{w}^E)^T H D_y \mathrm{T} \right).$$

Recall that $\boldsymbol{w}^E = -\frac{1}{c_v \mathrm{T}}$, such that, by using the discrete quotient rule (40), the above is equivalent to

$$A_1 = -\frac{\kappa}{c_v} \left(((\overset{x}{\mathrm{T}^2})^{-1} . D_x \mathrm{T})^T H D_x \mathrm{T} + ((\overset{y}{\mathrm{T}^2})^{-1} . D_y \mathrm{T})^T H D_y \mathrm{T} \right) = -\frac{\kappa}{c_v} \left(\| \sqrt{(\overset{x}{\mathrm{T}^2})^{-1}} . D_x \mathrm{T} \|_H^2 + \| \sqrt{(\overset{y}{\mathrm{T}^2})^{-1}} . D_y \mathrm{T} \|_H^2 \right),$$

where $(\overset{x}{\mathrm{T}^2})^{-1}$ and $(\overset{y}{\mathrm{T}^2})^{-1}$ are vectors containing the coefficients produced by the quotient rule (40).

Next, consider the boundary terms A_2, and insert the specific form of the SAT:

$$A_2 = \kappa \left((\boldsymbol{w}^E)^T (H_y \otimes B_N) D_x \mathrm{T} + (\boldsymbol{w}^E)^T (B_M \otimes H_x) D_y \mathrm{T} \right)$$
$$- \kappa (\boldsymbol{w}^E)^T (H_y \otimes H_x) \left((I_M \otimes H_x^{-1} B_N)(D_x \mathrm{T}) + (H_y^{-1} B_M \otimes I_N)(D_y \mathrm{T}) \right),$$
$$= \kappa (\boldsymbol{w}^E)^T \left((H_y \otimes B_N) D_x \mathrm{T} + (B_M \otimes H_x) D_y \mathrm{T} - (H_y \otimes B_N) D_x \mathrm{T} - (B_M \otimes H_x) D_y \mathrm{T} \right) = 0.$$

Hence, we have

$$\boldsymbol{w}^T \mathbf{H} \mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\kappa + \boldsymbol{w}^T \mathbf{H} \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\kappa + \boldsymbol{w}^T \mathbf{H} \mathsf{SAT} = -\frac{\kappa}{c_v} \left(\| \sqrt{(\overset{x}{\mathrm{T}^2})^{-1}} . D_x \mathrm{T} \|_H^2 + \| \sqrt{(\overset{y}{\mathrm{T}^2})^{-1}} . D_y \mathrm{T} \|_H^2 \right) \leq 0. \quad \square$$

Proposition 5.7. *The semi-discrete scheme* (69) *is entropy stable in the sense of* (52).

Proof. Contract (69) with $\boldsymbol{w}_{i,j}^T = -\frac{1}{c_v \mathrm{T}_{i,j}} \left(\frac{(v_{i,j}^1)^2 + (v_{i,j}^2)^2}{2} + c_v \mathrm{T}_{i,j} (\mathcal{S}_{i,j} - \gamma), -v_{i,j}^1, -v_{i,j}^2, 1 \right)$, and the corresponding diagonal norm matrix element, H_k, $(k = (j(N+1) + i))$. Then sum over all grid points:

$$\sum_{i,j=0}^{N,M} \boldsymbol{w}_{i,j}^T H_k (\mathbf{u}_{i,j})_t + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{l}} \mathbf{f}^{\mathrm{l}})_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{l}} \mathbf{g}^{\mathrm{l}})_{i,j},$$
$$= \sum_{i,j=0}^{N,M} \left(\boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\kappa)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\kappa)_{i,j} + \boldsymbol{w}_{i,j}^T H_k \mathsf{SAT}_{i,j} \right).$$

By Lemma 5.4 the inviscid flux approximations on the left-hand side have been demonstrated to be entropy stable, hence we have

$$\sum_{i,j=0}^{N,M} H_k (U_{i,j})_t \leq \sum_{i,j=0}^{N,M} \left(\boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\kappa)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\kappa)_{i,j} + \boldsymbol{w}_{i,j}^T H_k \mathsf{SAT}_{i,j} \right).$$

Note that the sum on the right-hand side is equivalent to the matrix multiplications:

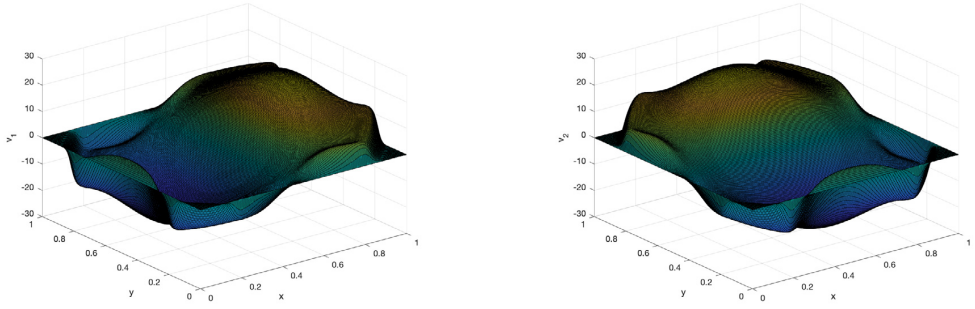
$$\sum_{i,j=0}^{N,M} \left(\boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\kappa)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu)_{i,j} + \boldsymbol{w}_{i,j}^T H_k (\mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\kappa)_{i,j} + \boldsymbol{w}_{i,j}^T H_k \mathsf{SAT}_{i,j} \right),$$
$$= \boldsymbol{w}^T \mathbf{H} \mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\mu + \boldsymbol{w}^T \mathbf{H} \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\mu + \boldsymbol{w}^T \mathbf{H} \mathcal{D}_x^{\mathrm{v}} \mathbf{f}^\kappa + \boldsymbol{w}^T \mathbf{H} \mathcal{D}_y^{\mathrm{v}} \mathbf{g}^\kappa + \boldsymbol{w}^T \mathbf{H} \mathsf{SAT},$$

such that we can utilise the results of Lemma 5.5 and 5.6, and obtain

$$\sum_{i,j=0}^{N,M} H_k (U_{i,j})_t \leq 0. \quad \square$$

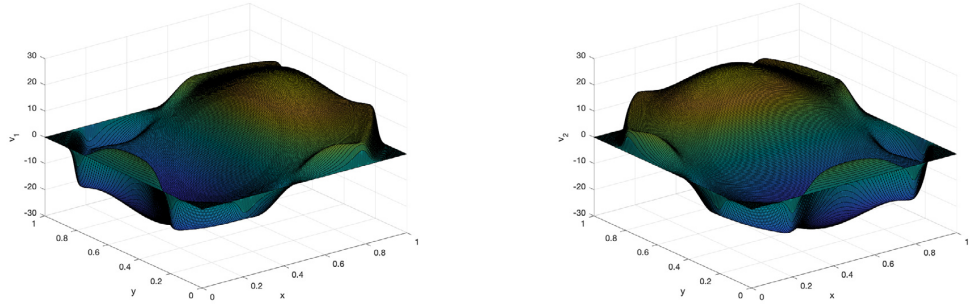
6. Numerical simulations

To demonstrate the properties of the schemes with special emphasis on the no-slip condition, we consider both a subsonic and a supersonic case.



(a) (b)

Fig. 1. (a) v_1 at $t = 0.01$ obtained with 257^2 grid points and $\alpha = 1$. (b) v_2 at $t = 0.01$ obtained with 257^2 grid points and $\alpha = 1$.



(a) (b)

Fig. 2. (a) v_1 at $t = 0.01$ obtained with 257^2 grid points and $\alpha = 0.4$. (b) v_2 at $t = 0.01$ obtained with 257^2 grid points and $\alpha = 0.4$.

6.1. Blast wave

Let $\Omega = [0, 1] \times [0, 1]$, with homogeneous no-slip condition at all walls. We use a similar setup as in [26] with the following initial conditions

$$\rho = 1, \qquad v_1 = 0, v_2 = 0, \qquad p = \begin{cases} 0.01, & \text{if } (x, y) \in \Omega \setminus \mathcal{B}((0.5, 0.5), 0.35), \\ 1000, & \text{if } (x, y) \in \mathcal{B}((0.5, 0.5), 0.35), \end{cases}$$

where $\mathcal{B}((0.5, 0.5), 0.35)$ denotes a disk centred at $(x, y) = (0.5, 0.5)$ with radius $r = 0.35$. Furthermore, we use the following parameters

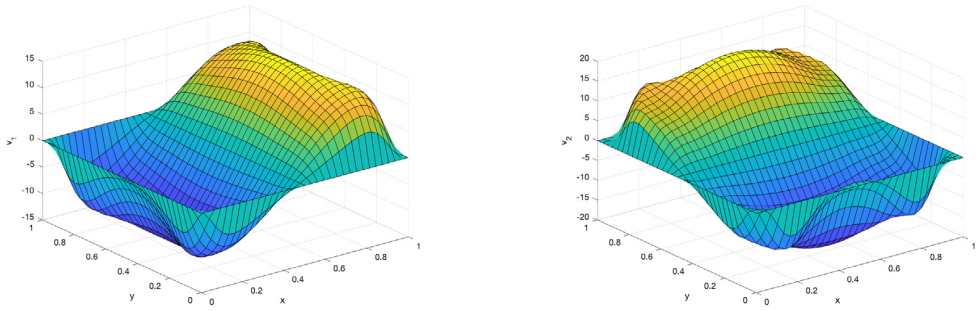
$$\gamma = 1.4, \qquad \mathcal{R} = 286.84, \qquad \mu = 0.1, \qquad \mathrm{Pr} = 0.72, \qquad c_p = 1005, \qquad \kappa = \frac{\mu c_p}{\mathrm{Pr}}.$$

We use (69) with, $\delta_{i+1/2,j} = \alpha \max\left(\left|v^1_{i,j}\right| + c_{i,j}, \left|v^1_{i+1,j}\right| + c_{i+1,j}\right)$. For $\alpha = 1$, this is the entropy stable local Lax-Friedrichs scheme, but to stress test the scheme we also run the non-provably entropy stable choice $\alpha = 0.4$. For time discretisation, we apply the third-order strong stability preserving Runge-Kutta method (see [7]).

The entropy-stable numerical results computed with 257^2 grid points and $\alpha = 1$ are displayed in Fig. 1a and 1b at time $t = 0.01$. Fig. 2a and 2b display the numerical results obtained for the same problem, but with reduced artificial diffusion, $\alpha = 0.4$.

Lastly, we have run a simulation on a coarse mesh (33^2 grid points) as a further demonstration of the robustness of the boundary treatment. The results for the velocity components are displayed in Fig. 3a and 3b.

We have furthermore compared the entropy decay for the cases $\alpha = 1$ and $\alpha = 0.4$ (for the coarse mesh to highlight the differences). The plot of the total entropy, i.e., $\int_\Omega U(u)\, d\Omega$ is depicted in Fig. 4. We have normalised the entropy at every time step by subtracting the initial entropy, $\int_\Omega U(u)\, d\Omega\big|_{t=0}$. As we see from the plot, the entropy is decaying for both values of α, but the decay is faster for larger diffusion, which is as expected.



(a) (b)

Fig. 3. (a) v_1 at $t = 0.01$ obtained with 33^2 grid points and $\alpha = 1$. (b) v_2 at $t = 0.01$ obtained with 33^2 grid points and $\alpha = 0.4$.

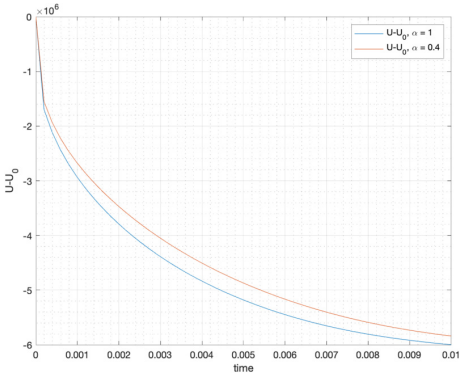


Fig. 4. Plot of the (normalised) total entropy $\int_\Omega U(u)\,d\Omega|_{t=t} - \int_\Omega U(u)\,d\Omega|_{t=0}$ for the coarse grid with $\alpha = 1$ and $\alpha = 0.4$.

6.2. Lid-driven cavity flow

We have run a similar problem as in [3] on the spatial domain $\Omega = [0, 1] \times [0, 1]$. The upper wall of the cavity is moving at a constant speed to the right, such that the boundary conditions for the velocity components become

$$\begin{cases} v_1 = 1, v_2 = 0, & \partial\Omega \cap \{y = 1\}, \\ v_{1,2} = 0, & \partial\Omega \setminus \{y = 1\}. \end{cases} \tag{77}$$

The boundary condition for the temperature is given by (62). Furthermore, the problem parameters are given by $Re = 100$, $Ma = 0.1$, $\Pr = 0.72$, $\gamma = 1.4$, and it is initialised by the conditions

$$\rho = 1, \qquad v_1, v_2 = 0, \qquad p = \frac{1}{Ma^2\gamma}.$$

Note that at one wall, the lid-driven cavity problem has a non-homogeneous no-slip condition for one of the velocity components. Still, an entropy bound for the continuum solution is obtained as only the normal components of the velocities enter the estimate. (We have not been able to prove entropy stability for the discrete scheme with the boundary conditions (77).)

Fig. 5a shows the solution at $t = 2$ when using the scheme (69). We have also run the same problem using a 3rd-order scheme. (See remark at the end of Section 5.1.) We have verified linear stability in one spatial dimension, and the extension to two dimensions is straightforward. The 3rd-order numerical solution is shown in Fig. 5b.

Fig. 6a shows the solution at $t = 2$ for the lid-driven cavity flow with the heat-entropy flow boundary condition $\frac{\kappa}{T}\frac{\partial T}{\partial n} = g = 2$. The solution is qualitatively similar to the adiabatic case where $\frac{\partial T}{\partial n} = 0$. Fig. 6b depicts the total entropy $\int_\Omega U(u)\,d\Omega$, normalised by subtracting the initial entropy $\int_\Omega U(u)|_0\,d\Omega$. We note that the entropy increases initially. This does not violate the entropy inequality since the system is not closed; there is a heat-entropy flow through the boundary.

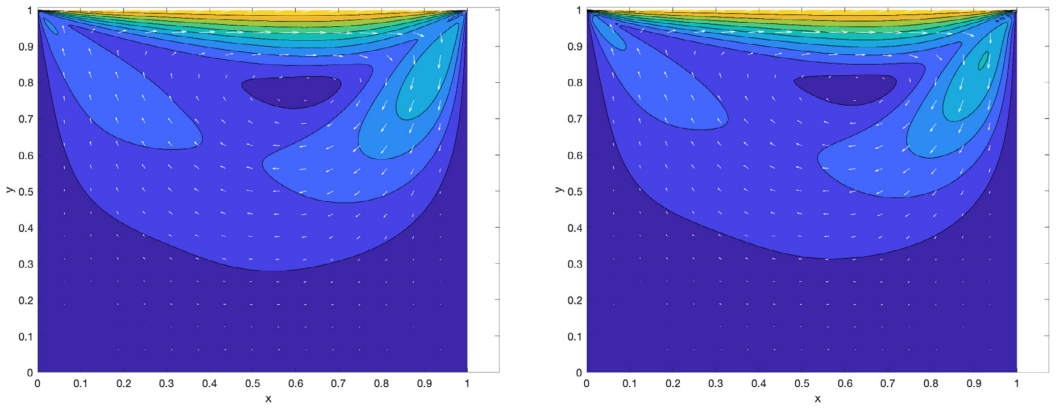


Fig. 5. (a) The velocity field displayed at $t = 2$ using 257^2 grid points and $\alpha = 0.15$. (Second-order scheme.) (b) The velocity field displayed at $t = 2$ using 129^2 grid points. (3rd-order SBP scheme.)

