INSTITUTE OF MATHEMATICS

THE CZECH ACADEMY OF SCIENCES

Convergence of a finite volume scheme for the compressible Navier-Stokes system

Eduard Feireisl

Hana Mizerová

Bangwei She

Preprint No. 41-2018

PRAHA 2018

Convergence of a finite volume scheme for the compressible Navier–Stokes system

Eduard Feireisl * Hana Mizerová *,† Bangwei She *

September 18, 2018

* Institute of Mathematics of the Academy of Sciences of the Czech Republic
Žitná 25, CZ-115 67 Praha 1, Czech Republic
feireisl@math.cas.cz, mizerova@math.cas.cz, she@math.cas.cz

† Department of Mathematical Analysis and Numerical Mathematics
Faculty of Mathematics, Physics and Informatics of the Comenius University
Mlynská dolina, 842 48 Bratislava, Slovakia

Abstract

We propose a finite volume scheme for the compressible (isentropic) Navier–Stokes system. We show that the numerical solutions generate a dissipative measure-valued solution of the limit system by deriving suitable stability and consistency estimates. By virtue of the weak-strong uniqueness principle in the class of dissipative measure-valued solutions, the limit coincides with the strong solution as long as the latter exists.

Keywords: compressible Navier–Stokes system, finite volume method, dissipative measure–valued solution

1 Introduction

We study the flow of an *isentropic* viscous fluid governed by the compressible Navier–Stokes system:

$$\begin{aligned} \partial_t \varrho + \mathrm{div}_x(\varrho \mathbf{u}) = 0, \\ \partial_t(\varrho \mathbf{u}) + \mathrm{div}_x(\varrho \mathbf{u} \otimes \mathbf{u}) + \nabla_x p = \mu \Delta_x \mathbf{u} + (\mu + \lambda) \nabla_x \mathrm{div}_x \mathbf{u} \end{aligned} \tag{1.1}$$

*The research of E.F. and H.M. leading to these results has received funding from the Czech Sciences Foundation (GAČR), Grant Agreement 18–05974S. The research of B.S. leading to these results has received funding from the Czech Sciences Foundation (GAČR), Grant Agreement 16–03230S. The Institute of Mathematics of the Academy of Sciences of the Czech Republic is supported by RVO:67985840.

in the time–space domain $(0,T)\times\Omega$. Here $\varrho=\varrho(t,x)$, and $\mathbf{u}=\mathbf{u}(t,x)$ are the fluid density and velocity, constants $\mu>0$, $\lambda\geq 0$ are the viscosity coefficients. The pressure p is assumed to satisfy the *isentropic* state equation

$$p(\varrho)=a\varrho^{\gamma},\quad a>0,\quad \gamma>1. \tag{1.2}$$

For the sake of simplicity, we impose the space periodic boundary conditions, meaning, the physical domain can be identified with the flat torus $\Omega=([0,1]|_{0,1})^d$, $d=1,2,3$. The problem is (formally) closed by prescribing the initial conditions

$$\varrho(0)=\varrho_0\in L^{\gamma}(\Omega),\ \varrho_0>0,\ \mathbf{u}(0)=\mathbf{u}_0\in L^2(\Omega;R^d). \tag{1.3}$$

Although many numerical methods have been developed to solve the isentropic Navier-Stokes system, see, e.g., the monographs by Dolejší and Feistauer [4, 5], the article papers [1, 12, 13, 14] and references therein, their mathematical properties are not well understood. In particular, the convergence of approximate solutions towards the solution of the continuous system remains open in many cases. A pioneering work was done by Karper who proved the convergence of a combined finite volume (FV) – finite element (FE) method in [16] under the assumption on the adiabatic exponent $\gamma>3$, see [16]. The physically relevant range of adiabatic coefficient $\gamma\in(1,2)$ was successfully handled in [9] via the dissipative measure-valued (DMV) solutions introduced in [7].

Pursuing the strategy of [9], we introduce the concept of (DMV) solution to problem (1.1).

Definition 1.1 (Dissipative measure-valued solution)**.** We say that a parametrized family of probability measures $\{\mathcal{V}_{t,x}\}_{(t,x)\in(0,T)\times\Omega}$,

$$\mathcal{V}_{t,x}\in L^{\infty}_{weak}\Big((0,T)\times\Omega;\mathcal{P}(Q)\Big),\ Q=\Big\{[\varrho,\mathbf{u}]\ \Big|\ \varrho\in[0,\infty),\ \mathbf{u}\in R^N\Big\},$$

is *a dissipative measure-valued (DMV) solution* of the Navier–Stokes system in $(0,T)\times\Omega$, with the initial condition $\mathcal{V}_{0,x}\in\mathcal{P}(Q)$ and dissipative defect $\mathcal{D}\in L^{\infty}(0,T)$, $\mathcal{D}\geq 0$, if the following holds:

- $$\left[\int_{\Omega}\langle\mathcal{V}_{t,x};\varrho\rangle\phi(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau}=\int_0^{\tau}\int_{\Omega}[\langle\mathcal{V}_{t,x};\varrho\rangle\partial_t\phi+\langle\mathcal{V}_{t,x};\varrho\mathbf{u}\rangle\cdot\nabla_x\phi]\,\mathrm{d}x\,\mathrm{d}t$$
  for any $0\leq\tau\leq T$, and any $\phi\in C^1\big([0,T]\times\Omega\big)$;

- $$\left[\int_{\Omega}\langle\mathcal{V}_{t,x};\varrho\mathbf{u}\rangle\boldsymbol{\phi}(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau}=\int_0^{\tau}\int_{\Omega}[\langle\mathcal{V}_{t,x};\varrho\mathbf{u}\rangle\partial_t\boldsymbol{\phi}+\langle\mathcal{V}_{t,x};\varrho\mathbf{u}\otimes\mathbf{u}+p(\varrho)\mathbb{I}\rangle:\nabla_x\boldsymbol{\phi}]\,\mathrm{d}x\,\mathrm{d}t$$
  $$-\int_0^{\tau}\int_{\Omega}\mathcal{S}(\nabla_x\mathbf{u}):\nabla_x\boldsymbol{\phi}\,\mathrm{d}x\,\mathrm{d}t+\int_0^{\tau}\langle r^M;\nabla_x\boldsymbol{\phi}\rangle\,\mathrm{d}t$$
  for any $0\leq\tau\leq T$, and any $\boldsymbol{\phi}\in C^1_c\big([0,T]\times\Omega;R^d\big)$, where
  $$\mathbf{u}_{t,x}=\langle\mathcal{V}_{t,x};\mathbf{u}\rangle\,,\ \mathbf{u}\in L^2(0,T;W^{1,2}(\Omega;R^d)),$$
  $$\mathcal{S}(\nabla_x\mathbf{u})=\mu(\nabla_x\mathbf{u}+\nabla_x^t\mathbf{u})+\lambda\mathrm{div}_x\mathbf{u}\mathbb{I},\ \text{and}\ r^M\in L^1\big(0,T;\mathcal{M}(\Omega)\big);$$

- $$\left[\int_\Omega \langle \mathcal{V}_{t,x}; \frac{1}{2}\varrho \mathbf{u}^2 + \mathcal{H}(\varrho)\rangle \,\mathrm{d}x\right]_{t=0}^{t=\tau} + \int_0^\tau \int_\Omega \mathcal{S}(\nabla_x \mathbf{u}) : \nabla_x \mathbf{u} \,\mathrm{d}x \,\mathrm{d}t + \mathcal{D}(\tau) \leq 0,$$

  for a.a. $0 \leq \tau \leq T$, where $\mathcal{H}(\varrho) = \frac{p(\varrho)}{\gamma - 1}$. The dissipation defect $\mathcal{D}$ dominates the concentration measure R^M, specifically,

$$\left|\langle r^M(\tau); \phi\rangle\right| \lesssim \xi(\tau)\mathcal{D}(\tau) \left\|\phi\right\|_{C(\Omega)}, \text{ for some } \xi \in L^1(0,T).$$

The interested reader may consult [7] for thorough discussion of the concept of (DMV) solutions and their basic properties, in particular, the weak–strong uniqueness principle used in the present paper. Recently, the approach to convergence proof via the (DMV) solutions has been successfully applied to the two finite difference Marker-and-Cell schemes, see [17, 18]. To the best of our knowledge, there is no convergence proof of any (FV) scheme for the Navier–Stokes system (1.1).

Our numerical method is inspired by a semi-discrete (FV) scheme proposed for the complete Euler system in [10], where the convergence of piecewise constant numerical solutions via the (DMV) solutions was proved. We adapt this approach to a fully discrete scheme for the isentropic Navier–Stokes system (1.1) and aim to show the stability as well as the convergence of numerical solutions to the smooth solution of the limit system.

The paper is organized as follows. In Section 2, we introduce the necessary preliminaries including the properties of the mesh, basic notation, the numerical method, and some (in)equalities. Next, in Section 3, we show the energy stability of the scheme and derive all necessary *a priori* bounds. Then we establish the consistency formulation of the scheme in Section 4. Finally, we address the convergence of approximate solutions in Section 5.

2 Numerical scheme

We introduce the basic notation, mesh, space and time discretization, and, finally, we define the numerical scheme along with some useful (in)equalities.

2.1 Space discretization

Mesh. A discretization of Ω is given by $\mathcal{M} = (\mathcal{T}, \mathcal{E}, \mathcal{D})$, where:

- The primary grid $\mathcal{T}$ is the set of all compact regular quadrilateral elements K such that

$$\Omega = \bigcup_{K \in \mathcal{T}} K.$$

Let h_i be the mesh size in the i^{th} Cartesian direction, and $h = \max_{i=1,\dots,d} h_i$ be the mesh size. The mesh is regular in the sense that there exists a positive η_h such that $\eta_h = \max_{i=1,\dots,d} \left\{\frac{h}{h_i}\right\}$.

• We denote by $\mathcal{E}$ the set of all faces, and by $\mathcal{E}_i$ the set of all faces that are orthogonal to the standard basis vector $\mathbf{e}_i$ ($i \in \{1, \dots, d\}$) of the Cartesian coordinate system. By $\mathcal{E}(K)$ we denote the set of faces of an element K, and define $\mathcal{E}_i(K) = \mathcal{E}(K) \cap \mathcal{E}_i$. Each face $\sigma \in \mathcal{E}$ is associated with a normal vector $\mathbf{n}$. The points $\mathbf{x}_K$ and $\mathbf{x}_\sigma$ stand for the centers of mass of an element $K \in \mathcal{T}$ and a face $\sigma \in \mathcal{E}$, respectively.

• The intersection $K \cap L$, for $K,\ L \in \mathcal{T},\ K \neq L$, is either a vertex, or an edge, or a face $\sigma \in \mathcal{E}$. For any $\sigma \in \mathcal{E}$ we write $\sigma = K|L$ if $\sigma = \mathcal{E}(K) \cap \mathcal{E}(L)$, and further write $\sigma = \overrightarrow{K|L}$ if $\mathbf{x}_L = \mathbf{x}_K + h_i \mathbf{e}_i$ or $\mathbf{x}_L = \mathbf{x}_K + (h_i - 1)\mathbf{e}_i$ for any $\sigma \in \mathcal{E}_i$. Similarly, we write $K = \overrightarrow{[\sigma\sigma']}$ for $\sigma, \sigma' \in \mathcal{E}_i(K)$ if $\mathbf{x}_{\sigma'} = \mathbf{x}_\sigma + h_i \mathbf{e}_i$. For any $\sigma = K|L \in \mathcal{E}_i,\ i \in 1, \dots, d$, we also denote by $d_\sigma = h_i$ the periodic distance between the points $\mathbf{x}_K$ and $\mathbf{x}_L$.

