Journal of Computational Physics 517 (2024) 113365

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Efficient energy stable schemes for incompressible flows with variable density

Buyang Li [a], Jie Shen [b], Zongze Yang [a], Yanrong Zhang [a,*]

[a] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Hong Kong
[b] School of Mathematical Science, Eastern Institute of Technology, Ningbo, Zhejiang, 315200, China

ARTICLE INFO

MSC:
35Q35
65N30
65N12
76D07

Keywords:
Incompressible flows with variable density
Preserving lower bound
Energy stability
Upwind discontinuous Galerkin

ABSTRACT

We present novel numerical schemes for the incompressible Navier-Stokes equations with variable density and address two critical concerns, i.e., the preservation of lower density bounds and the preservation of energy inequality under the gravitational force. Spatial discretization is performed using upwind discontinuous Galerkin methods for density, and continuous finite element methods are used for velocity and pressure. Additionally, velocity fields are projected onto divergence-free Raviart-Thomas finite element space in the transport equation. This relaxes using divergence-free finite elements and therefore simplifies the implementation. The proposed schemes are tested on benchmark examples to demonstrate the excellent performance in accurately capturing the complex dynamics of these problems. Overall, the proposed numerical schemes exhibit enhanced stability, efficiency, and accuracy, making them suitable for modeling and simulating incompressible flows with variable density in practical applications.

1. Introduction

Incompressible viscous flows with variable density are described by the hyperbolic-parabolic system of partial differential equations

$$\rho_t + \mathbf{u}\cdot\nabla\rho = 0 \quad \text{in } \Omega\times(0,T], \tag{1.1}$$

$$\rho\left(\mathbf{u}_t + (\mathbf{u}\cdot\nabla)\mathbf{u}\right) + \nabla p - \nu\Delta\mathbf{u} = \mathbf{f} \quad \text{in } \Omega\times(0,T], \tag{1.2}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in } \Omega\times(0,T], \tag{1.3}$$

where Ω is a bounded domain in $\mathbb{R}^d$ ($d=2$ or 3) and $T>0$ is a fixed time. The unknowns are the density $\rho>0$, the velocity field $\mathbf{u}$ and the pressure p; ν is the dynamic viscosity coefficient; $\mathbf{f}$ represents gravity term in practical physical problems, which is expressed as $\mathbf{f}=[0,-\rho g]^T$ in two dimension. The system (1.1)-(1.3) is prescribed with initial conditions

$$\begin{cases} \rho(\mathbf{x},0) = \rho_0(\mathbf{x}) & \text{in } \Omega, \\ \mathbf{u}(\mathbf{x},0) = \mathbf{u}_0(\mathbf{x}) & \text{in } \Omega, \end{cases} \tag{1.4}$$

* Corresponding author.
E-mail addresses: buyang.li@polyu.edu.hk (B. Li), jshen@eitech.edu.cn (J. Shen), zongze.yang@polyu.edu.hk (Z. Yang), yanrongzhang_math@163.com (Y. Zhang).

https://doi.org/10.1016/j.jcp.2024.113365
Received 12 September 2023; Received in revised form 25 June 2024; Accepted 16 August 2024

Available online 22 August 2024

and boundary conditions

$$\begin{cases} \rho(\mathbf{x},t)|_{\Gamma_{\mathbf{u}(\mathbf{x},t)}} = h(\mathbf{x},t), \\ \mathbf{u}(\mathbf{x},t)|_{\Gamma} = \mathbf{g}(\mathbf{x},t), \end{cases} \tag{1.5}$$

where $\Gamma = \partial\Omega$, and for any velocity field $\mathbf{u}$, $\Gamma_{\mathbf{u}}$ is the inflow boundary defined by

$$\Gamma_{\mathbf{u}} := \{\mathbf{x} \in \Gamma : \mathbf{u}(\mathbf{x}) \cdot \mathbf{n} < 0\}$$

with $\mathbf{n}$ being the outward unit normal vector.

There exists an extensive literature on numerical schemes for solving the Navier-Stokes equations with constant density, including particularly decoupled methods, such as projection methods [7,10,31] and consistent splitting approaches [14,15,25,33], which decouple the computation of velocity and pressure. The Navier-Stokes equations with variable density, as given in equations (1.1)-(1.3), incorporate all the challenges of the Navier-Stokes equations with constant density, including the handling of nonlinear terms and the coupling of velocity and pressure through the divergence-free condition, also present additional challenges due to the presence of a transport equation. This could result in oscillations near the distinct interface, rendering it essential to maintain the lower bound of density for the sake of numerical stability and the prevention of unphysical phenomena.