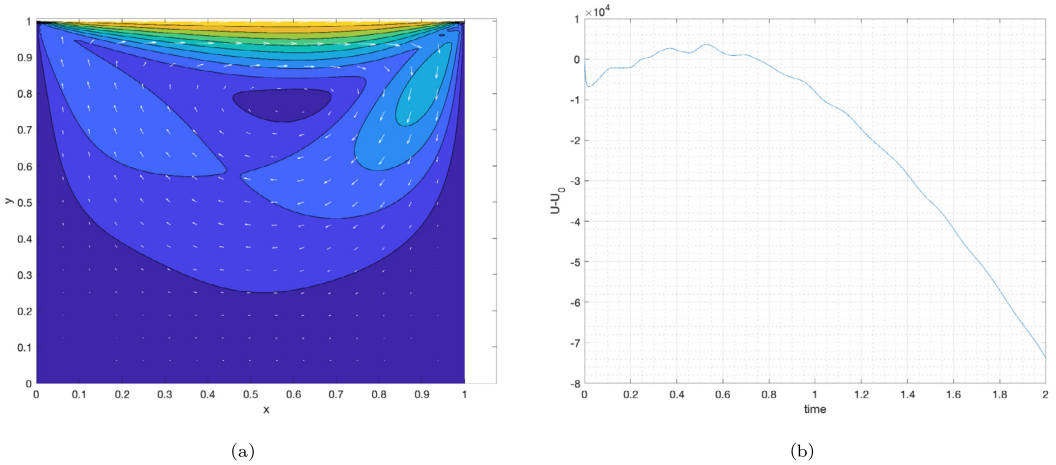


Fig. 6. (a) The velocity field displayed at $t = 2$ using 129^2 grid points. (Second-order scheme using the heat entropy flow boundary condition.) (b) The velocity field displayed at $t = 2$ using 129^2 grid points. (Second-order scheme using the heat entropy flow boundary condition.)

6.2.1. Comments for the implementation

Since one of the velocity components is non-zero at the boundary $y = 1$ for the lid-driven cavity, we must manually update this boundary after each Runge-Kutta stage even when using the proposed scheme with the Dirichlet-SBP operator. This is to take into account the contribution from the continuity equation into the momentum equation at the boundary. The momentum equation is updated as $m|_{\partial\Omega,\{y=1\}} = \rho|_{\partial\Omega,\{y=1\}} v_1|_{\partial\Omega,\{y=1\}}$, where $\rho|_{\partial\Omega,\{y=1\}}$ is given by the continuity equation and $v_1|_{\partial\Omega,\{y=1\}} = 1$.

7. Conclusions

In this article, we have investigated the injection method for strongly imposing the no-slip condition for finite-difference approximations of the compressible Navier-Stokes equations in 1-D and 2-D. Based on standard SBP operators, spatial operators (which we have termed Dirichlet-SBP operators) facilitating the injection procedure were introduced. The temperature condition, on the other hand, was enforced by a SAT. Thus, density, pressure and temperature are updated on the boundary while the momentum is no longer a variable in the boundary points. In particular, we have considered the stability properties of the proposed schemes taking the mixed boundary treatment into account.

When proving linear stability of non-linear problems, it is common to immediately associate the scheme with a linear symmetric constant-coefficient version. Herein, we have rigorously performed all linearisation steps for two different

schemes; one with second-order interior accuracy and one with fourth-order. We prove that the procedure is valid for the proposed 1-D scheme, including the strong-weak imposition of the wall boundary conditions. The linearisation of the 2-D scheme requires considerable more work, but we do not see any additional difficulties beyond more involved algebra and it should also reduce to the same form as the 1-D scheme. Moreover, under the assumption that the interior scheme is entropy stable (52), we have proven that both the proposed 1-D and 2-D 2nd-order schemes, with the mixed boundary treatment, are non-linearly (entropy) stable. The non-linear stability proofs are straightforwardly extendable to 3-D.

Although our proofs rely on the introduction of the Dirichlet-SBP operator, we stress that this operator is not necessary in practice, and has only been introduced for purpose of the proofs. In implementations one can simply overwrite the velocity at the boundary nodes after each Runge-Kutta stage. (This makes the code significantly simpler than with SATs enforcing no-slip.)

Two numerical test cases have been considered; a blast wave and a lid-driven cavity flow. For the blast wave, two types of local Lax-Friedrichs type diffusions were considered: an entropy stable diffusion ($\alpha = 1$) and a non-provably entropy stable diffusion ($\alpha = 0.4$). In both cases, the total entropy was decaying, although a faster decay was observed for the more diffusive scheme (which is as expected). For the lid-driven cavity flow, a reduced local Lax-Friedrichs diffusion ($\alpha = 0.15$) was considered for the 2nd-order scheme. Thereafter, the 3rd-order, linearly stable (but not provably non-linearly stable), scheme was run. The solutions were similar to those obtained in [3]. All test cases demonstrated that the combination of strongly and weakly imposed boundary conditions is robust, and corroborate the claim that the 2-D scheme is stable.

CRediT authorship contribution statement

Anita Gjesteland: Conceptualization, Formal analysis, Software, Writing – original draft, Writing – review & editing. **Magnus Svärd:** Conceptualization, Formal analysis, Methodology, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Appendix A. Linearisation procedure

A.1. Linearisation of the compressible Navier-Stokes equations in 1D

We present the derivation of the linearised and symmetrised Navier-Stokes equations (13), since the details are not presented in [1].

We write the Navier-Stokes equations (10) in primitive variables $\mathsf{v} = (\rho, v, p)^T$:

$$\rho_t + (\rho v)_x = 0, \tag{A.1}$$

$$v_t + v v_x + \tfrac{1}{\rho} p_x = \tfrac{2\mu+\lambda}{\rho} v_{xx}, \tag{A.2}$$

$$p_t + \gamma p v_x + v p_x = (\gamma - 1)(2\mu + \lambda) v_x^2 + \kappa(\gamma - 1)\mathtt{T}_{xx}. \tag{A.3}$$

We decompose each variable into its exact (known smooth and bounded) solution and a small smooth perturbation: $\rho = \rho_{\text{ex}} + \rho'$, etc.

$$(\rho_{\text{ex}} + \rho')_t + \big((\rho_{\text{ex}} + \rho')(v_{\text{ex}} + v')_x\big) = 0,$$

$$(\rho_{\text{ex}})_t + \rho'_t + \big(\rho_{\text{ex}} v_{\text{ex}} + \rho_{\text{ex}} v' + \rho' v_{\text{ex}} + \rho' v'\big)_x = 0,$$

$$(\rho_{\text{ex}})_t + \rho'_t + (\rho_{\text{ex}} v_{\text{ex}})_x + (\rho_{\text{ex}})_x v' + \rho_{\text{ex}} v'_x + \rho'_x v_{\text{ex}} + \rho'(v_{\text{ex}})_x + \rho'_x v' + \rho' v'_x = 0.$$

By definition $(\rho_{\text{ex}})_t + (\rho_{\text{ex}} v_{\text{ex}})_x = 0$, and hence

$$\rho'_t + \underline{(\rho_{\text{ex}})_x v'} + \rho_{\text{ex}} v'_x + \rho'_x v_{\text{ex}} + \underline{\rho'(v_{\text{ex}})_x} + \rho'_x v' + \rho' v'_x = 0.$$

The underlined terms are zeroth-order derivatives of ρ' and v', and hence do not affect the well-posedness of the problem (see [9]), hence they are omitted. The linearisation is done by neglecting non-linear terms, i.e. $\rho'_x v' + \rho' v'_x$. The final result is

$$\rho'_t + \rho_{\text{ex}} v'_x + v_{\text{ex}} \rho'_x = 0.$$

For the velocity equation, we have

$$(v_{\text{ex}}+v')_t+(v_{\text{ex}}+v')(v_{\text{ex}}+v')_x+\frac{1}{\rho_{\text{ex}}+\rho'}(p_{\text{ex}}+p')_x=\frac{2\mu+\lambda}{\rho_{\text{ex}}+\rho'}(v_{\text{ex}}+v')_{xx},$$
$$(v_{\text{ex}})_t+v'_t+v_{\text{ex}}(v_{\text{ex}})_x+v_{\text{ex}}v'_x+v'(v_{\text{ex}})_x+v'v'_x+\frac{(p_{\text{ex}})_x}{\rho_{\text{ex}}+\rho'}+\frac{p'_x}{\rho_{\text{ex}}+\rho'}=(2\mu+\lambda)\left(\frac{(v_{\text{ex}})_{xx}}{\rho_{\text{ex}}+\rho'}+\frac{v'_{xx}}{\rho_{\text{ex}}+\rho'}\right).$$

Factorise $\frac{1}{\rho_{\text{ex}}+\rho'}=\frac{1}{\rho_{\text{ex}}}\frac{1}{1+\frac{\rho'}{\rho_{\text{ex}}}}$, and Taylor expand the second factor; $\frac{1}{1+\frac{\rho'}{\rho_{\text{ex}}}}=1-\frac{\rho'}{\rho_{\text{ex}}}+\mathcal{O}((\rho'/\rho_{\text{ex}})^2)$. Using that the exact solution satisfies Equation (A.2), we have

$$\begin{aligned}&v'_t+v_{\text{ex}}v'_x+v'(v_{\text{ex}})_x+v'v'_x+(p_{\text{ex}})_x\left(-\frac{\rho'}{\rho_{\text{ex}}^2}+\mathcal{O}(\rho'/\rho_{\text{ex}}^2)\right)+p'_x\left(\frac{1}{\rho_{\text{ex}}}-\frac{\rho'}{\rho_{\text{ex}}^2}+\mathcal{O}(\rho'/\rho_{\text{ex}}^2)\right),\\&=(2\mu+\lambda)\left((v_{\text{ex}})_{xx}\left(-\frac{\rho'}{\rho_{\text{ex}}^2}+\mathcal{O}((\rho')^2/\rho_{\text{ex}}^3)\right)+v'_{xx}\left(\frac{1}{\rho_{\text{ex}}}-\frac{\rho'}{\rho_{\text{ex}}^2}+\mathcal{O}((\rho')^2/\rho_{\text{ex}}^3)\right)\right).\end{aligned}$$

We neglect the non-linear terms and omit zeroth-order terms. This yields

$$v'_t+v_{\text{ex}}v'_x+\frac{p'_x}{\rho_{\text{ex}}}=\frac{2\mu+\lambda}{\rho_{\text{ex}}}v'_{xx}.$$

In the same way, the equation (A.3) becomes

$$\begin{aligned}&(p_{\text{ex}}+p')_t+\gamma(p_{\text{ex}}+p')(v+v')_x+(v_{\text{ex}}+v')(p_{\text{ex}}+p')_x\\&=(\gamma-1)(2\mu+\lambda)(v_{\text{ex}}+v')_x^2+\kappa(\gamma-1)(\mathrm{T}_{\text{ex}}+\mathrm{T}')_{xx},\end{aligned}$$

that after expansion becomes,

$$\begin{aligned}&(p_{\text{ex}})_t+p'_t+\gamma(p_{\text{ex}})(v_{\text{ex}})_x+\gamma(p_{\text{ex}})v'_x+\gamma p'(v_{\text{ex}})_x+\gamma p'v'_x+v_{\text{ex}}(p_{\text{ex}})_x+v_{\text{ex}}p'_x+v'(p_{\text{ex}})_x+v'p'_x,\\&=(\gamma-1)(2\mu+\lambda)\left((v_{\text{ex}})_x^2+2(v_{\text{ex}})_xv'_x+v_x'^2\right)+\kappa(\gamma-1)(\mathrm{T}_{\text{ex}}+\mathrm{T})_{xx}.\end{aligned}$$

Next, consider the linearisation of the temperature

$$\begin{aligned}\mathcal{R}((\mathrm{T}_{\text{ex}})_{xx}+\mathrm{T}'_{xx})&=\frac{(p_{\text{ex}})_{xx}+p'_{xx}}{\rho_{\text{ex}}+\rho'}-2\frac{((p_{\text{ex}})_x+p'_x)((\rho_{\text{ex}})_x+\rho'_x)}{(\rho_{\text{ex}}+\rho')^2}\\&\quad+2\frac{(p_{\text{ex}}+p')((\rho_{\text{ex}})_x+\rho'_x)^2}{(\rho_{\text{ex}}+\rho')^3}-\frac{(p_{\text{ex}}+p')((\rho_{\text{ex}})_{xx}+\rho'_{xx})}{(\rho_{\text{ex}}+\rho')^2},\\&=\frac{(p_{\text{ex}})_{xx}+p'_{xx}}{\rho_{\text{ex}}+\rho'}-2\frac{(p_{\text{ex}})_x(\rho_{\text{ex}})_x+(p_{\text{ex}})_x\rho'_x+p'_x(\rho_{\text{ex}})_x+p'_x\rho'_x}{(\rho_{\text{ex}}+\rho')^2}\\&\quad+2\frac{p_{\text{ex}}(\rho_{\text{ex}})_x^2+2p_{\text{ex}}(\rho_{\text{ex}})_x\rho'_x+p_{\text{ex}}\rho_x'^2+p'(\rho_{\text{ex}})_x^2+2p'(\rho_{\text{ex}})_x\rho'_x+p'\rho_x'^2}{(\rho_{\text{ex}}+\rho')^3}\\&\quad-\frac{p_{\text{ex}}(\rho_{\text{ex}})_{xx}+p_{\text{ex}}\rho'_{xx}+p'(\rho_{\text{ex}})_{xx}+p'\rho'_{xx}}{(\rho_{\text{ex}}+\rho')^2}.\end{aligned}$$

Taylor expanding yields

$$\begin{aligned}&\mathcal{R}((\mathrm{T}_{\text{ex}})_{xx}+\mathrm{T}'_{xx})\\&=((p_{\text{ex}})_{xx}+p'_{xx})\left(\frac{1}{\rho_{\text{ex}}}-\frac{\rho'}{\rho_{\text{ex}}^2}+\mathcal{O}((\rho')^2/\rho_{\text{ex}}^3)\right)\\&\quad-2\left((p_{\text{ex}})_x(\rho_{\text{ex}})_x+(p_{\text{ex}})_x\rho'_x+p'_x(\rho_{\text{ex}})_x+p'_x\rho'_x\right)\left(\frac{1}{\rho_{\text{ex}}^2}-\frac{\rho'}{\rho_{\text{ex}}^3}+\mathcal{O}((\rho')^2/\rho_{\text{ex}}^4)\right)\\&\quad+2\left(p_{\text{ex}}(\rho_{\text{ex}})_x^2+2p_{\text{ex}}(\rho_{\text{ex}})_x\rho'_x+p_{\text{ex}}\rho_x'^2+p'(\rho_{\text{ex}})_x^2+2p'(\rho_{\text{ex}})_x\rho'_x+p'\rho_x'^2\right)\left(\frac{1}{\rho_{\text{ex}}^3}-\frac{\rho'}{\rho_{\text{ex}}^4}+\mathcal{O}((\rho')^2/\rho_{\text{ex}}^5)\right)\\&\quad-\left(p_{\text{ex}}(\rho_{\text{ex}})_{xx}+p_{\text{ex}}\rho'_{xx}+p'(\rho_{\text{ex}})_{xx}+p'\rho'_{xx}\right)\left(\frac{1}{\rho_{\text{ex}}^2}-\frac{\rho'}{\rho_{\text{ex}}^3}+\mathcal{O}((\rho')^2/\rho_{\text{ex}}^4)\right).\end{aligned}$$

The exact solution disappears and the quadratic terms are neglected in the linearisation procedure to obtain

$$\mathcal{R}\mathrm{T}'_{xx} = \frac{p'_{xx}}{\rho_{\mathrm{ex}}} - 2\frac{(p_{\mathrm{ex}})_x\rho'_x + p'_x(\rho_{\mathrm{ex}})_x}{\rho_{\mathrm{ex}}^2} + \frac{4p_{\mathrm{ex}}(\rho_{\mathrm{ex}})_x\rho'_x}{\rho_{\mathrm{ex}}^3} - \frac{p_{\mathrm{ex}}\rho'_{xx}}{\rho_{\mathrm{ex}}^2}.$$

Turning back to the pressure equation, and using that the exact solution satisfies (A.3), we end up with

$$\begin{aligned} &p'_t + \gamma(p_{\mathrm{ex}})v'_x + \gamma p'(v_{\mathrm{ex}})_x + \gamma p'v'_x + v_{\mathrm{ex}}p'_x + v'(p_{\mathrm{ex}})_x + v'p'_x, \\ &= (\gamma-1)(2\mu+\lambda)\left(2(v_{\mathrm{ex}})_xv'_x + v'^2_x\right) + \kappa(\gamma-1)\mathrm{T}'_{xx}. \end{aligned}$$

Since non-principal parts of the viscous flux can be bounded by the principal part in the interior and do not affect the number of boundary conditions needed for linear well-posedness, we drop them together with all non-linear terms, and obtain

$$p'_t + \gamma p_{\mathrm{ex}}v'_x + v_{\mathrm{ex}}p'_x = \frac{\kappa}{\mathcal{R}}(\gamma-1)\left(\frac{p'_{xx}}{\rho_{\mathrm{ex}}} - \frac{p}{\rho_{\mathrm{ex}}^2}\rho'_{xx}\right).$$

Next, we freeze the coefficients (the exact solutions), and denote them by the superscript star. We end up with the linearised system

$$\begin{aligned} &\rho'_t + v^*\rho'_x + \rho^*v'_x = 0, \\ &v'_t + v^*v'_x + \tfrac{1}{\rho^*}p'_x = \tfrac{2\mu+\lambda}{\rho^*}v'_{xx}, \\ &p'_t + \gamma p^*v'_x + v^*p'_x = -\tfrac{\gamma\mu p^*}{\mathrm{Pr}\rho^{*2}}\rho'_{xx} + \tfrac{\gamma\mu}{\mathrm{Pr}\rho}p'_{xx}. \end{aligned}$$

(This is the starting point in [1].)