• The dual grid $\mathcal{D}$ is the set of all dual cells. A dual element D_σ is associated to a generic face $\sigma = K|L \in \mathcal{E}$, where $D_\sigma = D_{\sigma,K} \cup D_{\sigma,L}$, and $D_{\sigma,K}$ (resp. $D_{\sigma,L}$) is built by half of K (resp. L), see Figure 1 for an example of such cell. Furthermore, we define $\mathcal{D}_i = \{D_\sigma\}_{\sigma \in \mathcal{E}_i}, i \in \{1, \dots, d\}$. Note that the dual grid verifies for each fixed i the equality

$$\Omega = \bigcup_{\sigma \in \mathcal{E}_i} D_\sigma.$$

We emphasize that the dual grid is not used for the implementation of the scheme, but only for the theoretical proof of convergence.

• By $|K|$, $|D_\sigma|$ and $|\sigma|$ we denote the (d–, d– and $(d-1)$–dimensional) Lebesgue measure of an element K, a dual cell D_σ, and a face σ, respectively. Obviously, $|K| = h_i|\sigma|$ for any $\sigma \in \mathcal{E}_i(K)$ and $|D_\sigma| = |\sigma| d_\sigma$ for any $\sigma \in \mathcal{E}_i$. In what follows, we shall suppose

$$|K| = |D_\sigma| \approx h^d,\ |\sigma| \approx h^{d-1} \text{ for any } K \in \mathcal{T},\ D_\sigma \in \mathcal{D},\ \sigma \in \mathcal{E}.$$

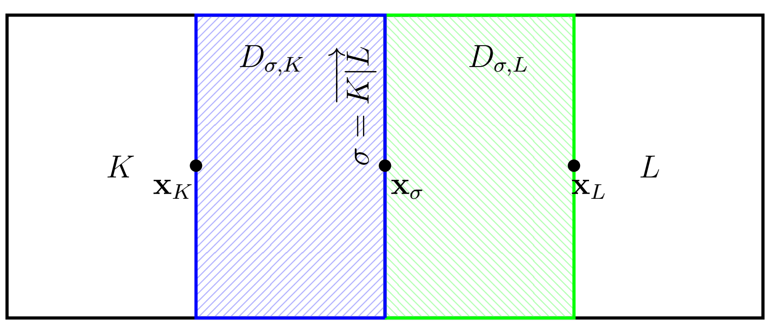

Figure 1: Dual grid

Function spaces. In order to define a finite volume scheme we introduce the spaces Q_h and $W_h^{(i)}$ ($i \in \{1, \dots, d\}$) of piecewise constant functions defined on the primary grid $\mathcal{T}$ and the dual grid

$\mathcal{D}_i$, respectively. By $\mathbf{q} = (q_1, \dots, q_d) \in W_h := \left(W_h^{(1)}, \dots, W_h^{(d)}\right)$, we mean that $q_i \in W_h^{(i)}$, for all $i = 1, \dots, d$. We define the standard projections of $\phi \in L^1(\Omega)$ into the discrete functional spaces Q_h and W_h,

$$\Pi_{\mathcal{T}} : \ L^1(\Omega) \to Q_h. \quad \Pi_{\mathcal{T}} \phi = \sum_{K \in \mathcal{T}} 1_K \frac{1}{|K|} \int_K \phi \, \mathrm{d}x,$$

$$\Pi_{\mathcal{D}} : \quad L^1(\Omega) \to W_h. \quad \Pi_{\mathcal{D}} = (\Pi_{\mathcal{D}}^{(1)}, \dots, \Pi_{\mathcal{D}}^{(d)}), \quad \Pi_{\mathcal{D}}^{(i)} \phi = \sum_{\sigma \in \mathcal{E}_i} \frac{1_{D_\sigma}}{|\sigma|} \int_\sigma \phi \, \mathrm{d}Sx.$$

For a piecewise (elementwise) continuous function v we define

$$v^{\text{out}}(x) = \lim_{\delta \to 0+} v(x + \delta \mathbf{n}), \ v^{\text{in}}(x) = \lim_{\delta \to 0+} v(x - \delta \mathbf{n}), \ \overline{v}(x) = \frac{v^{\text{in}}(x) + v^{\text{out}}(x)}{2}, \ [\![v]\!] = v^{\text{out}}(x) - v^{\text{in}}(x)$$

whenever $x \in \sigma \in \mathcal{E}$. Hereafter we mean by $\mathbf{v} \in Q_h$ that $\mathbf{v} \in Q_h(\Omega; R^d)$, i.e., $v_i \in Q_h$, for all $i = 1, \dots, d$.

Diffusive upwind flux. Given the velocity filed $\mathbf{v} \in Q_h$, the upwind flux for any function $r \in Q_h$ is defined at each face $\sigma \in \mathcal{E}$ by

$$Up[r, \mathbf{v}] = r^{\text{up}} \mathbf{v} \cdot \mathbf{n} = r^{\text{in}} [\overline{\mathbf{v}} \cdot \mathbf{n}]^+ + r^{\text{out}} [\overline{\mathbf{v}} \cdot \mathbf{n}]^- = \overline{r} \ \overline{\mathbf{v}} \cdot \mathbf{n} - \frac{1}{2} |\overline{\mathbf{v}} \cdot \mathbf{n}| \, [\![r]\!] \,,$$

where

$$[f]^\pm = \frac{f \pm |f|}{2} \quad \text{and} \quad r^{\text{up}} = \begin{cases} r^{\text{in}} & \text{if } \overline{\mathbf{u}} \cdot \mathbf{n} \geq 0, \\ r^{\text{out}} & \text{if } \overline{\mathbf{u}} \cdot \mathbf{n} < 0. \end{cases}$$

Furthermore, we consider a diffusive numerical flux function of the following form

$$F_h(r, \mathbf{v}) = Up[r, \mathbf{v}] - h^\varepsilon \, [\![r]\!] \,, \ \varepsilon > 0. \tag{2.1}$$

Discrete differential operators. We define the discrete differential operators with respect to both the primary and the dual grid. The divergence operator based on the primary grid that appears in the numerical scheme (2.3) below is given by

$$\mathrm{div}_h \mathbf{u}_h(\mathbf{x}) := \sum_{K \in \mathcal{T}} (\mathrm{div}_h \mathbf{u}_h)_K 1_K, \quad (\mathrm{div}_h \mathbf{u}_h)_K := \frac{1}{|K|} \sum_{\sigma \in \mathcal{E}(K)} |\sigma| \overline{\mathbf{u}_h} \cdot \mathbf{n}, \ \forall \ \mathbf{u}_h \in Q_h. \tag{2.2}$$

We also need some discrete operators that are not directly used to discretize the Navier-Stokes system, but are essential to establish the consistency formulation in Section 4. Thus, for any $r_h \in Q_h$ and $\mathbf{q}_h = (q_{1,h}, \dots, q_{d,h}) \in W_h$, we define the difference operators based on the dual grid,

$$\eth_{\mathcal{E}}^{(i)} r_h(\mathbf{x}) := \sum_{\sigma \in \mathcal{E}_i} 1_{D_\sigma} \left(\eth_{\mathcal{E}}^{(i)} r_h \right)_{D_\sigma}, \quad \left(\eth_{\mathcal{E}}^{(i)} r_h \right)_{D_\sigma} := \frac{r_L - r_K}{d_\sigma}, \ \forall \ \sigma = \overrightarrow{K|L} \in \mathcal{E}_i,$$

and the primary grid

$$\eth_{\mathcal{T}}^{(i)} q_{i,h}(\mathbf{x}) := \sum_{K\in\mathcal{T}} \left(\eth_{\mathcal{T}}^{(i)} q_{i,h}\right)_K 1_K,\ i,\in\{1,\ldots,d\},$$

where

$$\left(\eth_{\mathcal{T}}^{(i)} q_{i,h}\right)_K := \frac{q_{i,\sigma'} - q_{i,\sigma}}{h_i},\ \forall\, \sigma,\sigma' \in \mathcal{E}_i \text{ and } K = [\overrightarrow{\sigma\sigma'}].$$

Using the above notations, we define the gradient operators for $r_h \in Q_h$, and $\mathbf{q}_h \in W_h$, by

$$\nabla_{\mathcal{E}} r_h(\mathbf{x}) := (\eth_{\mathcal{E}}^{(1)} r_h, \ldots, \eth_{\mathcal{E}}^{(d)} r_h)(\mathbf{x}), \quad \text{and} \quad \nabla_{\mathcal{T}}\mathbf{q}_h := (\eth_{\mathcal{T}}^{(1)} q_{1,h}, \ldots, \eth_{\mathcal{T}}^{(d)} q_{d,h})(\mathbf{x}),$$

respectively. Note that the divergence operator div_h defined in (2.2) can be written as

$$\mathrm{div}_h \mathbf{u}_h = \sum_{i=1}^{d} \eth_{\mathcal{T}}^{(i)} \overline{u_{i,h}}, \quad \forall\ \mathbf{u}_h \in Q_h.$$

Finally, we define the Laplace operator for $r_h \in Q_h$ on the primary grid

$$\Delta_h r_h(\mathbf{x}) = \sum_{i=1}^{d} \Delta_h^{(i)} r_h(\mathbf{x}) = \sum_{K\in\mathcal{T}} (\Delta_h r_h)_K 1_K, \quad \Delta_h^{(i)} r_h(\mathbf{x}) = \sum_{K\in\mathcal{T}} (\Delta_h^{(i)} r_h)_K 1_K,$$

where $i \in \{1,\ldots,d\}$, and

$$(\Delta_h^{(i)} r_h)_K := \frac{1}{|K|} \sum_{\sigma\in\mathcal{E}_i(K)} |\sigma| \frac{[\![r_h]\!]}{d_\sigma}, \quad (\Delta_h r_h)_K := \frac{1}{|K|} \sum_{\sigma\in\mathcal{E}(K)} |\sigma| \frac{[\![r_h]\!]}{d_\sigma},\ \forall\ K \in \mathcal{T}.$$

In addition, it is worth mentioning that

$$\Delta_h^{(i)} r_h = \eth_{\mathcal{T}}^{(i)}(\eth_{\mathcal{E}}^{(i)} r_h), \forall\, i \in \{1,\ldots,d\}.$$

2.2 Time discretization

For a given time step $\Delta t \approx h > 0$, we denote the approximation of a function v_h at time $t^k = k\Delta t$ by v_h^k for $k = 1,\ldots,N_T(= T/\Delta t)$. The time derivative is discretized by the backward Euler method,

$$D_t v_h^k = \frac{v_h^k - v_h^{k-1}}{\Delta t}, \quad \text{for } k = 1,2,\ldots,N_T.$$

Furthermore, we introduce the piecewise constant extension of discrete values,

$$\begin{aligned}\varrho_h(t,\cdot) = \varrho_h^0 \text{ for } t < \Delta t,\ \varrho_h(t,\cdot) \quad &= \varrho_h^k \text{ for } t \in [k\Delta t,(k+1)\Delta t),\ k = 1,2,\ldots,N_T,\\ \mathbf{u}_h(t,\cdot) = \mathbf{u}_h^0 \text{ for } t < \Delta t,\ \mathbf{u}_h(t,\cdot) \quad &= \mathbf{u}_h^k \text{ for } t \in [k\Delta t,(k+1)\Delta t),\ k = 1,2,\ldots,N_T,\end{aligned}$$

and $p_h = p(\varrho_h)$, for which the discrete time derivative then reads

$$D_t v_h = \frac{v_h(t,\cdot) - v_h(t-\Delta t,\cdot)}{\Delta t}.$$

We shall write $A \lesssim B$ if $A \le cB$ for a generic positive constant c independent of h.

2.3 Numerical scheme

Using the above notation we introduce the implicit (FV) scheme to approximate system (1.1).