Several algorithms have been proposed to extend the decoupled schemes to variable density fluid flows, e.g., see [1,3,8,11,12,16,22,27,28] and the references therein. In addition, [5] proposed a hybrid scheme combining finite volume and finite element methods for numerical simulation of variable density incompressible flows, which is flexible and allows for unstructured meshes, and the method proposed by [19] is a filtered time-stepping approach that enhances time accuracy from first to second order that requires only one backward Euler solve at each time step followed by adding a time filter.

On the other hand, there is only a limited literature on the issue of preserving bound of the density. For instance, a bound-preserving limiter is employed in [18] to preserve bounds for density. However, this limiter only provides a relatively weak form of bound preservation, as it only ensures that the average value of the density is bound preserving at the element level, rather than at every point within each element. The scarcity of previous research on this problem underscores its difficulty and the importance of further investigation in this area.

Another issue is the energy stability in the presence of a gravity term. Several numerical schemes [9,11,12,19,28] have been shown to be energy stable for variable density flows in the absence of a gravity term. To the best of the authors' knowledge, there is currently no available numerical scheme that can achieve unconditional energy stability in the presence of the gravity term. Additionally, the numerical analysis for this model poses a significant challenge, and the existing literature on this topic is relatively scarce. The works such as [2,4,6,13,17,24] have attempted to address this challenge. In particular, [24] introduces a mixed DG-CG method that ensures positivity and offers robust energy estimates along with a convergence analysis by adopting a discontinuous piecewise constant approximation for the density space.

The main purpose of this paper is to develop efficient, bound preserving and energy stable schemes for (1.1)-(1.3). Our schemes enjoy the following advantages:

- Without modifying the structure of the original equations or requiring any additional post-processing techniques, the proposed scheme preserves a lower bound on density;
- The schemes satisfy energy stability with added gravity terms based on the SAV approach [20,21,23,29,30];
- For the transport equation, spatial discretization is performed using upwind discontinuous Galerkin methods, which effectively control oscillations;
- The use of Raviart-Thomas elements to project the velocity fields in the transport equation eliminates the requirement of divergence-free finite elements (for solving the velocity equations), which are more computationally expensive and challenging to implement.

In addition to its enhanced stability, the overall algorithm is highly efficient, making it suitable for practical applications.

The remainder of the paper is structured as follows. Section 2 introduces equivalent formulations for incompressible Navier-Stokes equations with variable density. The numerical scheme is presented in Section 3. In Section 4, numerical experiments are provided to assess the performance of the proposed method. Finally, Some remarks are given in Section 5.

2. Reformulation of the model

A challenge in solving Navier-Stokes equations with variable density is that the trilinear term no longer satisfies the skew-symmetry property that is valid for Navier-Stokes equations with constant density. Specifically, for smooth enough $\mathbf{u}$ and $\mathbf{v}$ with $\mathbf{u} \cdot \mathbf{n}|_{\Gamma} = 0$, the following equation holds for incompressible Navier-Stokes equations with constant density:

$$\int_{\Omega} (\mathbf{u} \cdot \nabla)\mathbf{v} \cdot \mathbf{v} d\mathbf{x} = 0 \quad \text{for } \mathbf{u}, \mathbf{v} \text{ smooth enough and } \mathbf{u} \cdot \mathbf{n}|_{\Gamma} = 0.$$

This property is vital in the numerical schemes and analysis of the Navier-Stokes equations. However, it is no longer valid when ρ is a variable density.

To address this challenge, Guermond and Quartapelle [11] introduced a conserved form system

$$\rho_t + \nabla\cdot(\rho\mathbf{u}) = 0 \quad \text{in } \Omega\times(0,T], \tag{2.6}$$

$$\sigma(\sigma\mathbf{u})_t + (\rho\mathbf{u}\cdot\nabla)\mathbf{u} + \frac{\mathbf{u}}{2}\nabla\cdot(\rho\mathbf{u}) + \nabla p - \nu\Delta\mathbf{u} = \mathbf{f} \quad \text{in } \Omega\times(0,T], \tag{2.7}$$

$$\nabla\cdot\mathbf{u} = 0 \quad \text{in } \Omega\times(0,T], \tag{2.8}$$

where $\sigma = \sqrt{\rho}$. From (2.6) and (2.8), we can derive that

$$\sigma(\sigma\mathbf{u})_t = \rho\mathbf{u}_t + \frac{1}{2}\rho_t\mathbf{u} = \rho\mathbf{u}_t - \frac{1}{2}\nabla\cdot(\rho\mathbf{u})\mathbf{u}. \tag{2.9}$$