A.2. Linearised gas law

Recall that $p = \rho\mathcal{R}\mathrm{T}$. By the same procedure as above, we linearise this gas law as follows

$$p_{\mathrm{ex}} + p' = \mathcal{R}\left(\rho_{\mathrm{ex}} + \rho'\right)\left(\mathrm{T}_{\mathrm{ex}} + \mathrm{T}'\right) = \mathcal{R}\left(\rho_{\mathrm{ex}}\mathrm{T}_{\mathrm{ex}} + \rho_{\mathrm{ex}}\mathrm{T}' + \rho'\mathrm{T}_{\mathrm{ex}} + \rho'\mathrm{T}'\right).$$

Since $p_{\mathrm{ex}} = \mathcal{R}\rho_{\mathrm{ex}}\mathrm{T}_{\mathrm{ex}}$, by neglecting the non-linear term $\mathcal{R}\rho'\mathrm{T}'$, this reduces to

$$p' = \mathcal{R}\left(\rho_{\mathrm{ex}}\mathrm{T}' + \rho'\mathrm{T}_{\mathrm{ex}}\right).$$

Solving for T' and freezing the coefficients yields

$$\mathrm{T}' = \frac{1}{\mathcal{R}}\left(\frac{p'}{\rho^*} - \frac{p^*\rho'}{\rho^{*2}}\right).$$

Appendix B. Kronecker products

Let $\tilde{B}_N$ be the $(N+1)\times(N+1)$ matrix given by

$$\tilde{B}_N = \begin{pmatrix} 0 & -\frac{1}{2} & 0 & 0 & 0 & \dots & 0 \\ -\frac{1}{2} & 0 & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & 0 & 0 & 0 & \dots & 0 \\ \vdots & & & \ddots & & & \vdots \\ 0 & \dots & 0 & 0 & 0 & 0 & 0 \\ 0 & \dots & 0 & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \dots & 0 & 0 & 0 & \frac{1}{2} & 0 \end{pmatrix},$$

and $\tilde{B}_M$ the $(M+1)\times(M+1)$ matrix with the same form as $\tilde{B}_N$. Furthermore, let $\tilde{H}_x = H_x\tilde{I}_N$, where $H_x = h_x \cdot \mathrm{diag}(1/2, 1, \dots, 1, 1/2)$ and $\tilde{I}_N$ is the $(N+1)\times(N+1)$ identity matrix, with the upper left and lower right element set to zero. $\tilde{H}_y$ is defined similarly (see Section 5.6).

Next, for a two-dimensional grid, the solution vectors are ordered as

$$\boldsymbol{u}^T = \begin{pmatrix} \boldsymbol{u}_{i,1}^T & \boldsymbol{u}_{i,1}^T & \boldsymbol{u}_{i,2}^T & \dots & \boldsymbol{u}_{i,M}^T \end{pmatrix},$$

where $\boldsymbol{u}_{i,j}^T = \begin{pmatrix} u_{0,j} & u_{1,j} & u_{2,j} & \dots & u_{N,j} \end{pmatrix}$. This more compact form of writing the vectors will be convenient below.

The Kronecker products, $\tilde{H}_y \otimes \tilde{B}_N$ and $\tilde{B}_M \otimes \tilde{H}_x$ written as matrices can be stated more compactly in the following way

$$\tilde{H}_y \otimes \tilde{B}_N = h_y \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \begin{pmatrix} 0 & -\frac{1}{2} & \dots & 0 \\ -\frac{1}{2} & 0 & \dots & 0 \\ 0 & \dots & 0 & \frac{1}{2} \\ 0 & \dots & \frac{1}{2} & 0 \end{pmatrix} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \begin{pmatrix} 0 & -\frac{1}{2} & \dots & 0 \\ -\frac{1}{2} & 0 & \dots & 0 \\ 0 & \dots & 0 & \frac{1}{2} \\ 0 & \dots & \frac{1}{2} & 0 \end{pmatrix} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & & \ddots & & \vdots \\ \vdots & & & & \ddots & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

$$\tilde{B}_M \otimes \tilde{H}_x = h_x \begin{pmatrix} \mathbf{0} & -\frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & \ddots & & \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ -\frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & \ddots & & \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & & & \ddots & & \vdots \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & \ddots & & \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix} \\ \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \frac{1}{2}\begin{pmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ & & \ddots & & \\ 0 & \dots & 0 & 1 & 0 \\ 0 & \dots & 0 & 0 & 0 \end{pmatrix} & \mathbf{0} \end{pmatrix},$$

where the bold-face zero denotes a matrix of size $(N+1) \times (N+1)$ with all elements being zero.

Applying these products to a vector, $\boldsymbol{u}$, yields

$$(\tilde{H}_y \otimes \tilde{B}_N)\boldsymbol{u} = h_y \begin{pmatrix} \underline{\mathbf{0}} & \tilde{B}_N \boldsymbol{u}_{i,1} & \tilde{B}_N \boldsymbol{u}_{i,2} & \dots & \tilde{B}_N \boldsymbol{u}_{i,M-1} & \underline{\mathbf{0}} \end{pmatrix}^T, \tag{B.1}$$

$$(\tilde{B}_M \otimes \tilde{H}_x)\boldsymbol{u} = \begin{pmatrix} -\frac{1}{2}\tilde{H}_x \boldsymbol{u}_{i,1} & -\frac{1}{2}\tilde{H}_x \boldsymbol{u}_{i,0} & \underline{\mathbf{0}} & \dots & \underline{\mathbf{0}} & \frac{1}{2}\tilde{H}_x \boldsymbol{u}_{i,M} & \frac{1}{2}\tilde{H}_x \boldsymbol{u}_{i,M-1} \end{pmatrix}^T. \tag{B.2}$$

Note that in the above expressions, the underlined bold-face zeros denotes vectors of length $N+1$ with all elements being zero.

References

[1] S. Abarbanel, D. Gottlieb, Optimal time splitting for two- and three-dimensional Navier-Stokes equations with mixed derivatives, J. Comput. Phys. 41 (1) (1981) 1–33, https://doi.org/10.1016/0021-9991(81)90077-2.

[2] M.H. Carpenter, D. Gottlieb, S. Abarbanel, Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes, J. Comput. Phys. 111 (1994) 220–236.

[3] J. Chan, Y. Lin, T. Warburton, Entropy stable modal discontinuous Galerkin schemes and wall boundary conditions for the compressible Navier-Stokes equations, J. Comput. Phys. 448 (2022), https://doi.org/10.1016/j.jcp.2021.110723.

[4] L. Dalcin, D. Rojas, S. Zampini, D.C.D.R. Fernández, M.H. Carpenter, M. Parsani, Conservative and entropy stable solid wall boundary conditions for the compressible Navier-Stokes equations: adiabatic wall and heat entropy transfer, J. Comput. Phys. 397 (2019) 108775, https://doi.org/10.1016/j.jcp.2019.06.051.

[5] D.C.D.R. Fernández, J.E. Hicken, D.W. Zingg, Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations, Comput. Fluids 95 (2014) 171–196, https://doi.org/10.1016/j.compfluid.2014.02.016.

[6] G.J. Gassner, M. Svärd, F.J. Hindenlang, Stability issues of entropy-stable and/or split form high-order schemes, J. Sci. Comput. 90 (79) (2022), https://doi.org/10.1007/s10915-021-01720-8.

[7] S. Gottlieb, On high order strong stability preserving Runge-Kutta and multi step time discretizations, J. Sci. Comput. 25 (1) (2005), https://doi.org/10.1007/s10915-004-4635-5.

[8] B. Gustafsson, High Order Difference Methods for Time Dependent PDE, Springer Series in Computational Mathematics, Springer-Verlag Berlin Heidelberg, ISBN 978-3-540-74993-6, 2008.

[9] B. Gustafsson, H.-O. Kreiss, J. Oliger, Time-Dependent Problems and Difference Methods, 2nd edition, Pure and Applied Mathematics, John Wiley & Sons, Inc., ISBN 978-0-470-90056-7, 2013.

[10] A. Harten, On the symmetric form of systems of conservation laws with entropy, J. Comput. Phys. 49 (1983) 151–164.
[11] T.J.R. Hughes, L. Franca, M. Mallet, A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics, Comput. Methods Appl. Mech. Eng. 54 (1986) 223–234, https://doi.org/10.1016/0045-7825(86)90127-1.
[12] K.O. Friedrichs, P.D. Lax, System of conservation equations with a convex extension, Proc. Natl. Acad. Sci. USA 68 (8) (1971) 1686–1688.
[13] H.-O. Kreiss, J. Lorenz, Initial-Boundary Value Problems and the Navier-Stokes Equations, Academic Press, ISBN 0-89871-565-2, 1989.
[14] H.O. Kreiss, G. Scherer, Finite element and finite difference methods for hyperbolic partial differential equations, in: Mathematical Aspects of Finite Elements in Partial Differential Equations, 1974.
[15] H.O. Kreiss, L. Wu, On the stability definition of difference approximations for the initial boundary value problem, Appl. Numer. Math. 12 (1993) 213–227.
[16] K. Mattsson, J. Nordström, Summation by parts operators for finite difference approximations of second derivatives, J. Comput. Phys. 199 (2004) 503–540, https://doi.org/10.1016/j.jcp.2004.03.001.
[17] K. Mattsson, M. Svärd, J. Nordström, Stable and accurate artificial dissipation, J. Sci. Comput. 21 (1) (2004).
[18] K. Mattsson, F. Ham, G. Iaccarino, Stable and accurate wave-propagation in discontinuous media, J. Comput. Phys. 227 (2008) 8753–8767, https://doi.org/10.1016/j.jcp.2008.06.0.
[19] K. Mattsson, M. Svärd, M. Shoeybi, Stable and accurate schemes for the compressible Navier-Stokes equations, J. Comput. Phys. 227 (2008) 2293–2316, https://doi.org/10.1016/j.jcp.2007.10.018.
[20] S. Mishra, M. Svärd, On stability of numerical schemes via frozen coefficients and the magnetic induction equations, BIT Numer. Math. 50 (1) (2010) 85–108, https://doi.org/10.1007/s10543-010-0249-5.
[21] J. Nordström, K. Forsberg, C. Adamsson, P. Eliasson, Finite volume methods, unstructured meshes and strict stability for hyperbolic problems, Appl. Numer. Math. 45 (2003) 453–473.
[22] J. Nordström, S. Eriksson, P. Eliasson, Weak and strong wall boundary procedures and convergence to steady-state of the Navier-Stokes equations, J. Comput. Phys. 231 (2012) 4867–4884, https://doi.org/10.1016/j.jcp.2012.04.007.
[23] M. Parsani, M.H. Carpenter, E.J. Nielsen, Entropy stable wall boundary conditions for the three-dimensional compressible Navier-Stokes equations, J. Comput. Phys. 292 (2015) 88–113, https://doi.org/10.1016/j.jcp.2015.03.026.
[24] G. Strang, Accurate partial difference methods II. Non-linear problems, Numer. Math. 6 (1964) 37–46.
[25] M. Svärd, A convergent numerical scheme for the compressible Navier-Stokes equations, SIAM J. Numer. Anal. 54 (3) (2016) 1484–1506.
[26] M. Svärd, A new Eulerian model for viscous and heat conducting compressible flows, Physica A 506 (2018) 350–375, https://doi.org/10.1016/j.physa.2018.03.097.
[27] M. Svärd, Entropy stable boundary conditions for the Euler equations, J. Comput. Phys. (2020), https://doi.org/10.1016/j.jcp.2020.109947.
[28] M. Svärd, J. Nordström, A stable high-order finite difference scheme for the compressible Navier-Stokes equations, no-slip wall boundary conditions, J. Comput. Phys. 227 (2008) 4805–4824, https://doi.org/10.1016/j.jcp.2007.12.028.
[29] M. Svärd, J. Nordström, Review of summation-by-parts schemes for initial-boundary-value problems, J. Comput. Phys. 268 (2014) 17–38, https://doi.org/10.1016/j.jcp.2014.02.031.
[30] M. Svärd, M.H. Carpenter, J. Nordström, A stable high-order finite difference scheme for the compressible Navier-Stokes equations, far-field boundary conditions, J. Comput. Phys. 225 (2007) 1020–1038, https://doi.org/10.1016/j.jcp.2007.01.023.
[31] M. Svärd, M.H. Carpenter, M. Parsani, Entropy stability and the no-slip wall boundary condition, SIAM J. Numer. Anal. 56 (1) (2018) 256–273.
[32] E. Tadmor, Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems, Acta Numer. (2003) 451–512, https://doi.org/10.1017/S0962492902000156.

Paper B

Convergence of Chandrashekar's Second-Derivative Finite-Volume Approximation

A. Gjesteland and M. Svärd

Journal of Scientific Computing, **96:46**, 2023.

Journal of Scientific Computing (2023) 96:46
https://doi.org/10.1007/s10915-023-02256-9

Convergence of Chandrashekar's Second-Derivative Finite-Volume Approximation

Anita Gjesteland[1] · **Magnus Svärd**[1]

Received: 7 October 2022 / Revised: 9 March 2023 / Accepted: 21 May 2023 /
Published online: 20 June 2023

Abstract
We consider a slightly modified local finite-volume approximation of the Laplacian operator originally proposed by Chandrashekar (Int J Adv Eng Sci Appl Math 8(3):174–193, 2016, https://doi.org/10.1007/s12572-015-0160-z). The goal is to prove consistency and convergence of the approximation on unstructured grids. Consequently, we propose a semi-discrete scheme for the heat equation augmented with Dirichlet, Neumann and Robin boundary conditions. By deriving a priori estimates for the numerical solution, we prove that it converges weakly, and subsequently strongly, to a weak solution of the original problem. A numerical simulation demonstrates that the scheme converges with a second-order rate.

Keywords Finite volume · Second derivative · Convergence

1 Introduction

The compressible Navier–Stokes equations are the foundation of computational fluid dynamics (CFD) for modelling the flow of viscous compressible fluids. Consequently, numerical methods for approximating their solutions are vastly studied. For a numerical scheme to yield a convergent sequence of approximate solutions, it must be a stable discretisation of the well-posed continuous problem. For linear partial differential equations (PDEs), the energy method, which depends heavily on integration by parts (IBP), is often used to prove well-posedness. In the (semi-)discrete setting, analogous stability proofs can be obtained by using the discrete energy method, where IBP is mimicked using summation-by-parts (SBP). Numerical methods formulated to satisfy the SBP property are thus frequently used for various PDEs, including CFD problems (see e.g. [6, 9, 24, 31, 32]).

Different numerical methods can be formulated in the SBP framework. These include the finite-difference methods (see e.g. [20, 21, 28]), the finite-volume methods (see e.g. [7,

Anita Gjesteland
anita.gjesteland@uib.no

Magnus Svärd
magnus.svard@uib.no

[1] Department of Mathematics, University of Bergen, Postbox 7800, 5020 Bergen, Norway

23, 29]) and the discontinuous Galerkin spectral element methods (see e.g. [12]). The latter two may be formulated on unstructured grids, that are sometimes preferred for domains with complex geometries.

Herein, we focus the attention on local finite-volume methods that only use nearest neighbours to approximate the derivatives. These are still the workhorse methods in production CFD, due to their simplicity and robustness, and since the local structure allows for easier parallelisation. A well-known drawback is however the difficulty of finding consistent second-derivative approximations, which hampers their usability for the compressible Navier–Stokes equations. It was, for example, shown in [29] that a commonly used edge-based approximation is inconsistent on general unstructured grids. Although proofs of convergence exist for finite-volume methods, they often rely on *admissible meshes* (in the sense of Def. 3.1 in [11], see e.g. [3, 13]), that require normal derivative approximations at volume faces to be orthogonal to the face. This severely constrains mesh generation. Hence, it is desirable to design a local finite-volume scheme that runs on standard unstructured grids such as Delaunay triangulations.

In the interest of accurately discretising the viscous terms of the compressible Navier–Stokes equations on such grids, we study the Laplacian approximation proposed by Chandrashekar in [7]. His approximation incorporated the Dirichlet boundary conditions weakly, and the resulting operator was shown to satisfy the SBP property. The approximation was then used to discretise the heat equation, and numerical experiments showed that the scheme converged with second-order rate on triangulated grids.