Definition 2.1 (Numerical scheme)**.** Given the initial values $(\varrho_h^0, \mathbf{u}_h^0) = (\Pi_\mathcal{T}\varrho_0, \Pi_\mathcal{T}\mathbf{u}_0)$, find $(\varrho_h, \mathbf{u}_h) \in Q_h \times Q_h$ satisfying for $k = 1, \dots, N_T$ the following equations

$$\int_\Omega D_t \varrho_h^k \phi_h \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma F_h(\varrho_h^k, \mathbf{u}_h^k) \left[\!\left[\phi_h \right]\!\right] \mathrm{d}Sx = 0, \quad \text{for all } \phi_h \in Q_h, \tag{2.3a}$$

$$\int_\Omega D_t(\varrho_h^k \mathbf{u}_h^k) \cdot \boldsymbol{\phi}_h \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma \mathbf{F}_h(\varrho_h^k \mathbf{u}_h^k, \mathbf{u}_h^k) \cdot \left[\!\left[\boldsymbol{\phi}_h \right]\!\right] \mathrm{d}Sx - \sum_{\sigma\in\mathcal{E}} \int_\sigma \overline{p_h^k} \mathbf{n} \cdot \left[\!\left[\boldsymbol{\phi}_h \right]\!\right] \mathrm{d}Sx$$
$$= -\mu \sum_{\sigma\in\mathcal{E}} \int_\sigma \frac{1}{d_\sigma} \left[\!\left[\mathbf{u}_h^k \right]\!\right] \cdot \left[\!\left[\boldsymbol{\phi}_h \right]\!\right] \mathrm{d}Sx - (\mu + \lambda) \int_\Omega \mathrm{div}_h \mathbf{u}_h^k \, \mathrm{div}_h \boldsymbol{\phi}_h \,\mathrm{d}x, \quad \text{for all } \boldsymbol{\phi}_h \in Q_h. \tag{2.3b}$$

Approximate solutions resulting from the scheme (2.3) enjoy the following properties:

1. **Conservation of mass.**
Taking $\phi_h \equiv 1$ in the equation of continuity (2.3a) yields the total mass conservation

$$\int_\Omega \varrho_h(t, \cdot) \,\mathrm{d}x = \int_\Omega \varrho_h^0 \,\mathrm{d}x = M_0 > 0, \ t \geq 0.$$

2. **Existence of numerical solution.**
The discrete problem (2.3) admits, for any $k \in \{1, \dots, N_T\}$, a solution $(\varrho_h^k, \mathbf{u}_h^k)$. We refer the reader to [15, Theorem 3.5] for the proof.

3. **Positivity of numerical density.**
Any solution $(\varrho_h^k, \mathbf{u}_h^k)$ to (2.3) satisfies $\varrho_h^k > 0$ provided $\varrho_h^{k-1} > 0$, $k \in \{1, \dots, N_T\}$, see [15, Lemma 3.2] for the proof.

2.4 Preliminary material

Firstly, we recall the identity

$$\overline{u_h v_h} - \overline{u_h}\ \overline{v_h} = \frac{1}{4} \left[\!\left[u_h \right]\!\right] \left[\!\left[v_h \right]\!\right], \tag{2.4}$$

together with the product rule

$$\left[\!\left[u_h v_h \right]\!\right] = \overline{u_h} \left[\!\left[v_h \right]\!\right] + \left[\!\left[u_h \right]\!\right] \overline{v_h}, \tag{2.5}$$

which are valid for any $u_h, v_h \in Q_h$. A direct application of the product rule (2.5) further implies

$$\left[\!\left[r_h \mathbf{v}_h \right]\!\right] \left[\!\left[\mathbf{v}_h \right]\!\right] - \frac{1}{2} \left[\!\left[r_h \right]\!\right] \left[\!\left[|\mathbf{v}_h|^2 \right]\!\right] = \overline{r_h} \left[\!\left[\mathbf{v}_h \right]\!\right]^2, \ \text{for } r_h \in Q_h, \ \mathbf{v}_h \in Q_h, \tag{2.6}$$

and the following lemma.

Lemma 2.2. *For any $r_h \in Q_h$ and $\mathbf{v}_h \in Q_h$, it holds*

$$\sum_{\sigma\in\mathcal{E}} \int_\sigma (\overline{r_h} \llbracket \mathbf{v}_h \rrbracket + \overline{\mathbf{v}_h} \llbracket r_h \rrbracket) \cdot \mathbf{n} \, \mathrm{d}Sx = 0. \tag{2.7}$$

Proof. For the functions r_h, $\mathbf{v}_h$, constant on each element $K \in \mathcal{T}$, it holds that

$$\sum_{\sigma\in\mathcal{E}} \int_\sigma (\overline{r_h} \llbracket \mathbf{v}_h \rrbracket + \overline{\mathbf{v}_h} \llbracket r_h \rrbracket) \cdot \mathbf{n} \, \mathrm{d}Sx = \sum_{\sigma\in\mathcal{E}} \int_\sigma \llbracket r_h \mathbf{v}_h \rrbracket \cdot \mathbf{n} \, \mathrm{d}Sx = -\sum_{K\in\mathcal{T}} r_K \mathbf{v}_K \cdot \sum_{\sigma\in\mathcal{E}(K)} \int_\sigma \mathbf{n} \, \mathrm{d}Sx = 0.$$

□

Consequently, for any $r_h, \phi_h \in Q_h$ and $\mathbf{q}_h \in W_h$, it is easy to observe the following discrete integration by parts formulae

$$\int_\Omega \Delta_h r_h \phi_h \, \mathrm{d}x = -\int_\Omega \nabla_\mathcal{E} r_h \cdot \nabla_\mathcal{E} \phi_h \, \mathrm{d}x = \int_\Omega r_h \Delta_h \phi_h \, \mathrm{d}x, \tag{2.8a}$$

$$\int_\Omega q_{i,h} \eth_\mathcal{E}^{(i)} r_h \, \mathrm{d}x = -\int_\Omega r_h \eth_\mathcal{T}^{(i)} q_{i,h} \, \mathrm{d}x, \; i \in \{1, \ldots, d\}. \tag{2.8b}$$

Next, we list some basic inequalities used in the numerical analysis. We assume the reader is fairly familiar with this matter, for which we refer to the monograph [6], and the article paper [13]. If $\phi \in C^1(\overline{\Omega})$ we have

$$\left| \llbracket \Pi_\mathcal{T} \phi \rrbracket \right|_\sigma \lesssim h \|\phi\|_{C^1}, \quad \text{for any } x \in \sigma \in \mathcal{E}, \text{ and } \|\phi - \Pi_\mathcal{T}\phi\|_{L^p(\Omega)} \lesssim h\|\phi\|_{C^1}. \tag{2.9}$$

Furthermore, if $\phi \in C^2(\overline{\Omega})$, there hold for all $1 < p \leq \infty$ that

$$\|\nabla_x \phi - \nabla_\mathcal{E} \Pi_\mathcal{T} \phi\|_{L^p} \lesssim h, \quad \|\nabla_\mathcal{E} \Pi_\mathcal{T} \phi\|_{L^p} \lesssim \|\phi\|_{C^1} + h, \tag{2.10}$$

$$\left\|\nabla_x \phi - \nabla_\mathcal{T} \Pi_\mathcal{D}\big(\Pi_\mathcal{T}\phi\big)\right\|_{L^p} \lesssim h, \quad \|\mathrm{div}_x \phi - \mathrm{div}_h(\Pi_\mathcal{T}\phi)\|_{L^p} \lesssim h. \tag{2.11}$$

If in addition, $\phi \in C^3(\overline{\Omega})$, we get

$$\|\Delta_h \Pi_\mathcal{T} \phi - \Delta_x \phi\|_{L^p} \lesssim h\|\phi\|_{C^3}, \quad \|\Delta_h \Pi_\mathcal{T}\phi\|_{L^p} \lesssim \|\phi\|_{C^2} + h\|\phi\|_{C^3}, \; \forall \; 1 < p \leq \infty. \tag{2.12}$$

The inverse estimates [3] for $r_h \in Q_h$ read

$$\|r_h\|_{L^p(\Omega)} \lesssim h^{d(\frac{1}{p} - \frac{1}{q})} \|r_h\|_{L^q(\Omega)} \; \text{ for any } \; 1 \leq q \leq p \leq \infty. \tag{2.13}$$

Finally, we need a discrete version of the Sobolev-type inequality that can be proved exactly as [11, Theorem 10.17].

Lemma 2.3 (Sobolev inequality)**.** *Let the function $r \geq 0$ be such that*

$$0 < \int_\Omega r \,\mathrm{d}x = c_M, \text{ and } \int_\Omega r^\gamma \,\mathrm{d}x \leq C_E \text{ for } \gamma > 1.$$

Then the following Poincaré-Sobolev type inequality holds true

$$\|v\|_{L^6(\Omega)} \leq c \,\|\nabla_{\mathcal{E}} v\|^2_{L^2(\Omega)} + c\left(\int_\Omega r|v| \,\mathrm{d}x\right)^2 \lesssim c\,\|\nabla_{\mathcal{E}} v\|^2_{L^2(\Omega)} + c_M + c\int_\Omega r|v|^2 \,\mathrm{d}x \tag{2.14}$$

for any $v \in Q_h$, where the constant c depends on c_M and c_E but not on the mesh parameter.

The following lemma shall be useful for analysing the error between the continuous convective term and its numerical analogue.

Lemma 2.4. *For any $r_h, \mathbf{v}_h \in Q_h$, and $\phi \in C^1(\Omega)$, it holds*

$$\begin{aligned}
&\int_\Omega r_h \mathbf{v}_h \cdot \nabla_x \phi \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma F_h[r_h, \mathbf{v}_h] \;\llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx \\
&= \sum_{\sigma\in\mathcal{E}} \int_\sigma \left(\frac{1}{2}|\overline{\mathbf{v}_h}\cdot\mathbf{n}| + h^\varepsilon + \frac{1}{4}\llbracket \mathbf{v}_h \rrbracket \cdot \mathbf{n}\right) \llbracket r_h \rrbracket \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx + \int_\Omega r_h \mathbf{v}_h \cdot \big(\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)\big) \,\mathrm{d}x.
\end{aligned}$$

Proof. Using the basic equalities (2.4)–(2.7), we have

$$\begin{aligned}
\int_\Omega r_h \mathbf{v}_h \cdot \nabla_x \phi \,\mathrm{d}x &= \sum_{K\in\mathcal{T}} \int_K r_h \mathbf{v}_h \cdot \nabla_x \phi \;\mathrm{d}x \\
&= \sum_{K\in\mathcal{T}} \int_K r_h \mathbf{v}_h \cdot (\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)) \,\mathrm{d}x + \sum_{K\in\mathcal{T}} \int_{\partial K} r_h \mathbf{v}_h \cdot \mathbf{n}\overline{\Pi_{\mathcal{T}}\phi} \,\mathrm{d}Sx \\
&= \int_\Omega r_h \mathbf{v}_h \cdot (\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)) \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma \llbracket r_h \mathbf{v}_h \rrbracket \cdot \mathbf{n}\overline{\Pi_{\mathcal{T}}\phi} \,\mathrm{d}Sx \\
&= \int_\Omega r_h \mathbf{v}_h \cdot (\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)) \,\mathrm{d}x + \sum_{\sigma\in\mathcal{E}} \int_\sigma \overline{r_h \mathbf{v}_h} \cdot \mathbf{n} \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx \\
&= \int_\Omega r_h \mathbf{v}_h \cdot (\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)) \,\mathrm{d}x + \sum_{\sigma\in\mathcal{E}} \int_\sigma (\overline{r_h \mathbf{v}_h} - \overline{r_h}\;\overline{\mathbf{v}_h}) \cdot \mathbf{n} \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx \\
&\quad + \sum_{\sigma\in\mathcal{E}} \int_\sigma \overline{r_h}\;\overline{\mathbf{v}_h} \cdot \mathbf{n} \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx \pm \sum_{\sigma\in\mathcal{E}} \int_\sigma \left(\frac{1}{2}|\overline{\mathbf{v}_h}\cdot\mathbf{n}| + h^\varepsilon\right) \llbracket r_h \rrbracket \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx \\
&= \int_\Omega r_h \mathbf{v}_h \cdot (\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)) \,\mathrm{d}x + \sum_{\sigma\in\mathcal{E}} \int_\sigma \frac{1}{4} \llbracket r_h \rrbracket \llbracket \mathbf{v}_h \rrbracket \cdot \mathbf{n} \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx \\
&\quad + \sum_{\sigma\in\mathcal{E}} \int_\sigma F_h[r_h, \mathbf{v}_h] \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx + \sum_{\sigma\in\mathcal{E}} \int_\sigma \left(\frac{1}{2}|\overline{\mathbf{v}_h}\cdot\mathbf{n}| + h^\varepsilon\right) \llbracket r_h \rrbracket \llbracket \Pi_{\mathcal{T}}\phi \rrbracket \,\mathrm{d}Sx.
\end{aligned}$$

□

3 Stability

We show the energy stability of the scheme and derive the estimates necessary for the consistency formulation in Section 4. For simplicity, hereafter we will write the norms $\|\cdot\|_{L^q(\Omega)}$ and $\|\cdot\|_{L^p(0,T;L^q(\Omega))}$ as $\|\cdot\|_{L^q}$ and $\|\cdot\|_{L^pL^q}$, respectively.