Hence, the system (2.6)-(2.8) is mathematically equivalent to the original system (1.1)-(1.3). In addition, the nonlinear terms now satisfy the following properties

$$\int_\Omega \rho\mathbf{u}\cdot\nabla\rho\,\mathrm{d}\mathbf{x} = 0 \quad \text{and} \quad \int_\Omega \rho\nabla\cdot\mathbf{u}\rho\,\mathrm{d}\mathbf{x} = 0, \tag{2.10}$$

$$\int_\Omega (\rho\mathbf{u}\cdot\nabla)\mathbf{v}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} + \frac{1}{2}\int_\Omega \nabla\cdot(\rho\mathbf{u})\mathbf{v}\cdot\mathbf{v}\,\mathrm{d}\mathbf{x} = 0, \tag{2.11}$$

assuming that $\mathbf{u}$ is divergence-free, and that $\rho, \mathbf{u}, \mathbf{v}$ are sufficiently smooth. Additionally, it is required that the normal component of $\mathbf{u}$ on the boundary Γ vanishes, i.e., $\mathbf{u}\cdot\mathbf{n}|_\Gamma = 0$. Furthermore, taking the inner product of (2.6) with 1 and $\rho(\mathbf{x},t)$ respectively, and taking the inner product of (2.7) with $\mathbf{u}(\mathbf{x},t)$, we obtain the following identities

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \rho(\cdot,t)\mathrm{d}\mathbf{x} = 0, \tag{2.12}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\rho(\cdot,t)\|^2 = 0, \tag{2.13}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{1}{2}\|\sigma(\cdot,t)\mathbf{u}(\cdot,t)\|^2 = -\nu\|\nabla\mathbf{u}(\cdot,t)\|^2 + \int_\Omega \mathbf{f}(\mathbf{x},t)\cdot\mathbf{u}(\mathbf{x},t)\mathrm{d}\mathbf{x}. \tag{2.14}$$

Next, we construct efficient and accurate numerical algorithms to preserve its inherent properties for incompressible flows with variable density.

Assume $\rho \geq \rho_m > 0$. To enforce the lower bound of the density, we introduce $\widetilde{\sigma}(\cdot,t) = \sqrt{\rho(\cdot,t) - \rho_m}$, then (2.6) can be rewritten as

$$\widetilde{\sigma}_t + \nabla\cdot(\widetilde{\sigma}\mathbf{u}) = 0, \quad \widetilde{\sigma}(\mathbf{x},0) = \sqrt{\rho(\mathbf{x},0) - \rho_m}, \quad \widetilde{\sigma}(\mathbf{x},t)\big|_{\Gamma_{\mathbf{u}(\mathbf{x},t)}} = \sqrt{h(\mathbf{x},t) - \rho_m}. \tag{2.15}$$

By applying the integration by parts formula and using (2.6), we can derive that

$$\begin{aligned} -\int_\Omega \mathbf{f}(\mathbf{x},t)\cdot\mathbf{u}(\mathbf{x},t)\mathrm{d}\mathbf{x} &= \int_\Omega \rho g\mathbf{e}_d\mathbf{u}(\mathbf{x},t)\mathrm{d}\mathbf{x} \\ &= \int_\Omega \rho g\nabla x_d\mathbf{u}(\mathbf{x},t)\mathrm{d}\mathbf{x} \\ &= -\int_\Omega \nabla\cdot(\rho\mathbf{u})gx_d\mathrm{d}\mathbf{x} \\ &= \int_\Omega \rho_t g x_d\mathrm{d}\mathbf{x}, \end{aligned} \tag{2.16}$$

where $\mathbf{e}_d = [0,\cdots,1]^T$, x_d represents the d-th component of $\mathbf{x} = (x_1,\cdots,x_d)$. Then (2.14) can be rewritten as the following energy dissipation law

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|\sigma(\cdot,t)\mathbf{u}(\cdot,t)\|^2 + \int_\Omega \rho(\cdot,t)gx_d\mathrm{d}\mathbf{x}\right) = -\nu\|\nabla\mathbf{u}(\cdot,t)\|^2. \tag{2.17}$$