In this work, we slightly modify the Laplacian operator from [7], by not including any boundary conditions directly in the operator. We mimic the proof of Chandrashekar, and demonstrate that the modified operator maintain the SBP property proved in [7]. To study the consistency and convergence of the Laplacian approximation, we use the heat equation as a model equation. We propose a numerical scheme for this equation where the Dirichlet boundary conditions are imposed strongly by injection (see e.g. [16] for more information about this technique), and the Neumann and Robin boundary coditions are imposed weakly similar to [7]. This approach is analogous to the one used in [15] to prove both linear and non-linear stability for the compressible Navier–Stokes equations augmented with the no-slip adiabatic wall boundary conditions on structured grids.

The main goal herein is to mathematically prove the convergence of the proposed scheme for the heat equation, thus also proving the consistency, in a weak sense, of the second-derivative approximation. By utilising the SBP properties of the Laplacian operator, we find a priori H^1 estimates for the numerical solution. These estimates guarantee that the numerical solution converges weakly (up to a subsequence) to a weak solution of the heat equation. Furthermore, we show that the numerical solution converges strongly by employing Aubin–Lions' lemma, and subsequently show that the weak solution is unique. The present proof is valid for general triangular grids with Lipschitz boundaries, and does not require *admissible meshes*. By using the method of manufactured solutions, we verify by a numerical experiment that the scheme is convergent.

Remark 1.1 To the best of our knowledge, this is the first convergence proof for a local finite-volume method for the second-derivative that does not require admissible meshes. We note that some multi-point flux approximations (MPFA) finite-volume methods have been proven convergent by identifying them as mixed finite-element approximations (see e.g. [2, 18]).

The proof presented herein is also easily adapted to weakly imposed boundary conditions. Stability for such a scheme for the heat equation was established in [7], and herein we

show that injected Dirichlet boundary conditions also yield a stable scheme. That is, both approaches are applicable, and we have chosen the strong imposition to provide an alternative.

The paper is further organised as follows. Section 2 defines the problem, whereas a priori estimates for the continuous solution are found in Sect. 3. In Sect. 4 we state the weak formulation of the problem. Section 5 concerns the spatial discretisation and provides the proof of the slightly altered Laplacian operator being SBP. In Sect. 6, the numerical scheme that approximates our problem is stated. Furthermore, the SBP properties of the Laplacian operator are utilised to obtain discrete a priori estimates similar to those found for the continuous solution. Using these estimates, we show in Sect. 7 that the approximate solution obtained by the proposed numerical scheme converges weakly to a weak solution of the original problem. Furthermore, we show in Sect. 8 that the solution indeed converges strongly by using Aubin–Lions' lemma. The solution is subsequently shown to be unique in Sect. 9. Finally, Sect. 10 provides a numerical example that demonstrates the convergence of the scheme.

2 Problem Statement

Consider the heat equation on a two-dimensional open polygonal Lipschitz domain, Ω, with boundary $\partial\Omega$:

$$\begin{aligned} v_t &= \nabla\cdot(\mu\nabla v), && \text{on } \Omega\times(0,T],\\ v &= g^D, && \text{on } \partial\Omega^D\times[0,T],\\ \mu\nabla v\cdot\boldsymbol{n} &= g^N, && \text{on } \partial\Omega^N\times[0,T],\\ \mu\nabla v\cdot\boldsymbol{n}+\alpha v &= g^R, && \text{on } \partial\Omega^R\times[0,T],\\ v|_{t=0} &= f, && \text{on } \Omega. \end{aligned} \tag{1}$$

The superscripts D, N, R indicate the Dirichlet, Neumann and Robin parts of the boundary with corresponding boundary data. We assume $\partial\Omega^D\cup\partial\Omega^N\cup\partial\Omega^R=\partial\Omega$, and $\partial\Omega^D\cap\partial\Omega^N=\partial\Omega^D\cap\partial\Omega^R=\partial\Omega^N\cap\partial\Omega^R=\emptyset$. Furthermore, $\boldsymbol{n}$ denotes the outward unit normal vector, $f\in L^2(\Omega)$ is the initial data, and $\mu>0$, $\alpha\geq 0$ are constants. We take $g^D\in H^1(0,T;H^{1/2}(\partial\Omega^D))$ and $g^{N,R}\in L^2(0,T;L^2(\partial\Omega^{N,R}))$.

To simplify the forthcoming analysis, we define a function, w, such that $w\in L^2(0,T;H^1(\Omega))$ and $w_t\in L^2(0,T;H^1(\Omega))$, and $w|_{\partial\Omega^D}=g^D$ (in the sense of traces). By the trace theorem, we know there exists such a $w\in H^1(\Omega)$ (see [1]). Lastly, we choose w to satisfy $w|_{t=0}=f$, and

$$\begin{aligned} \mu\nabla w\cdot\boldsymbol{n} &= 0 \quad \text{on } \partial\Omega^N,\\ \mu\nabla w\cdot\boldsymbol{n}+\alpha w &= 0 \quad \text{on } \partial\Omega^R. \end{aligned} \tag{2}$$

Then, by introducing $u=v-w$ (see e.g. [1, 17]), (1) can be recast to

$$u_t = \nabla\cdot(\mu\nabla u)+F, \qquad \text{on } \Omega\times(0,T], \tag{3a}$$
$$u = 0, \qquad \text{on } \partial\Omega^D\times[0,T], \tag{3b}$$
$$\mu\nabla u\cdot\boldsymbol{n} = g^N, \qquad \text{on } \partial\Omega^N\times[0,T], \tag{3c}$$
$$\mu\nabla u\cdot\boldsymbol{n}+\alpha u = g^R, \qquad \text{on } \partial\Omega^R\times[0,T], \tag{3d}$$
$$u|_{t=0} = 0, \qquad \text{on } \Omega. \tag{3e}$$

Here,

$$F = \nabla \cdot (\mu \nabla w) - w_t, \tag{4}$$

is a forcing function.

Remark 2.1 We could have made all boundary conditions homogeneous by defining w differently. However, we choose non-zero Neumann and Robin data to keep the regularity assumptions on the boundary data to a minimum.

3 A Priori Estimates for the Continuous Problem

To obtain a priori estimates on u, we use the energy method (see e.g. [17]). By inserting F given in (4) into (3a) and integrating by parts, we obtain

$$\begin{aligned}
\int_\Omega u u_t \, d\mathbf{x} &= \int_\Omega u \left(\nabla \cdot (\mu \nabla u) + \nabla \cdot (\mu \nabla w) - w_t\right) \, d\mathbf{x}, \\
\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} &= -\int_\Omega \left(\mu \nabla u \cdot (\nabla u + \nabla w) + u w_t\right) \, d\mathbf{x} \\
&\quad + \int_{\partial\Omega} u(\mu \nabla u \cdot \boldsymbol{n} + \mu \nabla w \cdot \boldsymbol{n}) \, ds, \\
&= -\mu \|\nabla u\|^2_{L^2(\Omega)} - \int_\Omega \left(\mu \nabla u \cdot \nabla w + u w_t\right) \, d\mathbf{x} \\
&\quad + \int_{\partial\Omega} u(\mu \nabla u \cdot \boldsymbol{n} + \mu \nabla w \cdot \boldsymbol{n}) \, ds.
\end{aligned}$$

Using Cauchy–Schwarz's and Young's inequality on the first integral on the right-hand side, we obtain

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} &\le -\mu \|\nabla u\|^2_{L^2(\Omega)} + \tfrac{\epsilon}{2}\mu \|\nabla u\|^2_{L^2(\Omega)} + \tfrac{1}{2\epsilon}\mu \|\nabla w\|^2_{L^2(\Omega)} + \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)} \\
&\quad + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)} \\
&\quad + \int_{\partial\Omega} u \left(\mu \nabla u \cdot \boldsymbol{n} + \mu \nabla w \cdot \boldsymbol{n}\right) \, ds.
\end{aligned} \tag{5}$$

By choosing $\epsilon = 1$, the term $-\mu \|\nabla u\|^2_{L^2(\Omega)} + \frac{\epsilon}{2}\mu \|\nabla u\|^2_{L^2(\Omega)} = -\frac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)}$. Since ϵ is now determined, $\frac{1}{2\epsilon}\mu \|\nabla w\|^2_{L^2(\Omega)} = \frac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)}$, which is bounded by definition. Hence, (5) reads

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} &\le -\tfrac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)} + \tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)} + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)} \\
&\quad + \int_{\partial\Omega} u \left(\mu \nabla u \cdot \boldsymbol{n} + \mu \nabla w \cdot \boldsymbol{n}\right) \, ds.
\end{aligned}$$

Inserting the boundary conditions for w and u given in (2) and (3b)–(3d), respectively, we obtain

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} &\le -\tfrac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)} + \tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)} + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)} \\
&\quad + \int_{\partial\Omega^N} \underline{u g^N} \, ds + \int_{\partial\Omega^R} \left(u \left(\underline{g^R} - \alpha u\right) \underline{- \alpha u w}\right) \, ds.
\end{aligned} \tag{6}$$

Consider the underlined boundary terms above. We follow [19], and bound these terms by first using the Cauchy–Schwarz inequality:

$$\int_{\partial\Omega^N} ug^N\,ds + \int_{\partial\Omega^R} ug^R - \alpha uw\,ds \leq \|u\|_{L^2(\partial\Omega^N)}\|g^N\|_{L^2(\partial\Omega^N)} + \|u\|_{L^2(\partial\Omega^R)}\|g^R\|_{L^2(\partial\Omega^R)} + \alpha\|u\|_{L^2(\partial\Omega^R)}\|w\|_{L^2(\partial\Omega^R)}, \tag{7}$$

and then by using the trace theorem, which states that $\|u\|_{L^2(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)}$, $C > 0$ (see e.g. [1]):

$$\begin{aligned}&\|u\|_{L^2(\partial\Omega^N)}\|g^N\|_{L^2(\partial\Omega^N)} + \|u\|_{L^2(\partial\Omega^R)}\left(\|g^R\|_{L^2(\partial\Omega^R)} + \alpha\|w\|_{L^2(\partial\Omega^R)}\right)\\ &\lesssim \|u\|_{H^1(\Omega)}\left(\|g^N\|_{L^2(\partial\Omega^N)} + \|g^R\|_{L^2(\partial\Omega^R)} + \alpha\|w\|_{H^1(\Omega)}\right).\end{aligned}$$

Here, we have introduced the notation $a \lesssim b$ for $a \leq Cb$, where $C > 0$ is a constant.

By employing Young's inequality, the boundary terms (7) finally read

$$\begin{aligned}&\int_{\partial\Omega^N} ug^N\,ds + \int_{\partial\Omega^R} ug^R - \alpha uw\,ds \lesssim \tfrac{\beta}{2}\|u\|^2_{H^1(\Omega)}\\ &\quad + \tfrac{1}{2\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right).\end{aligned}$$

The preliminary estimate (6), can then be stated as

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} &\lesssim -\tfrac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)} + \tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)}\\ &\quad + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)} + \tfrac{\beta}{2}\|u\|^2_{H^1(\Omega)}\\ &\quad + \tfrac{1}{2\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right) - \int_{\partial\Omega^R}\alpha u^2\,ds.\end{aligned}$$

The last term on the right-hand side is negative semi-definite, since $\alpha \geq 0$. We neglect it in the remaining analysis. Hence we have

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} + \tfrac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)} - \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)} - \tfrac{\beta}{2}\|u\|^2_{H^1(\Omega)}\\ &\qquad\lesssim \tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)}\\ &\qquad\quad + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)} + \tfrac{1}{2\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right).\end{aligned} \tag{8}$$

Consider the three last terms on the left-hand side of the above inequality. By adding and subtracting $\frac{\mu}{2}\|u\|^2_{L^2(\Omega)}$, they can be rewritten as

$$\begin{aligned}&\tfrac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)} - \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)} - \tfrac{\beta}{2}\|u\|^2_{H^1(\Omega)} + \tfrac{\mu}{2}\|u\|^2_{L^2(\Omega)} - \tfrac{\mu}{2}\|u\|^2_{L^2(\Omega)} = \tfrac{\mu-\beta}{2}\|\nabla u\|^2_{H^1(\Omega)}\\ &\quad - \tfrac{\mu+\delta}{2}\|u\|^2_{L^2(\Omega)}.\end{aligned} \tag{9}$$

By choosing β sufficiently small $\frac{\mu-\beta}{2}\|u\|^2_{H^1(\Omega)} \geq \frac{\mu-\beta}{2}\|u\|^2_{L^2(\Omega)} \geq 0$. From (8) we then have

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} &\lesssim \tfrac{\beta+\delta}{2}\|u\|^2_{L^2(\Omega)} + \tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)}\\ &\quad + \tfrac{1}{2\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right).\end{aligned}$$

Employing Grönwall's inequality (see e.g. [10]), we obtain

$$\|u(\cdot,\cdot,t)\|^2_{L^2(\Omega)} \lesssim e^{(\beta+\delta)t}\left(\|u(\cdot,\cdot,0)\|^2_{L^2(\Omega)} + \int_0^T \left(\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{1}{\delta}\|w_t\|^2_{L^2(\Omega)}\right) dt + \int_0^T \tfrac{1}{\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right) dt\right).$$

The inequality holds for all $0 \le t \le T$. In Sect. 2 we defined $w \in L^2(0,T;H^1(\Omega))$, $w_t \in L^2(0,T;H^1(\Omega))$ and $g^{N,R} \in L^2(0,T;L^2(\partial\Omega^{N,R}))$. Using this, we find that the right-hand side of the above inequality is bounded. Thus, $u \in L^\infty(0,T;L^2(\Omega))$. Lastly, we integrate (8) in time to obtain

$$\begin{aligned}
&\frac{1}{2}\|u(\cdot,\cdot,T)\|^2_{L^2(\Omega)} + \int_0^T \left(\tfrac{\mu}{2}\|\nabla u\|^2_{L^2(\Omega)} - \tfrac{\delta}{2}\|u\|^2_{L^2(\Omega)} - \tfrac{\beta}{2}\|u\|^2_{H^1(\Omega)}\right) dt \\
&\quad\lesssim \frac{1}{2}\|u(\cdot,\cdot,0)\|^2_{L^2(\Omega)} \\
&\qquad + \int_0^T \left(\tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)}\right. \\
&\qquad \left. + \tfrac{1}{2\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right)\right) dt \\
&\quad= \int_0^T \left(\tfrac{\mu}{2}\|\nabla w\|^2_{L^2(\Omega)} + \tfrac{1}{2\delta}\|w_t\|^2_{L^2(\Omega)}\right. \\
&\qquad \left. + \tfrac{1}{2\beta}\left(\|g^N\|^2_{L^2(\partial\Omega^N)} + \|g^R\|^2_{L^2(\partial\Omega^R)} + \alpha^2\|w\|^2_{H^1(\Omega)}\right)\right) dt,
\end{aligned}$$

where we have used $u|_{t=0} \equiv 0$ in the last step. Since $\int_0^T \|u(\cdot,\cdot,t)\|^2\, dt \le constant$, we observe from this inequality that $\nabla u \in L^2(0,T;L^2(\Omega))$, and thus we have $u \in L^2(0,T;H^1(\Omega))$.

4 Weak Formulation of the Heat Equation

Next, we derive the weak formulation of (3). Let $H^1_{\partial\Omega_0^D}(\Omega)$ denote the space of H^1 functions vanishing at the Dirichlet boundary. Furthermore, let $\phi \in H^1(0,T;H^1_{\partial\Omega_0^D}(\Omega))$ be a test function that satisfies $\phi(\mathbf{x},T) = 0$, $\mathbf{x} \in \Omega$. Multiply (3a) by ϕ and integrate over Ω.

$$\int_\Omega \phi u_t\, d\mathbf{x} = \int_\Omega \phi \nabla\cdot(\mu\nabla u)\, d\mathbf{x} + \int_\Omega \phi F\, d\mathbf{x}. \tag{10}$$

Integrating by parts and inserting the boundary conditions given in (3b)–(3d), give

$$\begin{aligned}
\int_\Omega \phi u_t\, d\mathbf{x} &= -\int_\Omega \nabla\phi\cdot\mu\nabla u\, d\mathbf{x} + \int_{\partial\Omega^D} \phi(\mu\nabla u\cdot\boldsymbol{n})\, ds + \int_{\partial\Omega^N} \phi g^N\, ds \\
&\quad + \int_{\partial\Omega^R} \phi(g^R - \alpha u)\, ds + \int_\Omega \phi F\, d\mathbf{x}, \\
&= -\int_\Omega \nabla\phi\cdot\mu\nabla u\, d\mathbf{x} + \int_{\partial\Omega^N} \phi g^N\, ds + \int_{\partial\Omega^R} \phi(g^R - \alpha u)\, ds + \int_\Omega \phi F\, d\mathbf{x},
\end{aligned}$$

where we have used $\phi|_{\partial\Omega^D} = 0$. Using $\phi|_{t=T} = 0$, $u|_{t=0} = 0$ and partially integrating the left-hand side in time further yields the weak form of (3):

$$\int_0^T \int_\Omega \phi_t u \, d\mathbf{x}dt = \int_0^T \int_\Omega \nabla\phi \cdot \mu\nabla u \, d\mathbf{x}dt - \int_0^T \int_{\partial\Omega^N} \phi g^N \, dsdt - \int_0^T \int_{\partial\Omega^R} \phi(g^R - \alpha u) \, dsdt - \int_0^T \int_\Omega \phi F \, d\mathbf{x}dt, \tag{11}$$

where F given by (4) satisfies

$$\int_\Omega \phi F \, d\mathbf{x} = -\int_\Omega \nabla\phi \cdot \mu\nabla w \, d\mathbf{x} - \int_{\partial\Omega^R} \alpha\phi w \, ds - \int_\Omega \phi w_t \, d\mathbf{x}. \tag{12}$$

Remark 4.1 Since the forcing function is not the main focus of this work, we use $\int_\Omega \phi F \, d\mathbf{x}$ as short-hand notation for (12) and make comments about it where necessary.