To begin, we recall the discrete internal energy balance, which is a result of the renormalization of the continuity equation, see, e.g. [8, Section 4.1] or [15, Lemma 3.1]. Indeed, multiplying (2.3a) by $\mathcal{H}'(\varrho_h^k)$ gives rise to the result of the following lemma.

Lemma 3.1 (Internal energy balance)**.** *Let $(\varrho_h, \mathbf{u}_h) \in Q_h \times Q_h$ satisfy the discrete continuity equation* (2.3a)*. Then there exists $\xi \in \mathrm{co}\{\varrho_h^{k-1}, \varrho_h^k\}$ and $\zeta \in \mathrm{co}\{\varrho_K^k, \varrho_L^k\}$ for any $\sigma = K|L \in \mathcal{E}$ such that*

$$\begin{aligned} &\int_\Omega D_t \mathcal{H}(\varrho_h^k) \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma \overline{\mathbf{u}_h^k} \cdot \mathbf{n} \left[\!\left[p(\varrho_h^k) \right]\!\right] \,\mathrm{d}Sx \\ &\qquad = -\frac{\Delta t}{2} \int_\Omega \mathcal{H}''(\xi) |D_t \varrho_h^k|^2 \,\mathrm{d}x - \frac{1}{2} \sum_{\sigma\in\mathcal{E}} \int_\sigma \mathcal{H}''(\zeta) \left[\!\left[\varrho_h^k \right]\!\right]^2 (h^\varepsilon + |\overline{\mathbf{u}_h} \cdot \mathbf{n}|) \,\mathrm{d}Sx. \end{aligned} \tag{3.1}$$

Next, we recall the renormalization of the transport equation, see [8, Lemma A.1, Section A.2].

Lemma 3.2 (Renormalized transport equation)**.** *Suppose that $b_h^k \in Q_h$, $\chi \in C^2(R)$. Then there exists $\xi \in \mathrm{co}\{b_h^{k-1}, b_h^k\}$, $\zeta \in \mathrm{co}\{b_h^k, (b_h^k)^{\mathrm{out}}\}$ for any $\phi_h \in Q_h$ such that*

$$\begin{aligned} &\int_\Omega D_t(\varrho_h^k b_h^k) \chi'(b_h^k) \phi_h \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma Up[\varrho_h^k b_h^k, \mathbf{u}_h^k] \left[\!\left[\chi'(b_h^k)\phi_h \right]\!\right] \,\mathrm{d}Sx \\ &= \int_\Omega D_t \left(\varrho_h^k \chi(b_h^k) \right) \phi_h \,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma Up[\varrho_h^k \chi(b_h^k), \mathbf{u}_h] \left[\!\left[\phi_h \right]\!\right] \,\mathrm{d}Sx + \frac{\Delta t}{2} \int_\Omega \chi''(\xi) \varrho_h^{k-1} |D_t b_h^k|^2 \phi_h \,\mathrm{d}x \\ &\quad + \sum_{\sigma\in\mathcal{E}} \int_\sigma h^\varepsilon \left[\!\left[\varrho_h^k \right]\!\right] \; \left[\!\left[\left(\chi(b_h^k) - \chi'(b_h^k) b_h^k \right) \phi_h \right]\!\right] \,\mathrm{d}Sx \\ &\quad - \frac{1}{2} \sum_{K\in\mathcal{T}} \sum_{\sigma\subset\partial K} \int_\sigma \phi_h \chi''(\zeta) \left[\!\left[b_h^k \right]\!\right]^2 (\varrho_h^k)^{\mathrm{out}} \left[\overline{\mathbf{u}_h^k} \cdot \mathbf{n} \right]^- \,\mathrm{d}Sx. \end{aligned} \tag{3.2}$$

3.1 Total energy balance

Now, we are ready to derive the discrete counterpart of the total energy balance.

Theorem 3.3 (Discrete energy balance)**.** *Let $(\varrho_h, \mathbf{u}_h)$ be a numerical solution obtained from the scheme* (2.3)*. Then, for any $k = 1, \ldots, N_T$, there exists $\xi \in \mathrm{co}\{\varrho_h^{k-1}, \varrho_h^k\}$ and $\zeta \in \mathrm{co}\{\varrho_K^k, \varrho_L^k\}$ such* *that, for any $\sigma = K|L \in \mathcal{E}$,*

$$D_t \int_\Omega \left(\frac{1}{2}\varrho_h^k|\mathbf{u}_h^k|^2 + \mathcal{H}(\varrho_h^k)\right) \mathrm{d}x + h^\varepsilon \sum_{\sigma\in\mathcal{E}} \int_\sigma \overline{\varrho_h^k} \left[\!\left[\mathbf{u}_h^k\right]\!\right]^2 \mathrm{d}Sx + \mu \left\|\nabla_\mathcal{E}\mathbf{u}_h^k\right\|_{L^2}^2 + (\mu+\lambda)\int_\Omega |\mathrm{div}_h \mathbf{u}_h^k|^2 \,\mathrm{d}x$$
$$= -\frac{\Delta t}{2}\int_\Omega \mathcal{H}''(\xi)|D_t\varrho_h^k|^2\,\mathrm{d}x - \frac{1}{2}\sum_{\sigma\in\mathcal{E}}\int_\sigma \mathcal{H}''(\zeta)\left[\!\left[\varrho_h^k\right]\!\right]^2\left(h^\varepsilon + |\overline{\mathbf{u}_h^k}\cdot\mathbf{n}|\right)\mathrm{d}Sx$$
$$-\frac{\Delta t}{2}\int_\Omega \varrho_h^{k-1}|D_t\mathbf{u}_h^k|^2\,\mathrm{d}x - \frac{1}{2}\sum_{\sigma\in\mathcal{E}}\int_\sigma (\varrho_h^k)^{\mathrm{up}}|\overline{\mathbf{u}_h^k}\cdot\mathbf{n}|\left[\!\left[\mathbf{u}_h^k\right]\!\right]^2\mathrm{d}Sx. \tag{3.3}$$

Proof. First, taking $\boldsymbol{\phi}_h = \mathbf{u}_h^k$ in (2.3b) we get

$$\int_\Omega D_t(\varrho_h^k\mathbf{u}_h^k)\cdot\mathbf{u}_h^k\,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}}\int_\sigma \mathbf{F}_h(\varrho_h^k\mathbf{u}_h^k,\mathbf{u}_h^k)\cdot\left[\!\left[\mathbf{u}_h^k\right]\!\right]\mathrm{d}Sx - \sum_{\sigma\in\mathcal{E}}\int_\sigma \overline{p_h^k}\mathbf{n}\cdot\left[\!\left[\mathbf{u}_h^k\right]\!\right]\mathrm{d}Sx$$
$$= -\mu\left\|\nabla_\mathcal{E}\mathbf{u}_h^k\right\|_{L^2} - (\mu+\lambda)\int_\Omega |\mathrm{div}_h\mathbf{u}_h^k|^2\,\mathrm{d}x$$

Next, we use relation (3.2) for $b_h = \mathbf{u}_h^k$, $\chi(|\mathbf{u}_h^k|) = \frac{1}{2}|\mathbf{u}_h^k|^2$, and $\phi_h = 1$ to compute

$$\int_\Omega D_t(\varrho_h^k\mathbf{u}_h^k)\cdot\mathbf{u}_h^k\,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}}\int_\sigma \mathbf{Up}[\varrho_h^k\mathbf{u}_h^k,\mathbf{u}_h^k]\cdot\left[\!\left[\mathbf{u}_h^k\right]\!\right]\mathrm{d}Sx$$
$$= \int_\Omega D_t\left(\frac{1}{2}\varrho_h^k|\mathbf{u}_h^k|^2\right)\mathrm{d}x - \sum_{\sigma\in\mathcal{E}}\int_\sigma \mathbf{Up}\left[\frac{1}{2}\varrho_h^k|\mathbf{u}_h^k|^2,\mathbf{u}_h^k\right]\underbrace{\left[\!\left[1\right]\!\right]}_{=0}\mathrm{d}Sx + \frac{\Delta t}{2}\int_\Omega \varrho_h^{k-1}|D_t\mathbf{u}_h^k|^2\,\mathrm{d}x$$
$$-\sum_{\sigma\in\mathcal{E}}\int_\sigma h^\varepsilon\left[\!\left[\varrho_h^k\right]\!\right]\left[\!\left[\frac{1}{2}|\mathbf{u}_h^k|^2\right]\!\right]\mathrm{d}Sx - \frac{1}{2}\sum_{K\in\mathcal{T}}\sum_{\sigma\subset\partial K}\int_\sigma(\varrho_h^k)^{\mathrm{out}}\left[\overline{\mathbf{u}_h^k}\cdot\mathbf{n}\right]^-\left[\!\left[\mathbf{u}_h^k\right]\!\right]^2\mathrm{d}Sx$$
$$= \int_\Omega D_t\left(\frac{1}{2}\varrho_h^k|\mathbf{u}_h^k|^2\right)\mathrm{d}x + \frac{\Delta t}{2}\int_\Omega \varrho_h^{k-1}|D_t\mathbf{u}_h^k|^2\,\mathrm{d}x - \sum_{\sigma\in\mathcal{E}}\int_\sigma h^\varepsilon\left[\!\left[\varrho_h^k\right]\!\right]\left[\!\left[\frac{1}{2}|\mathbf{u}_h^k|^2\right]\!\right]\mathrm{d}Sx$$
$$+\frac{1}{2}\sum_{\sigma\in\mathcal{E}}\int_\sigma(\varrho_h^k)^{\mathrm{out}}\left[\overline{\mathbf{u}_h^k}\cdot\mathbf{n}\right]^-\left[\!\left[\mathbf{u}_h^k\right]\!\right]^2\mathrm{d}Sx.$$

Further, summing up the previous two observations we infer that

$$D_t\int_\Omega \frac{1}{2}\varrho_h|\mathbf{u}_h^k|^2\,\mathrm{d}x + \mu\left\|\nabla_\mathcal{E}\mathbf{u}_h^k\right\|_{L^2} + (\mu+\lambda)\int_\Omega|\mathrm{div}_h\mathbf{u}_h^k|^2\,\mathrm{d}x$$
$$= \sum_{\sigma\in\mathcal{E}}\int_\sigma \overline{p_h^k}\mathbf{n}\cdot\left[\!\left[\mathbf{u}_h^k\right]\!\right]\mathrm{d}Sx - \sum_{\sigma\in\mathcal{E}}\int_\sigma h^\varepsilon\left[\!\left[\varrho_h^k\mathbf{u}_h^k\right]\!\right]\left[\!\left[\mathbf{u}_h^k\right]\!\right]\mathrm{d}Sx + \sum_{\sigma\in\mathcal{E}}\int_\sigma h^\varepsilon\left[\!\left[\varrho_h^k\right]\!\right]\left[\!\left[\frac{1}{2}|\mathbf{u}_h^k|^2\right]\!\right]\mathrm{d}Sx \tag{3.4}$$
$$-\frac{\Delta t}{2}\int_\Omega \varrho_h^{k-1}|D_t\mathbf{u}_h^k|^2\,\mathrm{d}x - \frac{1}{2}\sum_{\sigma\in\mathcal{E}}\int_\sigma(\varrho_h^k)^{\mathrm{up}}|\overline{\mathbf{u}_h^k}\cdot\mathbf{n}|\left[\!\left[\mathbf{u}_h^k\right]\!\right]^2\mathrm{d}Sx.$$

Finally, combining (3.4) with (3.1) and using the equalities (2.6)–(2.7) we get

$$
\begin{aligned}
&D_t \int_\Omega \left(\frac{1}{2} \varrho_h^k |\mathbf{u}_h^k|^2 + \mathcal{H}(\varrho_h^k) \right) \mathrm{d}x + h^\varepsilon \sum_{\sigma \in \mathcal{E}} \int_\sigma \overline{\varrho_h^k} \left[\!\left[\mathbf{u}_h^k \right]\!\right]^2 \mathrm{d}Sx + \mu \left\| \nabla_{\mathcal{E}} \mathbf{u}_h^k \right\|_{L^2}^2 + (\mu + \lambda) \int_\Omega |\mathrm{div}_h \mathbf{u}_h^k|^2 \, \mathrm{d}x \\
&= -\frac{\Delta t}{2} \int_\Omega \mathcal{H}''(\xi) |D_t \varrho_h^k|^2 \, \mathrm{d}x - \frac{1}{2} \sum_{\sigma \in \mathcal{E}} \int_\sigma \mathcal{H}''(\zeta) \left[\!\left[\varrho_h^k \right]\!\right]^2 \left(h^\varepsilon + |\overline{\mathbf{u}_h^k} \cdot \mathbf{n}| \right) \mathrm{d}Sx \\
&\quad - \frac{\Delta t}{2} \int_\Omega \varrho_h^{k-1} |D_t \mathbf{u}_h^k|^2 \, \mathrm{d}x - \frac{1}{2} \sum_{\sigma \in \mathcal{E}} \int_\sigma (\varrho_h^k)^{\mathrm{up}} |\overline{\mathbf{u}_h^k} \cdot \mathbf{n}| \left[\!\left[\mathbf{u}_h^k \right]\!\right]^2 \mathrm{d}Sx,
\end{aligned}
$$

which completes the proof. □

3.2 Uniform bounds

Having established all necessary ingredients, we are ready to discuss the available *a priori* bounds for solutions of scheme (2.3). From the total energy balance (3.3) and the Sobolev inequality (2.14), we directly get the estimates comprised in the following corollary.

Corollary 3.4. *Let* $(\varrho_h, \mathbf{u}_h)$ *satisfy the scheme* (2.3) *for* $\gamma > 1$. *Then the following estimates hold*

$$\left\| \varrho_h \mathbf{u}_h^2 \right\|_{L^\infty L^1} \lesssim 1, \tag{3.5a}$$

$$\left\| \varrho_h \right\|_{L^\infty L^\gamma} \lesssim 1, \tag{3.5b}$$

$$\left\| \varrho_h \mathbf{u}_h \right\|_{L^\infty L^{\frac{2\gamma}{\gamma+1}}} \lesssim 1, \tag{3.5c}$$

$$\left\| \nabla_{\mathcal{E}} \mathbf{u}_h \right\|_{L^2 L^2} \lesssim 1, \tag{3.5d}$$

$$\left\| \mathrm{div}_h \mathbf{u}_h \right\|_{L^2 L^2} \lesssim 1, \tag{3.5e}$$

$$\left\| \mathbf{u}_h \right\|_{L^2 L^6} \lesssim 1, \tag{3.5f}$$

$$h^\varepsilon \int_0^T \sum_{\sigma \in \mathcal{E}} \int_\sigma \overline{\varrho_h} \left[\!\left[\mathbf{u}_h \right]\!\right]^2 \mathrm{d}Sx \, \mathrm{d}t \lesssim 1, \tag{3.5g}$$

$$\int_0^T \sum_{\sigma \in \mathcal{E}} \int_\sigma \mathcal{H}''(\zeta) \left[\!\left[\varrho_h \right]\!\right]^2 \left(h^\varepsilon + |\overline{\mathbf{u}_h} \cdot \mathbf{n}| \right) \mathrm{d}Sx \, \mathrm{d}t \lesssim 1, \tag{3.5h}$$

where $\zeta \in \mathrm{co}\{\varrho_K, \varrho_L\}$ *for any* $\sigma = K|L \in \mathcal{E}$.

To show the consistency of the numerical scheme we shall need further bounds on the numerical solution, which can be derived provided the adiabatic coefficient in (1.2) lies in the physically realistic range $\gamma \in (1, 2)$.

Lemma 3.5. *Let* $(\varrho_h, \mathbf{u}_h)$ *satisfy the scheme* (2.3), *and let* $1 < \gamma < 2$. *Then there hold*

$$\| \sqrt{\varrho_h} \|_{L^2 L^\infty} \lesssim h^{-\frac{\varepsilon+2}{2\gamma}}, \tag{3.6a}$$

$$\|\varrho_h\|_{L^2L^2} \lesssim h^{-\frac{\varepsilon+2}{2\gamma}}, \tag{3.6b}$$

$$\|\varrho_h\mathbf{u}_h\|_{L^2L^2} \lesssim h^{-\frac{\varepsilon+2}{2\gamma}}. \tag{3.6c}$$

Proof. The estimates (3.6a) and (3.6b) hold true due to the bounds (3.5a), (3.5b), (3.5h). The detailed proof can be found in [18, Lemma 3.3], see also [9, Section 4.10]. The estimate (3.6c) is a direct consequence of the previous two, indeed,

$$\|\varrho_h\mathbf{u}_h\|_{L^2L^2} \lesssim \|\sqrt{\varrho_h}\|_{L^2L^\infty} \|\sqrt{\varrho_h}\mathbf{u}_h\|_{L^\infty L^2} \lesssim h^{-\frac{\varepsilon+2}{2\gamma}}.$$

□

4 Consistency

Another step towards the convergence of the approximate solutions is the consistency of the numerical scheme. In particular, we require the numerical solution to satisfy the weak formulation of the continuous problem up to a residual term vanishing for $h \to 0$.

Theorem 4.1. *Let $(\varrho_h, \mathbf{u}_h)$ be a solution of the approximate problem (2.3) on the time interval $[0,T]$ with $1 < \gamma < 2$ and $0 < \varepsilon < \min\{1, 2(\gamma-1)\}$. Then*

$$-\int_\Omega \varrho_h^0 \phi(0,\cdot)\,\mathrm{d}x = \int_0^\tau \int_\Omega \left[\varrho_h \partial_t \phi + \varrho_h \mathbf{u}_h \cdot \nabla_x \phi\right] \mathrm{d}x\,\mathrm{d}t + \int_0^T e_{1,h}(t,\phi)\,\mathrm{d}t, \tag{4.1}$$

for any $\phi \in C_c^3([0,T) \times \overline{\Omega})$;

$$\begin{aligned} -\int_\Omega \varrho_h^0 \mathbf{u}_h^0 \boldsymbol{\phi}(0,\cdot)\,\mathrm{d}x &= \int_0^T \int_\Omega \left[\varrho_h \mathbf{u}_h \cdot \partial_t \boldsymbol{\phi} + \varrho_h \mathbf{u}_h \otimes \mathbf{u}_h : \nabla_x \boldsymbol{\phi} + p_h \mathrm{div}_x \boldsymbol{\phi}\right] \mathrm{d}x\,\mathrm{d}t, \\ -\mu \int_0^T \int_\Omega \nabla_{\mathcal{E}} \mathbf{u}_h : \nabla_x \boldsymbol{\phi}\,\mathrm{d}x\,\mathrm{d}t &- (\mu+\lambda) \int_0^T \int_\Omega \mathrm{div}_h \mathbf{u}_h\, \mathrm{div}_x \boldsymbol{\phi}\,\mathrm{d}x\,\mathrm{d}t + \int_0^T e_{2,h}(t,\boldsymbol{\phi})\,\mathrm{d}t \end{aligned} \tag{4.2}$$

for any $\boldsymbol{\phi} \in C_c^3([0,T] \times \Omega; R^d)$;

$$\|e_{j,h}(\cdot,\phi)\|_{L^1(0,T)} \lesssim h^\beta \left(\|\phi\|_{C^2} + h\|\phi\|_{C^3}\right),\ j = 1, 2, \ \text{for some} \ \beta > 0.$$

Proof. Let $\phi \in C_c^3([0,T) \times \overline{\Omega})$ and $\boldsymbol{\phi} \in C_c^3([0,T) \times \Omega; R^d)$. We test the equations (2.3a) and (2.3b) with $\Pi_{\mathcal{T}}\phi$ and $\Pi_{\mathcal{T}}\boldsymbol{\phi}$, respectively. Then, we deal with each term in 4 steps.

Step 1 – time derivative terms:

$$\int_0^T \int_\Omega D_t r_h \Pi_{\mathcal{T}}\phi \,\mathrm{d}x\,\mathrm{d}t = \int_0^T \int_\Omega \frac{r_h(t) - r_h(t-\Delta t)}{\Delta t}\phi(t)\,\mathrm{d}x\,\mathrm{d}t$$

$$= \frac{1}{\Delta t}\int_0^T \int_\Omega r_h(t)\phi(t)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{\Delta t}\int_{-\Delta t}^{T-\Delta t}\int_\Omega r_h(t)\phi(t+\Delta t)\,\mathrm{d}x\,\mathrm{d}t$$

$$= -\int_0^T \int_\Omega r_h(t) D_t\phi(t)\,\mathrm{d}x\,\mathrm{d}t + \frac{1}{\Delta t}\int_{T-\Delta t}^{T}\int_\Omega r_h(t)\phi(t+\Delta t)\,\mathrm{d}x\,\mathrm{d}t - \frac{1}{\Delta t}\int_{-\Delta t}^{0}\int_\Omega r_h(t)\phi(t+\Delta t)\,\mathrm{d}x\,\mathrm{d}t$$

$$= -\int_0^T \int_\Omega r_h(t) D_t\phi(t)\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega r_h^0\phi(0)\,\mathrm{d}x,$$

where r_h stands for ϱ_h or $\varrho_h u_{i,h}$, $i = 1,\ldots,d$. Thus, we have

$$\int_0^T \int_\Omega D_t\varrho_h \Pi_{\mathcal{T}}\phi \,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \int_\Omega \varrho_h(t) D_t\phi(t)\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \varrho_h^0\phi(0)\,\mathrm{d}x, \tag{4.3a}$$

$$\int_0^T \int_\Omega D_t(\varrho_h\mathbf{u}_h) \Pi_{\mathcal{T}}\boldsymbol{\phi} \,\mathrm{d}x\,\mathrm{d}t = -\int_0^T \int_\Omega \varrho_h(t)\mathbf{u}_h(t) D_t\boldsymbol{\phi}(t)\,\mathrm{d}x\,\mathrm{d}t - \int_\Omega \varrho_h^0\mathbf{u}_h^0\boldsymbol{\phi}(0)\,\mathrm{d}x, \tag{4.3b}$$

for the continuity and the momentum equations, respectively.

Step 2 – convective terms:
To deal with the convective terms, it is convenient to recall Lemma 2.4:

$$\int_0^T \int_\Omega r_h\mathbf{u}_h\cdot\nabla_x\phi\,\mathrm{d}x\,\mathrm{d}t - \int_0^T \sum_{\sigma\in\mathcal{E}}\int_\sigma F[r_h,\mathbf{u}_h]\,[\![\Pi_{\mathcal{T}}\phi]\!]\,\mathrm{d}Sx\,\mathrm{d}t = \sum_{j=1}^4 E_j(r_h),$$

where

$$E_1(r_h) = \frac{1}{2}\int_0^T \sum_{\sigma\in\mathcal{E}}\int_\sigma |\overline{\mathbf{u}_h}\cdot\mathbf{n}|\,[\![r_h]\!]\,[\![\Pi_{\mathcal{T}}\phi]\!]\,\mathrm{d}Sx\,\mathrm{d}t,$$

$$E_2(r_h) = \frac{1}{4}\int_0^T \sum_{\sigma\in\mathcal{E}}\int_\sigma [\![\mathbf{u}_h]\!]\cdot\mathbf{n}\,[\![r_h]\!]\,[\![\Pi_{\mathcal{T}}\phi]\!]\,\mathrm{d}Sx\,\mathrm{d}t,$$

$$E_3(r_h) = \int_0^T \sum_{\sigma\in\mathcal{E}}\int_\sigma h^\varepsilon\,[\![r_h]\!]\,[\![\Pi_{\mathcal{T}}\phi]\!]\,\mathrm{d}Sx\,\mathrm{d}t,$$

$$E_4(r_h) = \int_0^T \int_\Omega r_h\mathbf{u}_h\cdot\Big(\nabla_x\phi - \nabla_{\mathcal{T}}\Pi_{\mathcal{D}}\big(\Pi_{\mathcal{T}}\phi\big)\Big)\,\mathrm{d}x\,\mathrm{d}t,$$

are the error terms to be estimated. Again, r_h is either ϱ_h or $\varrho_h u_{i,h}$, $i = 1,\ldots,d$.