Remark 4.2 From (12), we see that $w \in H^1(\Omega)$ is sufficient to bound the two first integrals on the right-hand side. Furthermore, the regularity of w_t is determined by the regularity of the boundary data (see e.g. [17]). Thus, for $\gamma(w_t) = g_t^D$ to be satisfied (where γ is the trace function), we must have $w_t \in H^1(\Omega)$, and that is why we assumed that $g^D \in H^1(0, T; H^{1/2}(\partial\Omega^D))$ in Sect. 2.

Definition 4.3 A function u satisfying (11) is called a weak solution of the problem (3).

5 Spatial Discretisation

Let $\bar{\Omega}_h$ be a discretisation of $\bar{\Omega} = \Omega \cup \partial\Omega$ into non-overlapping triangles K_n, $n = 1, \ldots, N$ such that $\bar{\Omega}_h = \cup_{n=1}^N K_n$, and such that there are no hanging nodes in $\bar{\Omega}_h$. The grid functions are defined on the vertices of the triangles. Furthermore, subdivide $\bar{\Omega}_h$ into a dual grid consisting of dual cells, V_i, $i = 1, \ldots, I$, such that $\bar{\Omega}_h = \cup_{i=1}^I V_i$. The dual cells are polygons surrounding a vertex, i. A dual-volume boundary consists of straight lines drawn from the mid-point of an edge adjacent to grid point i to the centroid of the triangles adjacent to the grid point (see Fig. 1). (These are the dual volumes of the standard node-centred finite-volume method, see e.g. [22]). We introduce the notation

$$\begin{aligned}
\bar{\Omega}_h^V &: \text{ the set of indices for interior and boundary nodes,}\\
\bar{\Omega}_h^K &: \text{ the set of triangles in} \bar{\Omega}_h,\\
\Omega_h^V &: \text{ the set of indices for interior nodes,}\\
\partial\Omega_h^V &: \text{ the set of indices for boundary nodes,}\\
\partial\Omega_h^N &: \text{ the set of indices for boundary nodes on } \partial\Omega^N,\\
\partial\Omega_h^R &: \text{ the set of indices for boundary nodes on } \partial\Omega^R.
\end{aligned}$$

The discretisation of the problem (3) utilises the approximation of the Laplacian and gradient operator proposed in [7] for the interior scheme. For triangles having at least one edge along the Dirichlet boundary, the Dirichlet condition was incorporated weakly in the gradient operator in [7]. Here, we use the approximation for *interior* triangles for every triangle in the grid. The approximation is given by

$$\nabla_h u^n = -\frac{1}{2|K_n|}\left[u_i \hat{\boldsymbol{n}}_i^n + u_j \hat{\boldsymbol{n}}_j^n + u_k \hat{\boldsymbol{n}}_k^n\right], \tag{13}$$

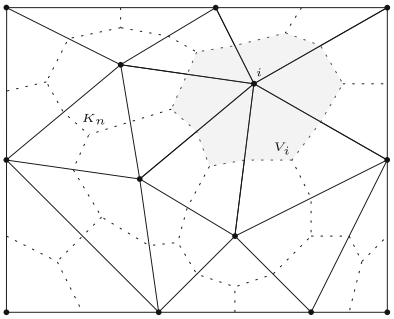


Fig. 1 Example of a triangulation and the corresponding dual cells

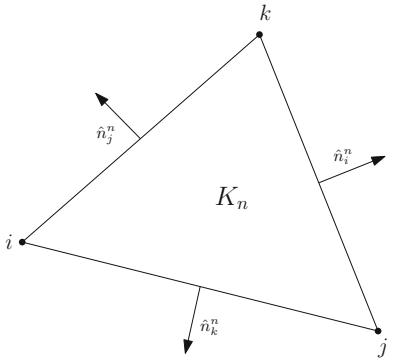


Fig. 2 Triangle depicting the components of the gradient approximation (13)

where $|K_n|$ is the area of triangle K_n; i, j, k are the vertices of the triangle, and $\hat{\boldsymbol{n}}^n_{i,j,k}$ are the outward pointing normal vectors of the triangle, opposite of the particular node (see Fig. 2). The length of the normal vectors, $\hat{\boldsymbol{n}}^n_{i,j,k}$, is equal to the length of the adjacent edge.

Next, we introduce the following notation.

$$\begin{aligned} I_n &= \{\text{all vertices of triangle} n\}, \\ N_i &= \{\text{all triangles with vertex} i\}, \\ E_i &= \{\text{all boundary edges having vertex} i \text{ as an endpoint}\}, \end{aligned}$$

Then the approximation of the Laplacian on a dual volume is found by approximating (see [7])

$$\int_{V_i} \Delta u \, d\mathbf{x} = \int_{\partial V_i \setminus \partial\Omega} \nabla u \cdot \boldsymbol{n} \, ds + \int_{\partial V_i \cap \partial\Omega} \nabla u \cdot \boldsymbol{n} \, ds, \tag{14}$$

by

$$(\Delta_h u)_i = \frac{1}{V_i} \left[\frac{1}{2} \sum_{n \in N_i} \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2} \sum_{e \in E_i} \nabla_h u^{n(e)} \cdot \hat{\boldsymbol{b}}(e) \right]. \tag{15}$$

Here, $\hat{\boldsymbol{b}}(e)$ denotes the outward pointing normal vector at boundary edge e (see Fig. 3a). The superscript $n(e)$ signifies the triangle that has the boundary edge e. The components of the approximation (15) is depicted in Fig. 3b.

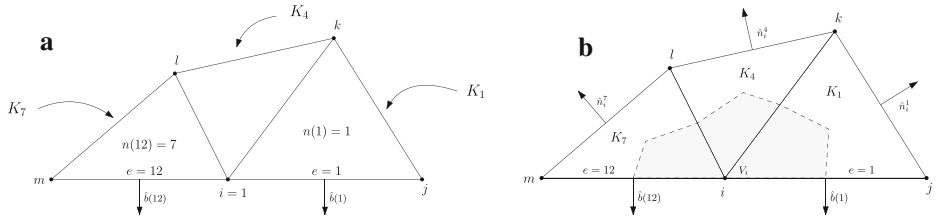


Fig. 3 **a** Example of a vertex i belonging to three triangles (K_1, K_4, K_7) where two of them (K_1, K_7) have an edge along the boundary, depicted with the corresponding boundary normals $\hat{\boldsymbol{b}}(e)$. **b** Example of a dual cell, V_i, and the components of the Laplace approximation (15)

The approximation of the Laplacian (15) with Dirichlet boundary conditions taken into account, was demonstrated to satisfy the Summation-by-Parts (SBP) property in Theorem 1 in [7]. Here, we state the analogous result without any boundary conditions.

Theorem 5.1 *Let u^h and v^h be two grid functions defined on $\bar{\Omega}_h$ such that $u^h = (u_1, u_2, \ldots, u_I)$, and correspondingly for v^h. Then the discrete approximation of the Laplacian operator (15) satisfies the SBP property*

$$\sum_{i\in\bar{\Omega}_h^V} v_i V_i (\Delta_h u^h)_i = -\sum_{n\in\bar{\Omega}_h^K} \nabla_h u^n \cdot \nabla_h v^n |K_n| + \frac{1}{2}\sum_{i\in\partial\Omega_h^V} v_i (\nabla_h u^{n_{i,1}(e)} \cdot \hat{\boldsymbol{b}}_{i,1}(e) + \nabla_h u^{n_{i,2}(e)} \cdot \hat{\boldsymbol{b}}_{i,2}(e)),$$

where the subscripts $\{i, 1\}$ and $\{i, 2\}$ indicate the two edges adjacent to the boundary node i.

Proof Multiply Eq. (15) by $v_i V_i$ and sum over all vertices in the grid.

$$\begin{aligned}\sum_{i\in\bar{\Omega}_h^V} v_i V_i (\Delta_h u^h)_i &= \frac{1}{2}\sum_{i\in\bar{\Omega}_h^V} v_i \sum_{n\in N_i} \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2}\sum_{i\in\bar{\Omega}_h^V} v_i \sum_{e\in E_i} \nabla_h u^{n(e)} \cdot \hat{\boldsymbol{b}}(e),\\ &= \frac{1}{2}\sum_{i\in\bar{\Omega}_h^V}\sum_{n\in N_i} v_i \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2}\sum_{i\in\partial\Omega_h^V}\sum_{e\in E_i} v_i \nabla_h u^{n(e)} \cdot \hat{\boldsymbol{b}}(e). \qquad (16)\end{aligned}$$

In the second equality, we have used that the set E_i is empty for interior nodes.

For the first term in the above equation, we change the order of summation and move $\nabla_h u^n$ outside the summation over the vertices of a triangle K_n in (16), to obtain

$$\frac{1}{2}\sum_{i\in\bar{\Omega}_h^V}\sum_{n\in K_i} v_i \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n = \frac{1}{2}\sum_{n\in\bar{\Omega}_h^K}\sum_{i\in I_n} v_i \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n = \frac{1}{2}\sum_{n\in\bar{\Omega}_h^K} \nabla_h u^n \cdot \sum_{i\in I_n} v_i \hat{\boldsymbol{n}}_i^n. \qquad (17)$$

For the boundary nodes, we have

$$\frac{1}{2}\sum_{i\in\partial\Omega_h^V}\sum_{e\in E_i} v_i \nabla_h u^{n(e)} \cdot \hat{\boldsymbol{b}}(e) = \sum_{i\in\partial\Omega_h^V} v_i (\nabla_h u^{n_{i,1}(e)} \cdot \hat{\boldsymbol{b}}_{i,1}(e) + \nabla_h u^{n_{i,2}(e)} \cdot \hat{\boldsymbol{b}}_{i,2}(e)) \qquad (18)$$

With (17) and (18), (16) can be written as

$$\sum_{i\in\bar{\Omega}_h^V} v_i V_i (\Delta_h u^h)_i = \frac{1}{2}\sum_{n\in\bar{\Omega}_h^K} \nabla_h u^n \cdot \sum_{i\in I_n} v_i \hat{\boldsymbol{n}}_i^n + \frac{1}{2}\sum_{i\in\partial\Omega_h^V} v_i (\nabla_h u^{n_{i,1}(e)} \cdot \hat{\boldsymbol{b}}_{i,1}(e)$$

$$
\begin{aligned}
&+ \nabla_h u^{n_{i,2}(e)} \cdot \hat{\boldsymbol{b}}_{i,2}(e)),\\
&= -\sum_{n\in\bar{\Omega}_h^K} \nabla_h u^n \cdot \nabla_h v^n |K_n| + \frac{1}{2}\sum_{i\in\partial\Omega_h^V} v_i(\nabla_h u^{n_{i,1}(e)} \cdot \hat{\boldsymbol{b}}_{i,1}(e)\\
&+ \nabla_h u^{n_{i,2}(e)} \cdot \hat{\boldsymbol{b}}_{i,2}(e)).
\end{aligned}
$$

In the last equality we have used the approximation of the gradient (13). □

6 The Numerical Scheme and Discrete A Priori Estimates

To approximate the problem (1) we use (15) for the Laplacian approximation at the interior nodes. The Dirichlet condition is imposed strongly by injection (see e.g. [15, 16]). The Neumann and Robin conditions are imposed weakly in the same way as in [7]. That is, by replacing the last term of (14) with the boundary data, we approximate the Neumann and Robin boundaries by:

$$
\int_{\partial V_i\cap\partial\Omega} \nabla u\cdot\boldsymbol{n}\, ds \approx \begin{cases} \frac{1}{2}\sum_{e\in E_i} g_i^N|\hat{\boldsymbol{b}}(e)| & \text{if } i \text{ is a Neumann boundary node,}\\ \frac{1}{2}\sum_{e\in E_i}(g_i^R-\alpha u_i)|\hat{\boldsymbol{b}}(e)| & \text{if } i \text{ is a Robin boundary node.}\end{cases}
$$

Remark 6.1 Imposing the Dirichlet condition by injection means in practice that the Dirichlet nodes are overwritten by the exact boundary data after each time step. (Equivalently, no equation is solved at these nodes, since u is equal to the boundary data.)

Remark 6.2 A boundary node is either of Dirichlet, Neumann or Robin type. The entire dual-cell boundary coinciding with the physical boundary is subsequently approximated as the same type as the boundary node, see Fig. 4. This means that in the junction between two boundary types, part of the computational boundary may be approximated as something different than the actual physical boundary. However, this is an $\mathcal{O}(h)$ error of the boundary integral which tends to zero with decreasing mesh sizes. Note that this is only necessary for the Dirichlet nodes where the boundary conditions are injected. For Neumann and Robin nodes, we could split the outer dual-cell boundary into a Neumann and a Robin part since these boundary conditions are set weakly. However, in the scheme (19) below, we use the first approach to reduce notation.

The above choices lead to the following discrete approximation scheme of (1)

$$
\begin{aligned}
\frac{dv_i}{dt} &= \left(g_i^D\right)_t, && i\in\partial\Omega_h^D,\\
\frac{dv_i}{dt} &= \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h v^n\cdot\hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i}\sum_{e\in E_i} g_i^N|\hat{\boldsymbol{b}}(e)|, && i\in\partial\Omega_h^N,\\
\frac{dv_i}{dt} &= \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h v^n\cdot\hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i}\sum_{e\in E_i}(g_i^R-\alpha u_i)|\hat{\boldsymbol{b}}(e)|, && i\in\partial\Omega_h^R,\\
\frac{dv_i}{dt} &= \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h v^n\cdot\hat{\boldsymbol{n}}_i^n, && i\in\Omega_h,\\
v_i|_{t=0} &= f_i, && i\in\Omega_h.
\end{aligned}
\tag{19}
$$

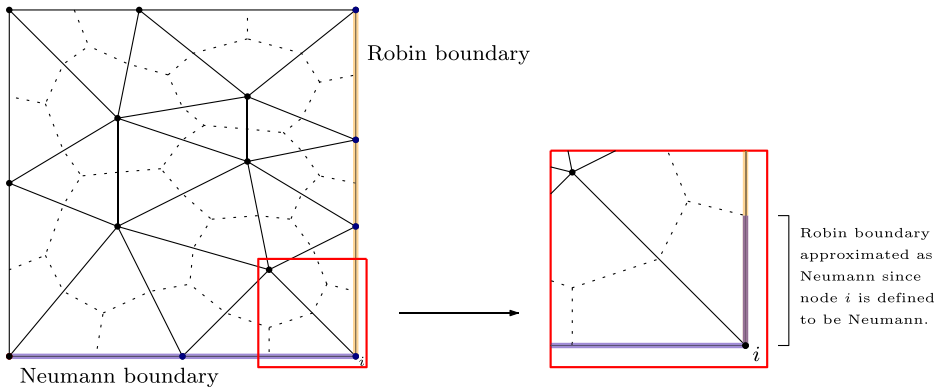


Fig. 4 Example of a grid with corresponding dual cells with an intersection of a Neumann and Robin boundary. For boundary nodes, the whole dual cell boundary is approximated as the type of the boundary node

Remark 6.3 For readers familiar with the simultaneous approximation term (SAT) (see e.g. the review papers [8, 30]) we remark that the schemes for the Neumann and Robin nodes are equivalent to

$$\begin{aligned}
\frac{dv_i}{dt} &= \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h v^n\cdot\hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i}\sum_{e\in E_i}\mu\nabla_h u^{n(e)}\cdot\hat{\boldsymbol{b}}(e) + \mathsf{SAT}_i^N, \quad i\in\partial\Omega_h^N,\\
\frac{dv_i}{dt} &= \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h v^n\cdot\hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i}\sum_{e\in E_i}\mu\nabla_h u^{n(e)}\cdot\hat{\boldsymbol{b}}(e) + \mathsf{SAT}_i^R, \quad i\in\partial\Omega_h^R,
\end{aligned}$$

where the SATs take the form

$$\begin{aligned}
\mathsf{SAT}_i^N &= -\frac{1}{2V_i}\sum_{e\in E_i}\left(\mu\nabla_h u^{n(e)}\cdot\hat{\boldsymbol{b}}(e) - g_i^N|\hat{\boldsymbol{b}}(e)|\right),\\
\mathsf{SAT}_i^R &= -\frac{1}{2V_i}\sum_{e\in E_i}\left(\mu\nabla_h u^{n(e)}\cdot\hat{\boldsymbol{b}}(e) - (g_i^R-\alpha u_i)|\hat{\boldsymbol{b}}(e)|\right).
\end{aligned} \tag{20}$$

Remark 6.4 To simplify following energy analysis, we have defined the Dirichlet nodes in (19) as $(v_i)_t = (g_i^D)_t$. We emphasise that when implementing the scheme, the Dirichlet nodes should take the form $v_i = g_i^D$ in order to avoid discretisation errors from the time-stepping algorithm.

As for the continuous problem, we transform the scheme (19) into one that imposes homogeneous Dirichlet boundary conditions. That is, we construct a function w as defined in Sect. 2 and introduce $u = v - w$ (see again [1, 17]). Inserting $v = u + w$ into the scheme (19), we obtain

$$\frac{du_i}{dt} = 0, \qquad i\in\partial\Omega_h^D \tag{21a}$$

$$\frac{du_i}{dt} = \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h u^n\cdot\hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i}\sum_{e\in E_i} g_i^N|\hat{\boldsymbol{b}}(e)| + F_i, \qquad i\in\partial\Omega_h^N \tag{21b}$$

$$\frac{du_i}{dt} = \frac{1}{2V_i}\sum_{n\in N_i}\mu\nabla_h u^n\cdot\hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i}\sum_{e\in E_i} (g_i^R-\alpha u_i)|\hat{\boldsymbol{b}}(e)| + F_i, \qquad i\in\partial\Omega_h^R \tag{21c}$$

$$\frac{du_i}{dt} = \frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + F_i \qquad i \in \Omega_h, \tag{21d}$$

$$u_i|_{t=0} = 0, \qquad i \in \Omega_h, \tag{21e}$$

where

$$F_i = \begin{cases} \frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h w^n \cdot \hat{\boldsymbol{n}}_i^n - \frac{dw_i}{dt}, & i \in \partial\Omega_h^N, \\ \frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h w^n \cdot \hat{\boldsymbol{n}}_i^n - \frac{1}{2V_i}\sum_{e\in E_i} \alpha w_i |\hat{\boldsymbol{b}}(e)| - \frac{dw_i}{dt}, & i \in \partial\Omega_h^R, \\ \frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h w^n \cdot \hat{\boldsymbol{n}}_i^n - \frac{dw_i}{dt}, & i \in \Omega_h. \end{cases} \tag{22}$$

Remark 6.5 By the Picard–Lindelöf theorem (see e.g. [25]), the ordinary differential equation (21) has a solution if the scheme is stable.

To obtain a priori estimates for the approximate solution $u^h = (u_1, u_2, \ldots, u_I)$, we use the discrete energy method (see e.g. [17] for more details on the energy method). That is, we multiply the scheme (21) in each node, i, by $u_i V_i$ and sum over all grid points.

$$\begin{aligned} \sum_{i\in\bar{\Omega}_h^V} u_i V_i \frac{du_i}{dt} = &\underline{\sum_{i\in\Omega_h^V} u_i V_i \left[\frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n\right]} \\ &+ \sum_{i\in\partial\Omega_h^N} u_i V_i \left[\underline{\frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n} + \frac{1}{2V_i}\sum_{e\in E_i} g_i^N |\hat{\boldsymbol{b}}(e)|\right] \\ &+ \sum_{i\in\partial\Omega_h^R} u_i V_i \left[\underline{\frac{1}{2V_i}\sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n} + \frac{1}{2V_i}\sum_{e\in E_i} (g_i^R - \alpha u_i)|\hat{\boldsymbol{b}}(e)|\right] \\ &+ \sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i. \end{aligned}$$

Since $\bar{\Omega}_h = \Omega_h^V \cup \partial\Omega_h^D \cup \partial\Omega_h^N \cup \partial\Omega_h^R$, and all the sets are disjoint, and since the scheme for the Dirichlet nodes is zero, the underlined terms amount to summing over all nodes in the grid. That is, the above is equivalent to

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\sum_{i\in\bar{\Omega}_h^V} V_i u_i^2 = &\frac{1}{2}\sum_{i\in\bar{\Omega}_h^V} u_i \sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2}\sum_{i\in\partial\Omega_h^N} u_i \sum_{e\in E_i} g_i^N |\hat{\boldsymbol{b}}(e)| \\ &+ \frac{1}{2}\sum_{i\in\partial\Omega_h^R} u_i \sum_{e\in E_i} (g_i^R - \alpha u_i)|\hat{\boldsymbol{b}}(e)| + \sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i. \end{aligned}$$

Using Theorem 5.1, we obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\sum_{i\in\bar{\Omega}_h} V_i u_i^2 = &-\sum_{n\in\bar{\Omega}_h^K} \nabla_h u^n \cdot \mu \nabla_h u^n |K_n| + \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} u_i g_i^N (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \\ &+ \sum_{i\in\partial\Omega_h^R} \tfrac{1}{2}\left(u_i g_i^R - \alpha u_i^2\right)(|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) + \sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i, \end{aligned}$$

$$\leq -\mu \sum_{n\in\bar{\Omega}_h^K} |\nabla_h u^n|^2 |K_n| + \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} u_i g_i^N (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|)$$
$$+ \sum_{i\in\partial\Omega_h^R} \tfrac{1}{2} u_i (g_i^R (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|)$$
$$+ \sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i, \tag{23}$$

where we in the last inequality have used that $\sum_{i\in\partial\Omega_h^R} -\frac{1}{2}\alpha u_i^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \leq 0$ since $\alpha \geq 0$. We can further manipulate the Neumann boundary terms as follows

$$\sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} u_i g_i^N (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \leq \sum_{i\in\partial\Omega_h^N} |u_i g_i^N| (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|).$$

Using Young's inequality, we obtain

$$\sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} u_i g_i^N (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \leq \tfrac{\beta}{2} \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} |u_i|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|)$$
$$+ \tfrac{1}{2\beta} \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} |g_i^N|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|).$$

The Robin boundary terms can be manipulated the same way. Thus, (23) reads

$$\begin{aligned}\frac{1}{2}\frac{d}{dt} \sum_{i\in\bar{\Omega}_h^V} V_i u_i^2 \leq & -\mu \sum_{n\in\bar{\Omega}_h^K} |\nabla_h u^n|^2 |K_n| + \tfrac{\beta}{2} \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} |u_i|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \\ & + \tfrac{1}{2\beta} \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2} |g_i^N|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \\ & + \tfrac{\beta}{2} \sum_{i\in\partial\Omega_h^R} \tfrac{1}{2} |u_i|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \\ & + \tfrac{1}{2\beta} \sum_{i\in\partial\Omega_h^R} \tfrac{1}{2} |g_i^R|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) + \sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i.\end{aligned} \tag{24}$$

We introduce the following notation for the discrete equivalents of the L^2-norms:

$$\|u^h\|^2_{L^2_V(\Omega)} = \sum_{i\in\bar{\Omega}_h^V} |u_i|^2 V_i, \tag{25}$$

$$\|\nabla_h u^h\|^2_{L^2_K(\Omega)} = \sum_{n\in\bar{\Omega}_h^K} |\nabla_h u^n|^2 |K_n|, \tag{26}$$

$$\|u^h\|^2_{L^2_B(\partial\Omega_h)} = \sum_{i\in\partial\Omega_h^B} \tfrac{1}{2} |u_i|^2 (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|). \tag{27}$$

Using the definitions (25)–(27), we can recast (24) as

$$\begin{aligned}\frac{d}{dt}\|u^h\|^2_{L^2_V(\Omega)} \leq & -\mu \|\nabla_h u^h\|^2_{L^2_K(\Omega)} + \tfrac{\beta}{2} \|u^h\|^2_{L^2_B(\partial\Omega_h^N)} + \tfrac{1}{2\beta} \|g^{N,h}\|^2_{L^2_B(\partial\Omega_h^N)} \\ & + \tfrac{\beta}{2} \|u^h\|^2_{L^2_B(\partial\Omega_h^R)} + \tfrac{1}{2\beta} \|g^{R,h}\|^2_{L^2_B(\partial\Omega_h^R)} + \sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i.\end{aligned}$$

To obtain an estimate analogous to (8), we must consider the forcing term $\sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i$. Except for the time-derivative term in (22), F_i takes the same form as the right-hand side of the scheme (21). By using the SBP property from Theorem 5.1 and Young's inequality, we obtain

$$\sum_{i\in\bar{\Omega}_h^V} u_i V_i F_i \le \tfrac{\epsilon}{2}\mu\|\nabla_h u^h\|^2_{L^2_K(\Omega)} + \tfrac{\mu}{2\epsilon}\|\nabla_h w^h\|^2_{L^2_K(\Omega)} + \tfrac{\beta}{2}\|u^h\|^2_{L^2_B(\partial\Omega_h^R)} + \tfrac{\alpha^2}{2\beta}\|w^h\|^2_{L^2_B(\partial\Omega_h^R)} + \tfrac{\delta}{2}\|u^h\|^2_{L^2_V(\Omega_h)} + \tfrac{1}{2\delta}\|w_t^h\|^2_{L^2_V(\Omega)}.$$

Thus, we have

$$\begin{aligned}\frac{d}{dt}\|u^h\|^2_{L^2_V(\Omega)} \le{}& -\mu\|\nabla_h u^h\|^2_{L^2_K(\Omega)} + \tfrac{\beta}{2}\|u^h\|^2_{L^2_B(\partial\Omega_h^N)} + \tfrac{1}{2\beta}\|g^{N,h}\|^2_{L^2_B(\partial\Omega_h^N)} + \tfrac{\beta}{2}\|u^h\|^2_{L^2_B(\partial\Omega_h^R)} \\ &+ \tfrac{1}{2\beta}\|g^{R,h}\|^2_{L^2_B(\partial\Omega_h^R)} \\ &+ \tfrac{\epsilon}{2}\mu\|\nabla_h u^h\|^2_{L^2_K(\Omega)} + \tfrac{\mu}{2\epsilon}\|\nabla_h w^h\|^2_{L^2_K(\Omega)} + \tfrac{\beta}{2}\|u^h\|^2_{L^2_B(\partial\Omega_h^R)} \\ &+ \tfrac{\alpha^2}{2\beta}\|w^h\|^2_{L^2_B(\partial\Omega_h^R)} + \tfrac{\delta}{2}\|u^h\|^2_{L^2_V(\Omega)} + \tfrac{1}{2\delta}\|w_t^h\|^2_{L^2_V(\Omega)}.\end{aligned} \tag{28}$$

Similarly as for the continuous problem, if we choose $\epsilon = 1$, we obtain $-\mu\|\nabla_h u^h\|^2_{L^2_K(\Omega)} + \frac{\epsilon}{2}\mu\|\nabla_h u^h\|^2_{L^2_K(\Omega)} = -\frac{\mu}{2}\|\nabla_h u^h\|^2_{L^2_K(\Omega)}$ in (28). Furthermore, $\beta\|u^h\|^2_{L^2_B(\partial\Omega_h^N)} + \frac{\beta}{2}\|u^h\|^2_{L^2_B(\partial\Omega_h^R)} \lesssim \beta\|u^h\|^2_{L^2_B(\partial\Omega)}$. Hence, we have

$$\begin{aligned}\frac{d}{dt}\|u^h\|^2_{L^2_V(\Omega)} \lesssim{}& -\tfrac{\mu}{2}\|\nabla_h u^h\|^2_{L^2_K(\Omega)} + \tfrac{\mu}{2}\|\nabla_h w^h\|^2_{L^2_K(\Omega)} \\ &+ \tfrac{\delta}{2}\|u^h\|^2_{L^2_V(\Omega)} + \tfrac{1}{2\delta}\|w_t^h\|^2_{L^2_V(\Omega)} + \beta\|u^h\|^2_{L^2_B(\partial\Omega)} \\ &\tfrac{1}{2\beta}\left(\|g^{N,h}\|^2_{L^2_B(\partial\Omega_h^N)} + \|g^{R,h}\|^2_{L^2_B(\partial\Omega_h^R)} + \|w^h\|^2_{L^2_B(\partial\Omega_h^R)}\right).\end{aligned}$$

Finally using the trace theorem, we arrive at a similar estimate as in (8):

$$\begin{aligned}&\frac{d}{dt}\|u^h\|^2_{L^2_V(\Omega)} + \tfrac{\mu}{2}\|\nabla_h u^h\|^2_{L^2_K(\Omega)} - \tfrac{\delta}{2}\|u^h\|^2_{L^2_V(\Omega)} - \beta\|u^h\|^2_{H^1_K(\Omega)} \\ &\quad\lesssim \tfrac{\mu}{2}\|\nabla_h w^h\|^2_{L^2_K(\Omega)} + \tfrac{1}{2\delta}\|w_t^h\|^2_{L^2_V(\Omega)} \\ &\qquad + \tfrac{1}{2\beta}\left(\|g^{N,h}\|^2_{L^2_B(\partial\Omega_h^N)} + \|g^{R,h}\|^2_{L^2_B(\partial\Omega_h^R)} + \|w^h\|^2_{H^1_K(\Omega)}\right).\end{aligned}$$

Note that we have arrived at a semi-discrete equivalent of (8). Thus, by using Grönnwall's inequality followed by integration in time, as done in Sect. 3, we obtain $u^h \in L^\infty(0,T;L^2_V(\Omega))$ and $\nabla_h u^h \in L^2(0,T;L^2_K(\Omega))$. We may extend the numerical solution, u^h, to the entire domain by a linear interpolation on the triangles. Let u_c^h denote this continuous piecewise linear function. We have that $\nabla_h u_c^h = \nabla u_c^h = \nabla_h u^h$. Hence $\nabla u_c^h \in L^2(0,T;L^2_K(\Omega))$ (and also, $\nabla u_c^h \in L^2(0,T;L^2(\Omega))$ since ∇u_c^h is piecewise constant). Furthermore, the norm $\|u_c^h\|^2_{L^2(\Omega)}$ can be bounded by $\|u^h\|^2_{L^2_V(\Omega)}$. Thus, we have $u_c^h \in L^2(0,T;H^1(\Omega))$.

7 Weak Convergence to a Weak Solution

Let $\phi \in H^1(0,T; C^\infty_{\partial\Omega_0^D}(\bar{\Omega}))$. Since ϕ is smooth (in space), $\phi|_{V_i}$, which is the restriction of ϕ to a dual cell, can be written as $\phi|_{V_i} = \phi(x_i, y_i, t) + hp_i = \phi_i + hp_i$, where h is a characteristic mesh size and $p_i(x_i, y_i, t)$ is a function of size $\mathcal{O}(1)$. The gradient approximation (13) is $\nabla_h \phi|_{K_n} = \nabla\phi|_{K_n} + \mathcal{O}(h)$. This can easily be checked for equilateral triangles. Thereafter, one can prove the relation for a general triangle by transforming it to an equilateral one using a linear transformation.