- Firstly, for the error term E_1 we can write

$$
\begin{aligned}
E_1(r_h) =& \frac{1}{2}\int_0^T \sum_{\sigma\in\mathcal{E}} \int_\sigma |\overline{\mathbf{u}_h}\cdot \mathbf{n}| \, [\![r_h]\!] \, [\![\Pi_{\mathcal{T}}\phi]\!] \, \mathrm{d}Sx \, \mathrm{d}t = \frac{1}{2}\int_0^T \sum_{\sigma\in\mathcal{E}} \int_\sigma |\overline{u_{i,h}}| \, [\![r_h]\!] \, [\![\Pi_{\mathcal{T}}\phi]\!] \, \mathrm{d}Sx \, \mathrm{d}t \\
&= \frac{1}{2}\int_0^T \sum_{i=1}^d \sum_{\sigma\in\mathcal{E}_i} \int_{D_\sigma} h_i |\overline{u_{i,h}}| \eth_{\mathcal{E}}^{(i)} r_h \eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi \, \mathrm{d}Sx \, \mathrm{d}t \\
&= -\frac{1}{2}\int_0^T \sum_{i=1}^d \sum_{K\in\mathcal{T}} \int_K h_i r_h \eth_{\mathcal{T}}^{(i)} \left(|\overline{u_{i,h}}| \eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi \right) \mathrm{d}x \, \mathrm{d}t \\
&= -\frac{1}{2}\int_0^T \sum_{i=1}^d \sum_{K\in\mathcal{T}} \int_K r_K \left(\widehat{|\overline{u_{i,h}}|}_K h_i \eth_{\mathcal{T}}^{(i)} \left(\eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi \right) + h_i \left(\eth_{\mathcal{T}}^{(i)} |\overline{u_{i,h}}| \right) \left(\widehat{\eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi} \right)_K \right) \mathrm{d}x \, \mathrm{d}t
\end{aligned}
$$

where we have used the integration by parts formula (2.8b), the product rule

$$
r_2 q_2 - r_1 q_1 = \frac{r_1 + r_2}{2}(q_2 - q_1) + \frac{q_1 + q_2}{2}(r_2 - r_1),
$$

and

$$
(\widehat{q_i})_K := \frac{q_{i,\sigma} + q_{i,\sigma'}}{2} \quad \text{for } \sigma, \sigma' \in \mathcal{E}_i(K), \ \sigma \neq \sigma'.
$$

Further, employing the inequality $\left(\frac{a+b}{2}\right)^2 \leq \frac{a^2+b^2}{2}$ twice, we claim $\left\| \widehat{\overline{u_{i,h}}} \right\|_{L^2} \lesssim \|u_{i,h}\|_{L^2}$. Similarly, we claim $\left\| \eth_{\mathcal{T}}^{(i)} \overline{u_{i,h}} \right\|_{L^2} \lesssim \left\| \eth_{\mathcal{E}}^{(i)} u_{i,h} \right\|_{L^2}$ as $\left(\eth_{\mathcal{T}}^{(i)} \overline{u_{i,h}} \right)_K = \left(\widehat{\eth_{\mathcal{E}}^{(i)} u_{i,h}} \right)_K$. Then applying Hölder's inequality, interpolation error estimates (2.10), (2.12), the velocity estimates (3.5d), (3.5f), the fact $|\partial_x u_i| \geq \partial_x |u_i|$, and noticing $\Delta_h^{(i)} r := \eth_{\mathcal{T}}^{(i)} \eth_{\mathcal{E}}^{(i)} r$, we derive

$$
\begin{aligned}
E_1(r_h) &= \int_0^T \sum_{i=1}^d \sum_{K\in\mathcal{T}} \int_K r_K \left(\widehat{|\overline{u_{i,h}}|}_K h_i \eth_{\mathcal{T}}^{(i)} \left(\eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi \right) + h_i \left(\eth_{\mathcal{T}}^{(i)} |\overline{u_{i,h}}| \right) \left(\widehat{\eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi} \right)_K \right) \mathrm{d}x \, \mathrm{d}t \\
&\lesssim \sum_{i=1}^d \left(\int_0^T \sum_K \int_K r_K^2 \right)^{1/2} \Bigg[\left(\int_0^T \sum_{K\in\mathcal{T}} \int_K \widehat{|\overline{u_{i,h}}|}_K^2 \right)^{1/2} \left\| \Delta_h^{(i)} \Pi_{\mathcal{T}}\phi \right\|_{L^\infty L^\infty} \\
&\qquad + \left(\int_0^T \sum_{K\in\mathcal{T}} \int_K \left(\eth_{\mathcal{T}}^{(i)} \overline{u_{i,h}} \right)^2 \right)^{1/2} \left\| \widehat{\eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi} \right\|_{L^\infty L^\infty} \Bigg] \\
&\lesssim h \sum_{i=1}^d \|r_h\|_{L^2L^2} \left(\left\| \Delta_h^{(i)} \Pi_{\mathcal{T}}\phi \right\|_{L^\infty L^\infty} \|u_{i,h}\|_{L^2L^2} + \left\| \eth_{\mathcal{E}}^{(i)} \Pi_{\mathcal{T}}\phi \right\|_{L^\infty L^\infty} \left\| \eth_{\mathcal{E}}^{(i)} u_{i,h} \right\|_{L^2L^2} \right) \\
&\lesssim h \|r_h\|_{L^2L^2} \left(\|\Delta_h \Pi_{\mathcal{T}}\phi\|_{L^\infty L^\infty} \|\mathbf{u}_h\|_{L^2L^2} + \|\nabla_{\mathcal{E}} \Pi_{\mathcal{T}}\phi\|_{L^\infty L^\infty} \|\nabla_{\mathcal{E}} \mathbf{u}_h\|_{L^2L^2} \right) \\
&\lesssim h \|r_h\|_{L^2L^2}
\end{aligned}
$$

Consequently, applying the density estimate (3.6b), and the momentum estimate (3.6c) indicates

$$E_1(r_h) \lesssim h^{\beta}, \quad \beta = 1 - \frac{\varepsilon + 2}{2\gamma} > 0, \text{ provided } \varepsilon < 2(\gamma - 1),$$

for r_h being ϱ_h or $\varrho_h u_{i,h}$, $i = 1, \dots, d$.

- Secondly, we deal with the error term E_2. In accordance with (2.9), we have

$$E_2(r_h) \lesssim h \sum_{\sigma \in \mathcal{E}} \int_{\sigma} | [\![\mathbf{u}_h]\!] \cdot \mathbf{n} \, [\![r_h]\!] | \, \mathrm{d}Sx \, \mathrm{d}t.$$

For r_h being ϱ_h, we further write

$$\begin{aligned} E_2(\varrho_h) &\lesssim h \left(\int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} [\![\mathbf{u}_h]\!]^2 \, \mathrm{d}Sx \, \mathrm{d}t \right)^{1/2} \left(\int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} [\![\varrho_h]\!]^2 \, \mathrm{d}Sx \, \mathrm{d}t \right)^{1/2} \\ &\lesssim h h^{1/2} \left(\int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} \overline{\varrho_h}^2 \, \mathrm{d}Sx \, \mathrm{d}t \right)^{1/2} \\ &\lesssim h^{3/2} h^{-1/2} \left\| \varrho_h \right\|_{L^2 L^2} \lesssim h^{\beta}, \quad \beta = 1 - \frac{\varepsilon + 2}{2\gamma} > 0, \text{ as soon as } \varepsilon < 2(\gamma - 1). \end{aligned}$$

Here we have used Hölder's inequality, (3.5d), (3.6b), and the fact $| [\![\varrho_h]\!] | < 2\overline{\varrho_h}$.

For r_h being $\varrho_h u_{i,h}$, we get

$$E_2(\varrho_h \mathbf{u}_h) \lesssim h \int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} | [\![\mathbf{u}_h]\!] \cdot \mathbf{n} | \; | [\![\varrho_h]\!] \, \overline{\mathbf{u}_h} + [\![\mathbf{u}_h]\!] \, \overline{\varrho_h} | \, \mathrm{d}Sx \, \mathrm{d}t := T_1 + T_2.$$

To control the residual term T_1 we apply Hölder's inequality, (3.5a), (3.5g), inverse estimate (2.13) and the inequality $| [\![\varrho_h]\!] | < 2\overline{\varrho_h}$ to obtain

$$\begin{aligned} T_1 &\lesssim h \int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} | [\![\mathbf{u}_h]\!] \cdot \mathbf{n} | \overline{\varrho_h} | \overline{\mathbf{u}_h} | \, \mathrm{d}Sx \, \mathrm{d}t \\ &\lesssim h \left(\int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} \overline{\varrho_h} \, [\![\mathbf{u}_h]\!]^2 \, \mathrm{d}Sx \right)^{1/2} \left(\int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} \overline{\varrho_h} | \overline{\mathbf{u}_h} |^2 \, \mathrm{d}Sx \right)^{1/2} \\ &\lesssim h^{(1-\varepsilon)/2}. \end{aligned}$$

Further, applying (3.5g) we can control the residual term T_2 as

$$T_2 = h \int_0^T \sum_{\sigma \in \mathcal{E}} \int_{\sigma} | [\![\mathbf{u}_h]\!] \cdot \mathbf{n} | \, [\![\mathbf{u}_h]\!] \, | \overline{\varrho_h} \, \mathrm{d}Sx \, \mathrm{d}t \lesssim h^{1-\varepsilon}.$$

Therefore, we claim that provided $\varepsilon < 2(\gamma-1)$ we have

$$E_2(r_h) \lesssim h^{\beta}, \quad \beta > 0$$

for r_h being ϱ_h or $\varrho_h u_{i,h}$, $i = 1,\dots,d$.

• Next, we consider the error term E_3. Analogously as above, the integration by parts formula (2.8a), Hölder's inequality, and the interpolation error (2.12) yield

$$\begin{aligned} E_3(r_h) = h^{\varepsilon} \int_0^T \sum_{\sigma\in\mathcal{E}} \int_\sigma [\![r_h]\!] \, [\![\Pi_{\mathcal{T}}\phi]\!] \, \mathrm{d}Sx \, \mathrm{d}t &= -h^{\varepsilon+1} \int_0^T \int_\Omega r_h \Delta_h \Pi_{\mathcal{T}}\phi \, \mathrm{d}x \, \mathrm{d}t \\ &\lesssim h^{\varepsilon+1} \left\| r_h \right\|_{L^1L^1} (\left\|\phi\right\|_{C^2} + h \left\|\phi\right\|_{C^3}) \lesssim h^{\varepsilon+1} \left\| r_h \right\|_{L^1L^1}. \end{aligned}$$

Furthermore, using the estimates (3.5b) and (3.5c) we can conclude for r_h being ϱ_h or $\varrho_h u_{i,h}$, $i = 1,\dots,d$, that

$$E_3(r_h) \lesssim h^{\varepsilon+1}.$$

• Finally, using the estimates of kinetic energy (3.5a) and momentum (3.5c) together with interpolation error (2.11) we obtain for r_h being ϱ_h or $\varrho_h u_{i,h}$, $i = 1,\dots,d$ that

$$E_4(r_h) = \int_0^T \int_\Omega r_h \mathbf{u}_h \cdot \big(\nabla_x \phi - \nabla_{\mathcal{T}} \Pi_{\mathcal{D}} (\Pi_{\mathcal{T}}\phi)\big) \, \mathrm{d}x \, \mathrm{d}t \lesssim h \left\|\phi\right\|_{C^2} \left\| r_h \mathbf{u}_h \right\|_{L^1L^1} \lesssim h \left\| r_h \mathbf{u}_h \right\|_{L^\infty L^1} \lesssim h.$$

Consequently, we conclude the consistency formulae of the convective terms in both equations (2.3a) and (2.3b), by collecting the above estimates of the four terms E_j ($j = 1,\dots,4$):

$$\int_\Omega \varrho_h \mathbf{u}_h \cdot \nabla_x \phi \, \mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma F[\varrho_h, \mathbf{u}_h] \, [\![\Pi_{\mathcal{T}}\phi]\!] \, \mathrm{d}Sx \lesssim h^{\beta_1}, \tag{4.4a}$$

$$\int_\Omega \varrho_h \mathbf{u}_h \otimes \mathbf{u}_h : \nabla_x \boldsymbol{\phi} \, \mathrm{d}x - \sum_{\sigma\in\mathcal{E}} \int_\sigma F[\varrho_h \mathbf{u}_h, \mathbf{u}_h] \, [\![\Pi_{\mathcal{T}}\boldsymbol{\phi}]\!] \, \mathrm{d}Sx \lesssim h^{\beta_2}, \tag{4.4b}$$

for some β_1, $\beta_2 > 0$ provided $\varepsilon < \min\{1, 2(\gamma-1)\}$.