We denote the right-hand side of the scheme (21) by $L_h u^h$. To prove convergence to a weak solution, we test the numerical scheme (21) against ϕ. That is, we calculate

$$\begin{aligned}\int_\Omega \phi u_t^h \, d\mathbf{x} &= \int_\Omega \phi (L_h u^h) \, d\mathbf{x}, \\ &= \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} \phi|_{V_i} (L_h u^h)|_{V_i} \, d\mathbf{x}.\end{aligned} \tag{29}$$

We now use that $\phi|_{V_i} = \phi_i + hp_i$ to obtain

$$\int_\Omega \phi u_t^h \, d\mathbf{x} = \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} (\phi_i + hp_i)(L_h u^h)_i \, d\mathbf{x} \tag{30}$$

$$= \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} \phi_i (L_h u^h)_i \, d\mathbf{x} + \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} hp_i (L_h u^h)_i \, d\mathbf{x}$$

$$= \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} \phi_i (L_h u^h)_i \, d\mathbf{x} + \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} hp_i (u_i)_t \, d\mathbf{x}, \tag{31}$$

where we have used $u_t^h = L_h u^h$ in the last step. Thus

$$\int_\Omega (\phi - hp) u_t^h \, d\mathbf{x} = \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} \phi_i (L_h u^h)_i \, d\mathbf{x}. \tag{32}$$

Inserting the specific form of $L_h u^h$ (that is, the right-hand side of the scheme (21)) yields

$$\begin{aligned}\int_\Omega (\phi - hp) u_t^h \, d\mathbf{x} = &\sum_{i\in\Omega_h^V} \int_{V_i} \phi_i \left[\frac{1}{2V_i} \sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n \right] d\mathbf{x} \\ &+ \sum_{i\in\partial\Omega_h^N} \int_{V_i} \phi_i \left[\frac{1}{2V_i} \sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i} \sum_{e\in E_i} g_i^N |\hat{\boldsymbol{b}}(e)| \right] d\mathbf{x} \\ &+ \sum_{i\in\partial\Omega_h^R} \int_{V_i} \phi_i \left[\frac{1}{2V_i} \sum_{n\in N_i} \mu \nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2V_i} \sum_{e\in E_i} (g_i^N - \alpha u_i) |\hat{\boldsymbol{b}}(e)| \right] d\mathbf{x} \\ &+ \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} \phi_i F_i \, d\mathbf{x}.\end{aligned} \tag{33}$$

Since ϕ_i is constant on each dual cell, V_i and the Laplacian approximation is a scalar constant, the right-hand side above can be integrated exactly, leading to

$$\begin{aligned}\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} = &\sum_{i\in\Omega_h^V} \phi_i V_i \underline{\left[\frac{1}{2V_i}\sum_{n\in N_i} \mu\nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n\right]} \\ &+ \sum_{i\in\partial\Omega_h^N} \phi_i V_i \left[\underline{\frac{1}{2V_i}\sum_{n\in N_i} \mu\nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n} + \frac{1}{2V_i}\sum_{e\in E_i} g_i^N |\hat{\boldsymbol{b}}(e)|\right] \\ &+ \sum_{i\in\partial\Omega_h^R} \phi_i V_i \left[\underline{\frac{1}{2V_i}\sum_{n\in N_i} \mu\nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n} + \frac{1}{2V_i}\sum_{e\in E_i} (g_i^N - \alpha u_i)|\hat{\boldsymbol{b}}(e)|\right] \\ &+ \sum_{i\in\bar{\Omega}_h^V} \phi_i V_i F_i .\end{aligned} \tag{34}$$

As in the discrete analysis in Sect. 6, the underlined terms can be written as the sum over all grid points in $\bar{\Omega}_h$ as follows

$$\begin{aligned}\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} = &\frac{1}{2}\sum_{i\in\bar{\Omega}_h^V} \phi_i \sum_{n\in N_i} \mu\nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2}\sum_{i\in\partial\Omega_h^N} \phi_i \sum_{e\in E_i} g_i^N |\hat{\boldsymbol{b}}(e)| \\ &+ \frac{1}{2}\sum_{i\in\partial\Omega_h^R} \phi_i \sum_{e\in E_i} (g_i^R - \alpha u_i)|\hat{\boldsymbol{b}}(e)| + \sum_{i\in\bar{\Omega}_h^V} \phi_i V_i F_i .\end{aligned} \tag{35}$$

Using the SBP properties from Theorem 5.1 yields

$$\begin{aligned}\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} = &-\sum_{n\in\bar{\Omega}_h^K} \nabla_h\phi^n \cdot \mu\nabla_h u^n |K_n| + \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2}\phi_i g_i^N (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \\ &+ \sum_{i\in\partial\Omega_h^R} \tfrac{1}{2}\phi_i (g_i^R - \alpha u_i)(|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) + \sum_{i\in\bar{\Omega}_h^V} \phi_i V_i F_i .\end{aligned} \tag{36}$$

Since $\nabla_h\phi^n$ and $\nabla_h u^n$ are constant on each triangle K, we have that $-\sum_{n\in\bar{\Omega}_h^K} \nabla_h\phi^n \cdot \mu\nabla_h u^n |K_n| = -\sum_{n\in\bar{\Omega}_h^K} \int_{K_n} \nabla_h\phi^h \cdot \mu\nabla_h u^h \, d\mathbf{x}$. Thus, (36) can be written as

$$\begin{aligned}\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} &= -\sum_{n\in\bar{\Omega}_h^K} \int_{K_n} \nabla_h\phi^n \cdot \mu\nabla_h u^n \, d\mathbf{x} \\ &\quad + \sum_{i\in\partial\Omega_h^N} \tfrac{1}{2}\phi_i g_i^N (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,1}(e)|) \\ &\quad + \sum_{i\in\partial\Omega_h^R} \tfrac{1}{2}\phi_i (g_i^R - \alpha u_i)(|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,1}(e)|) \\ &\quad + \sum_{i\in\bar{\Omega}_h^V} \int_{V_i} \phi_i F_i \, d\mathbf{x}, \\ &= -\int_\Omega \nabla_h\phi^h \cdot \mu\nabla_h u^h \, d\mathbf{x}\end{aligned}$$

$$+\int_\Omega \phi^h F^h \, d\mathbf{x} + \int_{\partial\Omega_h^N} \phi^h g^{N,h} \, ds + \int_{\partial\Omega_h^R} \phi^h (g^{R,h} - \alpha u^h) \, ds.$$

Partial integration in time yields

$$\begin{aligned}\int_0^T \int_\Omega (\phi - hp)_t u^h \, d\mathbf{x}dt = &\int_0^T \int_\Omega \nabla_h \phi^h \cdot \mu \nabla_h u^h \, d\mathbf{x}dt - \int_0^T \int_{\partial\Omega_h^N} \phi^h g^{N,h} \, dsdt \\ &- \int_0^T \int_{\partial\Omega_h^R} \phi^h (g^{R,h} - \alpha u^h) \, dsdt - \int_0^T \int_\Omega \phi^h F^h \, d\mathbf{x}dt \\ &- \int_\Omega hpu^h(T) \, d\mathbf{x},\end{aligned} \tag{37}$$

where we have used $u^h|_{t=0} = 0$ and $\phi^h|_{t=T} = 0$.

Remark 7.1 Here, $\int_\Omega \phi^h F^h \, d\mathbf{x}$ is the short-hand for the semi-discrete form of (12). By using the SBP property from Theorem 5.1, it can be written as

$$\begin{aligned}\int_\Omega \phi^h F^h \, d\mathbf{x} &= \sum_{i \in \bar{\Omega}_h^V} \phi_i V_i F_i, \\ &= -\sum_{n \in \bar{\Omega}_h^K} \nabla_h \phi^n \cdot \mu \nabla_h w^n |K_n| - \sum_{i \in \partial\Omega_h^R} \alpha \phi_i w_i (|\hat{\boldsymbol{b}}_{i,1}(e)| + |\hat{\boldsymbol{b}}_{i,2}(e)|) \\ &\quad - \sum_{i \in \bar{\Omega}_h^V} \phi_i V_i \frac{dw_i}{dt}, \\ &= -\int_\Omega \nabla_h \phi^h \cdot \mu \nabla_h w^h \, d\mathbf{x} - \int_{\partial\Omega_h^R} \alpha \phi^h w^h \, ds - \int_\Omega \phi^h w_t^h \, ds.\end{aligned} \tag{38}$$

We keep the symbolic expression to reduce notation.

Since $\phi|_{V_i} = \phi^h|_{V_i} + hp_i$ and $\nabla_h \phi|_{K_n} = \nabla\phi + \mathcal{O}(h)$, the weak formulation (37) becomes

$$\begin{aligned}\int_0^T \int_\Omega (\phi_t - hp_t)u^h \, d\mathbf{x}dt = &\int_0^T \int_\Omega (\nabla\phi + \mathcal{O}(h)) \cdot \mu \nabla_h u^h \, d\mathbf{x}dt \\ &- \int_0^T \int_{\partial\Omega_h^N} (\phi - hp) g^{N,h} \, dsdt \\ &- \int_0^T \int_{\partial\Omega_h^R} (\phi - hp)(g^{R,h} - \alpha u^h) \, dsdt \\ &- \int_0^T \int_\Omega (\phi - hp) F^h \, d\mathbf{x}dt \\ &- \int_\Omega hpu^h(T) \, d\mathbf{x}.\end{aligned} \tag{39}$$

We utilise the following functional analysis theorem (see e.g. [10], and [5] for a proof).

Theorem 7.2 *Let $\Omega_T \subset \mathbb{R}^n$ be an open domain and let $\{u_n\} \in L^2(\Omega_T)$ be a bounded sequence. Then there exists a subsequence, $\{u_{n_i}\} \in L^2(\Omega_T)$ that converges weakly to $\bar{u} \in L^2(\Omega_T)$. That is,*

$$\int_{\Omega_T} \phi u_{n_i} \, d\mathbf{x} \to \int_{\Omega_T} \phi \bar{u} \, d\mathbf{x} \quad as n_i \to \infty, for all \phi \in L^2(\Omega_T).$$

Here, we take $\Omega_T = \Omega \times [0, T]$. Consider the $\mathcal{O}(1)$ term on the left-hand side of (39). Using Theorem 7.2, we have that

$$\int_0^T \int_\Omega \phi_t u^h \, d\mathbf{x}dt \to \int_0^T \int_\Omega \phi_t \bar{u} \, d\mathbf{x}dt.$$

The other $\mathcal{O}(1)$ terms can be treated in a similar way. Turning to the second term in (39), we have

$$\int_0^T \int_\Omega h p_t u^h \, d\mathbf{x}dt \to 0,$$

as $h \to 0$, since $u^h \in L^\infty(0, T; L^2(\Omega))$. Using the available bounds, similar arguments imply that $\mathcal{O}(h)\mu\nabla_h u^h$, $hpg^{N,h}$, $hpg^{R,h}$, hpu^h and hpF^h vanish.

Remark 7.3 Since all terms in F (see (38)) are known and bounded in $L^2(\Omega_T)$ (see the assumptions in Sect. 2), the weak convergence of the symbolic expression (38) follows trivially.

In summary, letting $h \to 0$, (39) becomes

$$\begin{aligned}\int_0^T \int_\Omega \phi_t \bar{u} \, d\mathbf{x}dt = &\int_0^T \int_\Omega \nabla\phi \cdot \mu\overline{\nabla u} \, d\mathbf{x}dt \\ &- \int_0^T \int_{\partial\Omega^N} \phi \bar{g}^N \, dsdt - \int_0^T \int_{\partial\Omega} \phi(\bar{g}^R - \alpha\bar{u}) \, dsdt \\ &- \int_0^T \int_\Omega \phi \bar{F} \, d\mathbf{x}dt, \end{aligned} \tag{40}$$

which is satisfied for all $\phi \in H^1(0, T; C^\infty_{\partial\Omega_0^D}(\bar{\Omega}))$.

Remark 7.4 Note that the boundary integrals over the computational boundaries converge to the integrals over the physical boundaries as $h \to 0$. That is,

$$\int_{\partial\Omega_h^N} \phi^h g^{N,h} \, ds \to \int_{\partial\Omega^N} \phi^h g^{N,h} \, ds \quad \text{and}$$
$$\int_{\partial\Omega_h^R} \phi^h (g^{R,h} - \alpha u^h) \, ds \to \int_{\partial\Omega^R} \phi^h (g^{R,h} - \alpha u^h) \, ds,$$

as $h \to 0$.

Remark 7.5 The term $\int_\Omega \phi_t u^h \, d\mathbf{x}$ in (39) satisfies

$$\int_\Omega \phi_t u^h \, d\mathbf{x} = \int_\Omega \phi_t u_c^h \, d\mathbf{x} + \mathcal{O}(h),$$

(this can be verified by using the specific form of u_c^h on each triangle). Since $u_c^h \in H^1(\Omega)$, Theorem 7.2 gives

$$\int_\Omega \phi_t u_c^h \, d\mathbf{x} \to \int_\Omega \phi_t \bar{u}_c^h \, d\mathbf{x},$$

in $H^1(\Omega)$. Thus, $\overline{\nabla u} = \nabla\bar{u}$ in (40).

Theorem 7.6 *Equation (40) holds for all* $\phi \in H^1(0, T; H^1_{\partial\Omega_0^D}(\Omega))$.

Proof Since the space $H^1(0,T;C^\infty_{\partial\Omega_0^D}(\bar{\Omega}))$ is dense in $H^1(0,T;H^1_{\partial\Omega_0^D}(\Omega))$ (see [1]), the equality (40) holds for all $\phi \in H^1(0,T;H^1_{\partial\Omega_0^D}(\Omega))$. □

Hence, $\bar{u}$ is a weak solution of the problem (3).

8 Strong Convergence to a Weak Solution

Next, we prove strong convergence to the weak solution.

Definition 8.1 (*Strong convergence*, [10]) A sequence $\{u_n\}_{n=1}^\infty \subset X$ is said to converge to $u \in X$, i.e., $u_n \to u$, if $\lim_{n\to\infty}\|u_n - u\|_X = 0$.

We also need the following definition.

Definition 8.2 [25, Definition 6.76] We say that a domain, $\Omega \subset \mathbb{R}^d$, has the k-extension property if there exists a bounded linear mapping $E : H^k(\Omega) \to H^k(\mathbb{R}^d)$ such that $Eu|_\Omega = u$ for every $u \in H^k(\Omega)$.

As we have assumed the spatial domain to be Lipschitz, the following result applies.

Theorem 8.3 (see e.g. [1] or [27]) *Any Lipschitz domain has the k-extension property.*

For a bounded domain, Ω, with the k-extension property, we have that $H^1(\Omega)$ is compactly embedded in $L^2(\Omega)$ (see e.g. [25]), which in turn is continuously embedded in $H^{-1}(\Omega)$. To prove strong convergence, we need the Aubin–Lions Lemma:

Lemma 8.4 (Aubin–Lions, see e.g. [26]) *Let X, B and Y be Banach spaces such that $X \subset B \subset Y$, where the embedding, $X \subset B$ is compact and $B \subset Y$ is continuous. Let $U = \{u \in L^p(0,T;X) \mid u_t \in L^q(0,T;Y)\}$, $1 \le p,q < \infty$. Then U is compactly embedded in $L^p(0,T;B)$.*

A Banach space X is compactly embedded in another Banach space Y, if the following two conditions hold (see [10]):

(i) $\|u\|_Y \le C\|u\|_X$, $(u \in X)$, for some constant C.
(ii) each bounded sequence in X is precompact in Y, i.e., for a bounded sequence $\{u_n\}_{n=1}^\infty$, there exists a subsequence, $\{u_{n_i}\}_{n_i=1}^\infty \subseteq \{u_n\}_{n=1}^\infty$ that converges to a $\bar{u}$ in Y.

Herein, we use $X = H^1(\Omega)$, $B = L^2(\Omega)$ and $Y = H^{-1}(\Omega)$ in Lemma 8.4. Thus, since we have $u_c^h \in L^2(0,T,H^1(\Omega))$, it suffices to show that $(u_c^h)_t \in L^1(0,T;H^{-1}(\Omega))$ to establish the strong convergence. That is, we need to show that the norm (see e.g. [25])

$$\|(u_c^h)_t\|_{L^1(0,T;H^{-1}(\Omega))} = \int_0^T \sup_{\substack{\phi\in H_0^1(\Omega),\\ \|\phi\|_{H_0^1(\Omega)}=1}} \int_\Omega (u_c^h)_t \phi \, d\mathbf{x}dt, \tag{41}$$

is bounded. To this end, we test the scheme (21) with a function $\phi \in C_0^\infty(\bar{\Omega})$.

$$\int_\Omega \phi u_t^h \, dx = \int_\Omega \phi(L_h u^h)\, d\mathbf{x} = \sum_{i\in\bar{\Omega}_h} \int_{V_i} \phi|_{V_i}(L_h u^h)|_{V_i}\, d\mathbf{x}.$$

Note the resemblance to (29) (the only difference being the function ϕ that is now vanishing on the whole boundary $\partial\Omega$). From derivations analogous to (30)–(35), we can recast the above equation to

$$\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} = \frac{1}{2}\sum_{i\in\bar{\Omega}_h^V} \phi_i \sum_{n\in N_i} \mu\nabla_h u^n \cdot \hat{\boldsymbol{n}}_i^n + \frac{1}{2}\sum_{i\in\partial\Omega_h^N} \phi_i \sum_{e\in E_i} g_i^N |\hat{\boldsymbol{b}}(e)| + \sum_{i\in\partial\Omega_h^R} \phi_i \sum_{e\in E_i} (g_i^R - \alpha u_i)|\hat{\boldsymbol{b}}(e)| + \sum_{i\in\bar{\Omega}_h} \phi_i V_i F_i.$$

By using $\phi|_{\partial\Omega} = 0$ and the SBP property (see Theorem. 5.1) we have

$$\begin{aligned}\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} &= -\sum_{n\in\bar{\Omega}_h^K} \int_{K_n} \nabla_h\phi^n \cdot \mu\nabla_h u^n \, d\mathbf{x} + \sum_{i\in\bar{\Omega}_h^V} \phi_i V_i F_i \\ &= -\int_\Omega \nabla_h\phi^h \cdot \nabla_h u^h \, d\mathbf{x} + \int_\Omega \phi^h F^h \, d\mathbf{x}. \end{aligned} \tag{42}$$

Remark 8.5 Here, $\int_\Omega \phi^h F^h \, d\mathbf{x}$ takes the same form as in Remark 7.1, except for the boundary term $\int_{\partial\Omega_h^R} \alpha\phi^h w^h \, ds$ which is zero in (42) since ϕ is vanishing on the entire boundary $\partial\Omega$ in this case.

Inserting $\phi = \phi^h + hp$ and $\nabla_h\phi = \nabla\phi + \mathcal{O}(h)$, we obtain

$$\int_\Omega (\phi - hp)u_t^h \, d\mathbf{x} = -\int_\Omega \left(\nabla\phi \cdot \mu\nabla_h u^h + \mathcal{O}(h) \cdot \mu\nabla_h u^h\right) d\mathbf{x} + \int_\Omega \left(\phi F^h - hpF^h\right) d\mathbf{x}.$$

Since $\nabla_h u^h \in L^2(0, T; L_K^2(\bar{\Omega}_h))$ and all terms of F^h are properly bounded (see the assumptions in Sect. 2), letting $h \to 0$ yields

$$\int_\Omega \phi u_t^h \, d\mathbf{x} = -\int_\Omega \nabla\phi \cdot \mu\overline{\nabla u} \, d\mathbf{x} + \int_\Omega \phi\bar{F} \, d\mathbf{x},$$

as $\lim_{h\to 0}(\phi - hp) = \phi$. By inserting the specific form of $\int_\Omega \phi\bar{F} \, d\mathbf{x}$ and using the Cauchy–Schwarz inequality, we obtain

$$\begin{aligned}\int_\Omega \phi u_t^h \, dx \le \frac{1}{2}\Big(&\|\nabla\phi\|_{L^2(\Omega)}^2 + \mu\|\overline{\nabla u}\|_{L^2(\Omega)}^2 + \|\nabla\phi\|_{L^2(\Omega)}^2 + \|\overline{\nabla w}\|_{L^2(\Omega)}^2 \\ &+\|\phi\|_{L^2(\Omega)}^2 + \|\bar{w}_t\|_{L^2(\Omega)}^2\Big). \end{aligned} \tag{43}$$

This holds for all $\phi \in C_0^\infty(\bar{\Omega})$, and by density, it follows that the inequality holds for all $\phi \in H_0^1(\Omega)$. Integration in time finally yields

$$\begin{aligned}&\int_0^T \sup_{\substack{\phi\in H_0^1(\Omega),\\ \|\phi\|_{H_0^1(\Omega)}=1}} \int_\Omega \phi u_t^h \, d\mathbf{x}dt \\ &\le \int_0^T \sup_{\substack{\phi\in H_0^1(\Omega),\\ \|\phi\|_{H_0^1(\Omega)}=1}} \frac{1}{2}\Big(\|\nabla\phi\|_{L^2(\Omega)}^2 + \mu\|\overline{\nabla u}\|_{L^2(\Omega)}^2 \\ &\quad +\|\nabla\phi\|_{L^2(\Omega)}^2 + \|\overline{\nabla w}\|_{L^2(\Omega)}^2 + \|\phi\|_{L^2(\Omega)}^2 + \|\bar{w}_t\|_{L^2(\Omega)}^2\Big) \, dt. \end{aligned}$$

Hence $u_t^h \in L^1(0,T;H^{-1}(\Omega))$, and since $(u_c^h)_t$ is u_t^h extended to the entire domain using a linear interpolant on the triangles, we also have $(u_c^h)_t \in L^1(0,T;H^{-1}(\Omega))$. Thus, by Aubin–Lions' lemma 8.4, the family of functions, $U = \{u_c^h \in L^2(0,T;H^1(\Omega)) \mid (u_c^h)_t \in L^1(0,T;H^{-1}(\Omega))\}$, is compactly embedded in $L^2(0,T;L^2(\Omega))$, meaning that u_c^h converges strongly to the weak solution.

9 Uniqueness of the Weak Solution

Assume that there are two weak solutions u, v to the problem (1) satisfying the boundary and initial data. Then $w = u - v$ is also a weak solution with homogenous data ($F = g^D = g^N = g^R \equiv 0$). Take $\phi = w$ in (10) to obtain

$$\int_\Omega w w_t \, d\mathbf{x} = \int_\Omega w(\nabla \cdot \mu \nabla w) \, d\mathbf{x}.$$

Integrating the right-hand side by parts, and using the fact that the boundary data is zero, we obtain

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|w\|^2_{L^2(\Omega)} &= -\mu\|\nabla w\|^2_{L^2(\Omega)} - \alpha\|w\|^2_{L^2(\partial\Omega^R)} \le 0, \\ \|w(\cdot,\cdot,T)\|^2_{L^2(\Omega)} &\le \|w(\cdot,\cdot,0)\|^2_{L^2(\Omega)} \equiv 0. \end{aligned}$$

Hence, $\|w\|_{L^2(0,T;L^2(\Omega))} = \|u - v\|_{L^2(0,T;L^2(\Omega))} = 0$ and thus the weak solution is unique in $L^2(0,T;L^2(\Omega))$.

10 Numerical Simulations

We implement the scheme (1) and consider the manufactured solution used in [7]. That is, the exact solution is given by

$$u(x,y,t) = e^{-8\pi^2 t}\sin(2\pi x)\sin(2\pi y) + e^{-32\pi^2 t}\sin(4\pi x)\sin(4\pi y), \tag{44}$$

which yields a zero forcing function. Furthermore, we let $\mu = \alpha = 1$. We consider a square domain $\Omega = [0,1] \times [0,1]$ containing a hole. The hole is located at $(x,y) = (0.5, 0.5)$, and has radius $r = \frac{1}{8}$. We pose Dirichlet boundary conditions on the boundary of the hole, Neumann boundary conditions on $y = 0$, $y = 1$ and Robin boundary conditions on $x = 0$, $x = 1$. The boundary data is given by (44). $t = 0.05$ is used as the final time. The scheme was run on grids containing 398, 1394, 5097, 19457 and 76166 nodes. A typical grid is depicted in Fig. 5a. All grids were generated using Gmsh (see [14]). The scheme was implemented using the Julia programming language (see [4]).

Figure 5b shows the convergence rate together with a reference line representing second-order convergence. We conclude that the scheme converges at approximately a rate of two.

11 Conclusion

Herein, we have considered a slightly modified local finite-volume approximation of the Laplacian operator proposed by Chandrashekar in [7] for discretising the heat equation in

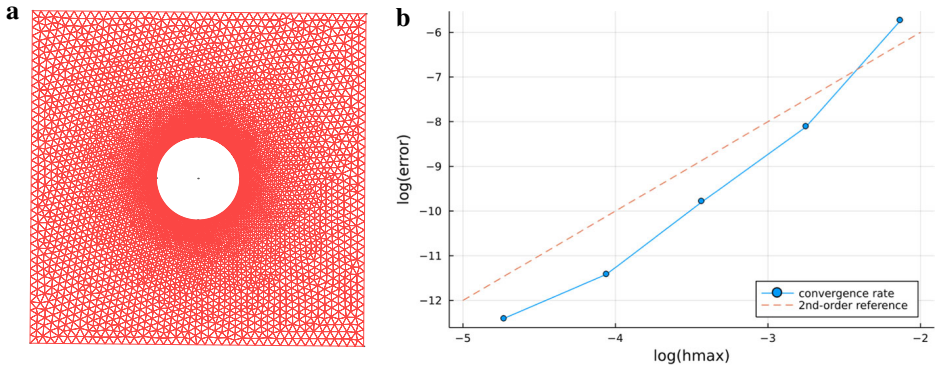


Fig. 5 **a** A typical mesh. **b** Convergence rate obtained for simulations using $N = 398, 1394, 5097, 19457, 76166$ grid points

two spatial dimensions on general triangular grids. The equation was augmented with Dirichlet, Neumann and Robin boundary conditions. The Dirichlet boundary condition was imposed strongly by injection, while the Neumann and Robin conditions were imposed weakly. We demonstrated that this modification satisfies the SBP property proved in [7]. By using the energy method, a priori estimates for the numerical solution were derived. From these estimates, we were able to prove the weak convergence of the numerical solution to a weak solution of the heat equation. Thus, consistency, in a weak sense, of the Laplacian operator was established. Subsequently, we demonstrated that the numerical solution converges strongly to a weak solution by using Aubin–Lions' lemma. Finally, the weak solution was shown to be unique. To the best of our knowledge, this is the first proof of convergence for a local finite-volume method for the Laplacian on general triangular grids. The theory presented here is straightforwardly applicable to three spatial dimensions, provided that the Laplacian approximation can be generalised to such domains.

A numerical simulation, which included Dirichlet, Neumann and Robin conditions was run on an unstructured triangulated grid containing a hole. By using the method of manufactured solutions, we demonstrated that the numerical solution converged with a second-order rate.

Funding Open access funding provided by University of Bergen (incl Haukeland University Hospital) The authors have not disclosed any funding.

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare that they have no known competing interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Atkinson, K., Han, W.: Theoretical numerical analysis, a functional analaysis framework. In: Texts in Applied Mathematics, 2 edn. Springer Science+Business Media Inc. (2005). ISBN 978-0387-25887-4
2. Bause, M., Hoffmann, J., Knabner, P.: First-order convergence of multi-point flux approximation on triangular grids and comparison with mixed finite element methods. Numer. Math. **116**, 1–29 (2010). https://doi.org/10.1007/s00211-010-0290-y
3. Bauzet, C., Nabet, F., Schmitz, K., Zimmermann, A.: Convergence of a finite-volume scheme for a heat equation with a multiplicative Lipschitz noise (2022). arXiv:2203.09851v1 [math.AP]
4. Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B.: Julia: a fresh approach to numerical computing. SIAM Rev **590**(1), 65–98 (2017)
5. Brezis, H.: Functional analysis, Sobolev spaces and partial differential equations. Springer (2011). https://doi.org/10.1007/978-0-387-70914-7
6. Chan, J., Del Rey Fernández, D.C., Carpenter, M.H.: Efficient entropy stable Gauss collocation methods. SIAM J. Sci. Comput. **410**(5), A2938–A2966 (2019). https://doi.org/10.1137/18M1209234
7. Chandrashekar, P.: Finite volume discretization of heat equation and compressible Navier–Stokes equations with weak Dirichlet boundary condition on triangular grids. Int. J. Adv. Eng. Sci. Appl. Math. **80**(3), 174–193 (2016). https://doi.org/10.1007/s12572-015-0160-z
8. Del Rey Fernández D.C., Hicken J.E., Zingg D.W.: Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations. Comput. Fluids pp. 171–196 (2014). https://doi.org/10.1016/j.compfluid.2014.02.016
9. Del Rey Fernández, D.C., Carpenter, M.H., Dalcin, L., Zampini, S., Parsani, M.: Entropy stable h/p-nonconforming discretization with the summation-by-parts property for the compressible Euler and Nacier–Stokes equations. SN Partial Differ. Equ. Appl. (2020). https://doi.org/10.1007/s42985-020-00009-z
10. Evans, L.C.: Partial Differential Equations, Volume 19 of Graduate Studies in Mathematics, 2 edn. American Mathematical Society (2010). ISBN 978-0-8218-4974-3
11. Eymard, R., Gallouët, T., Herbin, R.: Finite volume methods. Hanbook Numer. Anal. **7**, 713–1020 (2000). https://doi.org/10.1016/S1570-8659(00)07005-8
12. Friedrich, L., Winters, A.R., Del Rey Fernández, D.C., Gassner, G.J., Parsani, M., Carpenter, M.H.: An entropy stable h/p non-conforming dircontinuous Galerkin method with the summation-by-parts property. J. Sci. Comput. **77**, 689–725 (2018). https://doi.org/10.1007/s10915-018-0733-7
13. Gallouët, T., Larcher, A., Latché, J.: Convergence of a finite volume scheme for the convection-diffusion equation with L^1 data. Math. Comput. **810**(279), 1429–1454 (2012)
14. Geuzaine, C., Remacle, J.-F.: Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities. Int. J. Numer. Methods Eng. **79**, 1309–1331 (2009). https://doi.org/10.1002/nme.2579
15. Gjesteland, A., Svärd, M.: Entropy stability for the compressible Navier–Stokes equations with strong imposition of the no-slip boundary condition. J. Comput. Phys. **470** (2022). https://doi.org/10.1016/j.jcp.2022.111572
16. Gustafsson, B.: High order difference methods for time dependent PDE. In: Springer Series in Computational Mathematics. Springer, Berlin (2008). https://doi.org/10.1007/978-3-540-74993-6
17. Gustafsson, B., Kreiss, H.-O., Oliger, J.: Time-dependent problems and difference methods. In: Pure and Applied Mathematics, 2 edn. Wiley (2013). ISBN 978-0-470-90056-7
18. Klausen, R.A., Winther, R.: Robust convergence of multi point flux approximation on rough grids. Numer. Math. **104**, 317–337 (2006). https://doi.org/10.1007/s00211-006-0023-4
19. Knabner, P., Angermann, L.: Numerical methods for elliptic and parabolic partial differential equations, 2nd edn. With contributions by Andreas Rupp. In: Texts in Applied Mathematics. Springer-Verlag New York Inc (2021). https://doi.org/10.1007/978-3-030-79385-2
20. Kreiss, H.O., Scherer, G.: Finite element and finite difference methods for hyperbolic partial differential equations. In: Mathematical Aspects of Finite Elements in Partial Differential Equations (1974)
21. Mattsson, K., Nordström, J.: Summation by parts operators for finite-difference approximations of second derivatives. J. Comput. Phys. **199**, 503–540 (2004). https://doi.org/10.1016/j.jcp.2004.03.001
22. Moukalled, F., Mangani, L., Darwish, M.: The Finite Volume Method in Computational Fluid Dynamics: An Advanced Introduction with OpenFOAM and Matlab, Volume 113 of Fluid Mechanics and Its Applications. Springer International Publishing (2016). https://doi.org/10.1007/978-3-319-16874-6
23. Nordström, J., Forsberg, K., Adamsson, C., Eliasson, P.: Finite volume methods, unstructured meshes and strict stability for hyperbolic problems. Appl. Numer. Math. **45**, 453–473 (2003). https://doi.org/10.1016/S0168-9274(02)00239-8

24. Parsani, M., Carpenter, M.H., Nielsen, E.J.: Entropy stable wall boundary conditions for the three-dimensional compressible Navier–Stokes equations. J. Comput. Phys. **292**, 88–113 (2015). https://doi.org/10.1016/j.jcp.2015.03.026
25. Renardy, M., Rogers, R. C.: An Introduction to Partial Differential Equations, Volume 13 of Texts in Applied Mathematics. Springer-Verlag New York Inc (1993). ISBN 0-387-97952-2
26. Simon, J.: Compact sets in the space $L^p(0, T; B)$. Ann. Math. Pura Appl. **146**, 65–96 (1986)
27. Stein, E.M.: Singular Integrals and Differentiability Properties of Functions. Princeton University Press (1970). ISBN 0-691-08079-8
28. Strand, B.: Summation by parts for finite difference approximations for d/dx. J. Comput. Phys. **110**, 47–67 (1994)
29. Svärd, M., Nordström, J.: Stability of finite volume approximations for the Laplacian operator on quadrilateral and triangular grids. Appl. Numer. Math. **51**, 101–125 (2004). https://doi.org/10.1016/j.apnum.2004.02.001
30. Svärd, M., Nordström, J.: Review of summation-by-parts schemes for initial-boundary-value problems. J. Comput. Phys. **268**, 17–38 (2014). https://doi.org/10.1016/j.jcp.2014.02.031
31. Svärd, M., Carpenter, M.H., Parsani, M.: Entropy stability and the no-slip wall boundary condition. SIAM J. Numer. Anal. **560**(1), 256–273 (2018). https://doi.org/10.1137/16M1097225
32. Yamaleev, N.K., Del Rey Fernández, D.C., Lou, J., Carpenter, M.H.: Entropy stable spectral collocation schemes for the 3-D Navier–Stokes equations on dynamic unstructured grids. J. Comput. Phys. **399** (2019). https://doi.org/10.1016/j.jcp.2019.108897

Paper C

Injected Dirichlet boundary conditions for general diagonal-norm SBP operators

A. Gjesteland, D. Del Rey Fernández and M. Svärd

Under review, 2023.

Paper D

Entropy stable far-field boundary conditions for the compressible Navier-Stokes equations

M. Svärd and A. Gjesteland

Submitted, 2023.

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230843925 (print)
9788230855218 (PDF)