Step 3 – viscosity terms:
In accordance with (2.10) and (3.5d), we have the control of the viscosity terms. Indeed, we have

$$\begin{aligned} &\int_0^T \int_\Omega \nabla_{\mathcal{E}} \mathbf{u}_h : \nabla_x \boldsymbol{\phi} \, \mathrm{d}x \, \mathrm{d}t - \int_0^T \sum_{\sigma\in\mathcal{E}} \int_\sigma \frac{1}{d_\sigma} [\![\mathbf{u}_h]\!] \cdot [\![\Pi_{\mathcal{T}}\boldsymbol{\phi}]\!] \, \mathrm{d}Sx \, \mathrm{d}t \\ &= \int_0^T \int_\Omega \nabla_{\mathcal{E}} \mathbf{u}_h : (\nabla_x \boldsymbol{\phi} - \nabla_{\mathcal{E}} \Pi_{\mathcal{T}}\boldsymbol{\phi}) \, \mathrm{d}x \, \mathrm{d}t \lesssim \left\| \nabla_{\mathcal{E}} \mathbf{u}_h \right\|_{L^2L^2} h \left\|\boldsymbol{\phi}\right\|_{C^2} \lesssim h, \end{aligned} \tag{4.5a}$$

and for the divergence term, we get

$$\begin{aligned}&\int_0^T \int_\Omega \mathrm{div}_h \mathbf{u}_h \, \mathrm{div}_h\big(\Pi_{\mathcal{T}}\boldsymbol{\phi}\big) \, \mathrm{d}x - \int_0^T \int_\Omega \mathrm{div}_h \mathbf{u}_h \, \mathrm{div}_x \boldsymbol{\phi} \, \mathrm{d}x \, \mathrm{d}t \\ &= \int_0^T \int_\Omega \mathrm{div}_h \mathbf{u}_h \left(\mathrm{div}_h\big(\Pi_{\mathcal{T}}\boldsymbol{\phi}\big) - \mathrm{div}_x \boldsymbol{\phi} \right) \mathrm{d}x \, \mathrm{d}t \lesssim \left\| \mathrm{div}_h \mathbf{u}_h \right\|_{L^2 L^2} h \left\| \boldsymbol{\phi} \right\|_{C^2} \lesssim h,\end{aligned} \tag{4.5b}$$

by using (3.5e) and (2.11).

Step 4 – pressure term:
The pressure term can be controlled by using the integration by parts formula (2.7), the interpolation error (2.11), and the estimate (3.5b), i.e.,

$$\begin{aligned}&\int_0^T \sum_{\sigma \in \mathcal{E}} \int_\sigma \overline{p}_h \mathbf{n} \cdot [\![\Pi_{\mathcal{T}}\boldsymbol{\phi}]\!] \, \mathrm{d}Sx \, \mathrm{d}t - \int_0^T \int_\Omega p_h \mathrm{div}_x \boldsymbol{\phi} \, \mathrm{d}x \, \mathrm{d}t \\ &\qquad = -\int_0^T \sum_{\sigma \in \mathcal{E}} \int_\sigma \overline{\Pi_{\mathcal{T}}\boldsymbol{\phi}} \cdot \mathbf{n} \, [\![p_h]\!] \, \mathrm{d}Sx \, \mathrm{d}t - \int_0^T \sum_{K \in \mathcal{T}} \int_K p_h \mathrm{div}_x \boldsymbol{\phi} \, \mathrm{d}x \, \mathrm{d}t \\ &\qquad = \int_0^T \sum_{K \in \mathcal{T}} p_K \sum_{\sigma \in \mathcal{E}(K)} \int_\sigma \overline{\Pi_{\mathcal{T}}\boldsymbol{\phi}} \cdot \mathbf{n} \, \mathrm{d}Sx \, \mathrm{d}t - \int_0^T \sum_{K \in \mathcal{T}} \int_K p_h \mathrm{div}_x \boldsymbol{\phi} \, \mathrm{d}x \, \mathrm{d}t \\ &\qquad = \int_0^T \sum_{K \in \mathcal{T}} \int_K p_h \left(\mathrm{div}_h\big(\Pi_{\mathcal{T}}\boldsymbol{\phi}\big) - \mathrm{div}_x \boldsymbol{\phi} \right) \, \mathrm{d}x \, \mathrm{d}t \\ &\qquad \lesssim \left\| p_h \right\|_{L^\infty L^1} h \left\| \boldsymbol{\phi} \right\|_{C^2} \lesssim h.\end{aligned} \tag{4.6}$$

Collecting the inequalities (4.3)–(4.6) we complete the proof of Theorem 4.1. □

5 Convergence

5.1 Convergence to dissipative measure-valued solution

In this section, we show that any Young measure generated by a family of numerical solutions is a (DMV) solution in the sense of Definition in 1.1.

Theorem 5.1. *Let $\{(\varrho_h^n, \mathbf{u}_h^n)\}_{n=1}^{N_T}$ be a family of numerical solutions obtained by the scheme* (2.3), *with $\Delta t \approx h$, $1 < \gamma < 2$, $0 < \varepsilon < \min\{1, 2(\gamma - 1)\}$, and the initial data satisfying*

$$\varrho_0 \in L^\gamma(\Omega), \ \ \varrho_0 > 0, \ \ \mathbf{u}_0 \in L^2(\Omega; R^d).$$

Then any Young measure $\{\mathcal{V}_{t,x}\}_{(t,x)\in(0,T)\times\Omega}$ generated by $\varrho_h^n, \mathbf{u}_h^n$ for $h \to 0$ represents a dissipative measure-valued solution of the Navier–Stokes system (1.1) in the sense of Definition 1.1.

The rest of the section is devoted to the proof of Theorem 5.1.

5.1.1 Weak limit

We may use the energy estimates (3.3) to deduce that, at least for suitable subsequences,

$$\begin{aligned}
\varrho_h \to \varrho \text{ weakly-(*) in } L^\infty(0,T;L^\gamma(\Omega)),\ \varrho \geq 0 \\
\mathbf{u}_h \to \mathbf{u} \text{ weakly in } L^2((0,T)\times\Omega;R^d), \\
\text{where } \mathbf{u} \in L^2(0,T;W^{1,2}(\Omega)),\ \nabla_{\mathcal{E}}\mathbf{u}_h \to \nabla_x\mathbf{u} \text{ weakly in } L^2((0,T)\times\Omega;R^{d\times d}), \\
\varrho_h\mathbf{u}_h \to \widetilde{\varrho\mathbf{u}} \text{ weakly-(*) in } L^\infty(0,T;L^{\frac{2\gamma}{\gamma+1}}(\Omega;R^d)).
\end{aligned}$$

where the superscript '$\sim$' denotes the L^1-weak limit.

Note that, the limit functions satisfy the equation of continuity in the form

$$-\int_\Omega \varrho_0\phi(0,\cdot)\,\mathrm{d}x = \int_0^T\int_\Omega \left[\varrho\partial_t\phi + \widetilde{\varrho\mathbf{u}}\cdot\nabla_x\phi\right]\,\mathrm{d}x\,\mathrm{d}t \tag{5.1}$$

for any test function $\phi \in C_c^\infty([0,\infty)\times\Omega)$. It follows from (5.1) that $\varrho \in C_{\text{weak}}([0,T];L^\gamma(\Omega))$; whence (5.1) can be rewritten as

$$\left[\int_\Omega \varrho\phi(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau} = \int_0^\tau\int_\Omega \left[\varrho\partial_t\phi + \widetilde{\varrho\mathbf{u}}\cdot\nabla_x\phi\right]\,\mathrm{d}x\,\mathrm{d}t \tag{5.2}$$

for any $0 \leq \tau \leq T$ and any $\phi \in C^\infty([0,T]\times\Omega)$.

5.1.2 Young measure generated by numerical solutions

The energy stability (Theorem 3.3) together with the consistency (Theorem 4.1) provide a suitable platform for the use of the theory of measure-valued solutions developed in [7]. In accordance with the weak convergence statement derived in the preceding part, the family $[\varrho_h,\mathbf{u}_h]$ generates a Young measure - a parameterized measure [2, 19]

$$\mathcal{V}_{t,x} \in L^\infty((0,T)\times\Omega;\mathcal{P}([0,\infty)\times R^d)) \text{ for a.e. } (t,x)\in(0,T)\times\Omega,$$

such that

$$\langle\mathcal{V}_{t,x}, g(\varrho,\mathbf{u})\rangle = \widetilde{g(\varrho,\mathbf{u})}(t,x) \text{ for a.e. } (t,x)\in(0,T)\times\Omega,$$

for any $g \in C([0,\infty)\times R^d)$ such that

$$g(\varrho_h,\mathbf{u}_h) \to \widetilde{g(\varrho,\mathbf{u})} \text{ weakly in } L^1((0,T)\times\Omega).$$

Accordingly, the equation of continuity (5.2) can be written as

$$\left[\int_\Omega \varrho\phi(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau} = \int_0^\tau\int_\Omega \left[\varrho\partial_t\phi + \langle\mathcal{V}_{t,x},\varrho\mathbf{u}\rangle\cdot\nabla_x\phi\right]\,\mathrm{d}x\,\mathrm{d}t \tag{5.3}$$

For the consistency formulation of the momentum equation (4.2), we apply a similar treatment,

$$\varrho_h \mathbf{u}_h \otimes \mathbf{u}_h + p(\varrho_h)\mathbb{I} \to \{\varrho \mathbf{u} \otimes \mathbf{u} + p(\varrho)\mathbb{I}\} \text{ weakly-(*) in } [L^\infty(0,T;\mathcal{M}(\Omega))]^{d\times d};$$

whence letting $h \to 0$ in (4.2) gives rise to

$$\left[\int_\Omega \langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\rangle \cdot \boldsymbol{\phi}(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau} = \int_0^\tau \int_\Omega \Big[\langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\rangle \cdot \partial_t\boldsymbol{\phi} + \{\varrho\mathbf{u}\otimes\mathbf{u} + p(\varrho)\mathbb{I}\} : \nabla_x\boldsymbol{\phi}\Big]\,\mathrm{d}x\ \mathrm{d}t \\ - \int_0^\tau \int_\Omega \Big[\mu\nabla_x\mathbf{u} : \nabla_x\boldsymbol{\phi} + (\mu+\lambda)\mathrm{div}_x\mathbf{u}\,\mathrm{div}_x\boldsymbol{\phi}\Big]\,\mathrm{d}x\,\mathrm{d}t,$$

or, rewritten as

$$\left[\int_\Omega \langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\rangle \cdot \boldsymbol{\phi}(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau} = \int_0^\tau \int_\Omega \Big[\langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\rangle \cdot \partial_t\boldsymbol{\phi} + \{\varrho\mathbf{u}\otimes\mathbf{u} + p(\varrho)\mathbb{I}\} : \nabla_x\boldsymbol{\phi}\Big]\,\mathrm{d}x\ \mathrm{d}t \\ - \int_0^\tau \int_\Omega \Big[\mu\nabla_x\mathbf{u} : \nabla_x\boldsymbol{\phi} + (\mu+\lambda)\mathrm{div}_x\mathbf{u}\,\mathrm{div}_x\boldsymbol{\phi}\Big]\,\mathrm{d}x\,\mathrm{d}t, \tag{5.4}$$

for any $0 \le \tau \le T$, $\phi \in C_c^\infty([0,T]\times\Omega; R^d)$, where we have set

$$\mathcal{V}_{0,x} = \delta_{[\varrho_0(x),\mathbf{u}_0(x)]}.$$

Next, we introduce the *concentration remainder*

$$\mathcal{R} = \{\varrho\mathbf{u}\otimes\mathbf{u} + p(\varrho)\mathbb{I}\} - \left\langle \mathcal{V}_{t,x}; \widetilde{\varrho\mathbf{u}\otimes\mathbf{u}} + \widetilde{p(\varrho)}\mathbb{I}\right\rangle \in [L^\infty(0,T;\mathcal{M}(\Omega))]^{d\times d},$$

and rewrite (5.4) in the form

$$\left[\int_\Omega \langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\rangle \cdot \boldsymbol{\phi}(t,\cdot)\,\mathrm{d}x\right]_{t=0}^{t=\tau} \\ = \int_0^\tau \int_\Omega \Big[\langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\rangle \cdot \partial_t\boldsymbol{\phi} + \langle \mathcal{V}_{t,x}; \varrho\mathbf{u}\otimes\mathbf{u}\rangle : \nabla_x\boldsymbol{\phi} + \langle \mathcal{V}_{t,x}, p(\varrho)\rangle\,\mathrm{div}_x\boldsymbol{\phi}\Big]\,\mathrm{d}x\ \mathrm{d}t \\ - \int_0^\tau \int_\Omega \Big[\mu\nabla_x\mathbf{u} : \nabla_x\boldsymbol{\phi} + (\mu+\lambda)\mathrm{div}_x\mathbf{u}\cdot\mathrm{div}_x\boldsymbol{\phi}\Big]\,\mathrm{d}x\,\mathrm{d}t + \int_0^\tau\int_\Omega \mathcal{R} : \nabla_x\boldsymbol{\phi}\,\mathrm{d}x\,\mathrm{d}t \tag{5.5}$$

for any $0 \le \tau \le T$, $\boldsymbol{\phi} \in C_c^\infty([0,T]\times\Omega; R^d)$.

Similarly, the energy inequality (3.3) can be written as

$$\left[\int_\Omega \frac{1}{2}\left\langle \mathcal{V}_{t,x}; \varrho|\mathbf{u}|^2 + \mathcal{H}(\varrho)\right\rangle\,\mathrm{d}x\right]_{t=0}^{t=\tau} + \int_0^\tau\int_\Omega \left(\mu|\nabla_x\mathbf{u}|^2 + (\mu+\lambda)|\mathrm{div}_x\mathbf{u}|^2\right)\,\mathrm{d}x\,\mathrm{d}t + \mathcal{D}(\tau) \le 0 \tag{5.6}$$

for a.e. $\tau \in [0,T]$, with the *dissipation defect* $\mathcal{D}$ satisfying

$$\int_0^\tau \|\mathcal{R}\|_{\mathcal{M}(\Omega)}\ \mathrm{d}t \le \int_0^\tau \mathcal{D}(t)\ \mathrm{d}t,\ \mathcal{D}(\tau) \ge \liminf_{h\to 0}\int_0^\tau \|\nabla_\mathcal{E}\mathbf{u}_h\|_{L^2}^2\ \mathrm{d}t - \int_0^\tau\int_\Omega |\nabla_x\mathbf{u}|^2\,\mathrm{d}x\ \mathrm{d}t, \tag{5.7}$$

cf. [7, Lemma 2.1].

Relations (5.3), (5.5)–(5.7) imply that the Young measure $\{\mathcal{V}_{t,x}\}_{t,x\in(0,T)\times\Omega}$ represents a dissipative measure-valued solution of the Navier–Stokes system (1.1) in the sense of Definition 1.1. Seeing that validity of (5.5) as well as the bound on the dissipation remainder (5.7) can be extended to the class of test functions $\phi \in C^1([0,T]\times\Omega; R^d)$, we have proved Theorem 5.1.

5.2 Convergence to strong solution

In the previous subsection, we have shown that the numerical solution generates the dissipative measure-valued solution. We admit the conclusion of Theorem 5.1 is rather weak, also due to the non-uniqueness of Young measure, however, we may directly use the weak-strong uniqueness principle established in [7, Theorem 4.1] to obtain our final convergence result.

Theorem 5.2 (Convergence to strong solution)**.** *In addition to the hypotheses of Theorem 5.1, suppose that the Navier–Stokes system (1.1) endowed with the initial data $(\varrho_0, \mathbf{u}_0)$ admits a regular solution $(\varrho, \mathbf{u})$ belonging to the class*

$$\varrho, \nabla_x\varrho, \mathbf{u}, \nabla_x\mathbf{u} \in C([0,T]\times\Omega),\ \partial_t\mathbf{u} \in L^2\left(0,T; C(\Omega; R^d)\right),\ \varrho > 0.$$

Then

$$\varrho_h \to \varrho \text{ (strongly) in } L^\gamma\left((0,T)\times\Omega\right),\ \mathbf{u}_h \to \mathbf{u} \text{ (strongly) in } L^2\left((0,T)\times\Omega; R^d\right).$$

Indeed, the weak–strong uniqueness implies that the Young measure generated by the family of numerical solutions coincides at a.a. point (t,x) with the Dirac mass supported by the smooth solution of the problem. In particular, the numerical solutions converge strongly and no oscillations occur.

Remark 5.3. We have constructed solution on a space-periodic domain Ω. When considering a polyhedral domain, the existence of smooth solutions in the case of the no–slip boundary condition may be a delicate task. To avoid this problem, one has to approximate a smooth domain by a family of polyhedral domains exactly as in [9]. Note, however, this problem does not occur in the case of periodic boundary conditions.

Remark 5.4. If, in addition, we assume the density is uniformly bounded, the results of Theorems 5.1 and 5.2 remain valid on an unstructured grid as well. Indeed, the only difference of the proof would be in showing the consistency of the convective terms in (4.4). More precisely, since the discrete operators $\eth_{\mathcal{E}}^{(i)} r_h$, $\eth_{\mathcal{T}}^{(i)} q_{i,h}$ and $\Delta_h^{(i)} r_h$ can not be defined on an unstructured grid, the estimate of the error terms $E_1(\varrho_h)$ and $E_1(\varrho_h\mathbf{u}_h)$ would be done without the discrete integration by parts thanks to $L^\infty-$bound on the density. Moreover, in view of the conditional regularity result [20], we obtain the unconditional convergence to the strong solution as the (DMV) solution with bounded density is regular. We leave the details to the interested reader.

Conclusion

We have studied a finite volume method for the multi–dimensional compressible isentropic Navier–Stokes equations on regular quadrilateral mesh in a periodic domain. The main feature of the scheme is the artificial diffusion in the numerical flux function (2.1), which provides more regularity on the discrete density. The solutions of the scheme were shown to exist while preserving the positivity of the discrete density. Moreover, we have shown the stability of the scheme by deriving the unconditional balance of the discrete total energy in Theorem 3.3. Furthermore, we have established the consistency formulation provided the artificial diffusion coefficient is large enough, see Theorem 4.1. Finally, we have shown in Theorem 5.1 that the numerical solutions of the scheme (2.3) generate a (DMV) solution of the Navier–Stokes system (1.1). In addition, using the recent result on the (DMV)–strong uniqueness principle, we have proven the convergence to the strong solution on its lifespan, cf. Theorem 5.2. To the best of our knowledge, this is the first rigorous result concerning convergence of a finite volume method for the compressible isentropic Navier–Stokes equations in the multi–dimensional setting.

References

[1] G. Ansanay-Alex, F. Babik, J. C. Latché, and D. Vola An L2–stable approximation of the Navier–Stokes convection operator for low–order non-conforming finite elements *Int. J. Numer. Meth. Fluids*, **66**: 555–580, 2011.

[2] J.M. Ball. A version of the fundamental theorem for Young measures. *In Lect. Notes in Physics 344, Springer-Verlag*, pages 207–215, 1989.

[3] P. G. Ciarlet. The Finite Element Method for Elliptic Problems. *Classics in Applied Mathematics, Society for Industrial and Applied Mathematics*, 2002.

[4] V. Dolejší and M. Feistauer. Discontinuous Galerkin method. *volume 48 of Springer Series in Computational Mathematics, Springer*, Cham, 2015. xiv+572 pp. Analysis and applications to compressible flow.

[5] M. Feistauer. Mathematical methods in fluid dynamics. volume 67 of Pitman Monographs and Surveys in Pure and Applied Mathematics. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1993. xiv+657 pp.

[6] R. Eymard, T. Gallouët, and R. Herbin. Finite volume methods. *Handbook of numerical analysis* **7**: 713–1018, 2000.

[7] E. Feireisl, P. Gwiazda, A. Świerczewska-Gwiazda, and E. Wiedemann. Dissipative measure-valued solutions to the compressible Navier–Stokes system. *Calc. Var. Partial Dif.* **55**(6): 55–141, 2016.

[8] E. Feireisl, T. Karper, and A. Novotný. A convergent numerical method for the Navier–Stokes–Fourier system. *IMA J. Numer. Anal.* **36**(4): 1477–1535, 2016.

[9] E. Feireisl, and M. Lukáčová-Medvid'ová. Convergence of a mixed finite element–finite volume scheme for the isentropic Navier–Stokes system via the dissipative measure–valued solutions. *Found. Comput. Math.* **18**(3): 703–730, 2018.

[10] E. Feireisl, M. Lukáčová-Medvid'ová, and H. Mizerová. A finite volume scheme for the Euler system inspired by the two velocities approach. ArXiv: `arxiv.org/abs/1805.05072`, 2018.

[11] E. Feireisl, and A. Novotný. *Singular limits in thermodynamics of viscous fluids.* Birkhäuser–Basel, second edition, 2017.

[12] T. Gallouët, L. Gastaldo, R. Herbin, and J. C. Latché. An unconditionally stable pressure correction scheme for the compressible barotropic Navier–Stokes equations. *ESAIM: M2AN* , **42**(2): 303–331, 2008.

[13] T. Gallouët, R. Herbin, D. Maltese, and A. Novotný. Implicit MAC scheme for compressible Navier–Stokes equations: Unconditional error estimates. Preprint, 2016.

[14] J. Haack, S. Jin and J. G.Liu. An all–speed asymptotic–preserving method for the isentropic Euler and Navier-Stokes equations. *Commun. Comput. Phys.* **12**(4): 955–980, 2012.

[15] R. Hošek and B. She. Stability and consistency of a finite difference scheme for compressible viscous isentropic flow in multi-dimension. To appear in J. Numer. Math., 2018.

[16] T. Karper. A convergent FEM-DG method for the compressible NavierStokes equations. *Numer. Math.* **125**(3): 441-510, 2013.

[17] H. Mizerová and B. She. Convergence of a finite difference scheme for the compressible viscous isentropic flow in multi-dimension. Preprint, 2018.

[18] A. Novotný and B. She. Convergence of a finite difference MAC scheme for the compressible Navier–Stokes system via dissipative measure valued solutions. Preprint, 2018.

[19] P. Pedregal. *Parametrized measures and variational principles.* Birkhäuser, Basel, 1997.

[20] Y. Sun, C. Wang, and Z. Zhang. A BealeKatoMajda blow–up criterion for the 3–D compressible Navier–Stokes equations. *J. Math. Pures. Appl.*, **95**(1): 36–47, 2011.