We shall deal with the gravity term by using a SAV approach. More precisely, we introduce SAV

$$\eta(t) = \sqrt{E_1(\rho) + C_0} = \sqrt{\int_\Omega \rho(\cdot,t)gx_d\mathrm{d}\mathbf{x} + C_0}, \tag{2.18}$$

then we can rewrite the system (2.6)-(2.8) as the following equivalent form

$$\widetilde{\sigma}_t + \nabla\cdot(\widetilde{\sigma}\mathbf{u}) = 0, \tag{2.19}$$

$$\rho = \widetilde{\sigma}^2 + \rho_m, \quad \sigma = \sqrt{\rho}, \tag{2.20}$$

$$\sigma(\sigma\mathbf{u})_t + (\rho\mathbf{u}\cdot\nabla)\mathbf{u} + \frac{\mathbf{u}}{2}\nabla\cdot(\rho\mathbf{u}) + \nabla p - \nu\Delta\mathbf{u} - \frac{\eta(t)}{\sqrt{E_1(\rho)+C_0}}\mathbf{f} = \mathbf{0}, \tag{2.21}$$

$$\nabla\cdot\mathbf{u} = 0, \tag{2.22}$$

$$\eta_t = -\frac{1}{2\sqrt{E_1(\rho)+C_0}}(\mathbf{f},\mathbf{u}). \tag{2.23}$$

Note that the original idea of the SAV approach [20,21,23,29,30] is to introduce auxiliary variables in the nonlinear part of the equations to treat the nonlinear term explicitly while maintaining unconditional energy stability. However, in the above formulation, we introduce a SAV to handle the linear gravity term. This represents an extension of the SAV approach, and provides a new perspective on how to apply SAV to improve the numerical stability of Navier-Stokes equations with variable density.

3. Numerical schemes

In this section, we describe the numerical methods for the reformulated system (2.19)-(2.23).

3.1. Time discretization

We construct below two time discretization schemes with first- and second-order accuracy, respectively.

Scheme 1 (First-order scheme) Given $\left(\widetilde{\sigma}^n, \sigma^n, \mathbf{u}^n, \eta^n\right)$, we compute $\left(\widetilde{\sigma}^{n+1},\ \sigma^{n+1}, \rho^{n+1}, \mathbf{u}^{n+1}, p^{n+1}, \eta^{n+1}\right)$ via the following steps:

Step 1. Find a solution $\widetilde{\sigma}^{n+1}$ by solving a transport equation

$$\frac{\widetilde{\sigma}^{n+1}-\widetilde{\sigma}^n}{\delta t} + \nabla\cdot(\widetilde{\sigma}^{n+1}\mathbf{u}^n) = 0. \tag{3.24}$$

Step 2. Compute ρ^{n+1}, σ^{n+1} from

$$\rho^{n+1} = \left(\widetilde{\sigma}^{n+1}\right)^2 + \rho_m, \quad \sigma^{n+1} = \sqrt{\rho^{n+1}}. \tag{3.25}$$

Step 3. Solve $\mathbf{u}^{n+1}, p^{n+1}, \eta^{n+1}$ from

$$\sigma^{n+1}\frac{\sigma^{n+1}\mathbf{u}^{n+1}-\sigma^n\mathbf{u}^n}{\delta t} + (\rho^{n+1}\mathbf{u}^n\cdot\nabla)\mathbf{u}^{n+1} + \frac{\mathbf{u}^{n+1}}{2}\nabla\cdot(\rho^{n+1}\mathbf{u}^n) + \nabla p^{n+1} - \nu\Delta\mathbf{u}^{n+1} - \frac{\eta^{n+1}}{\sqrt{E_1(\rho^{n+1})+C_0}}\mathbf{f}^{n+1} = \mathbf{0}, \tag{3.26}$$

$$\nabla\cdot\mathbf{u}^{n+1} = 0, \tag{3.27}$$

$$\frac{\eta^{n+1}-\eta^n}{\delta t} = -\frac{1}{2\sqrt{E_1(\rho^{n+1})+C_0}}\left(\mathbf{f}^{n+1},\mathbf{u}^{n+1}\right). \tag{3.28}$$

Theorem 3.1. *The **Scheme 1** satisfies*

$$\rho^{n+1} \geq \rho_m, \tag{3.29}$$

$$\int_\Omega \widetilde{\sigma}^{n+1}\mathrm{d}\mathbf{x} = \int_\Omega \widetilde{\sigma}^n\mathrm{d}\mathbf{x} \tag{3.30}$$

and unconditionally stable in the sense that

$$E^{n+1} - E^n \leq -\delta t\nu\|\nabla u^{n+1}\|^2, \quad \forall\delta t, n \geq 0, \tag{3.31}$$

where

$$E^{n+1} = \frac{1}{2}\|\sigma^{n+1}\mathbf{u}^{n+1}\|^2 + |\eta^{n+1}|^2.$$

Proof. (3.29) is a direct consequence of (3.25). Taking the inner product of (3.24) with 1, we obtain (3.30).

Taking inner product of (3.26) with $\mathbf{u}^{n+1}$ and multiplying (3.28) by $2\eta^{n+1}$, summing them up, we obtain (3.31). □

Scheme 2 (Second-order scheme) Given $\left(\widetilde{\sigma}^{n-1}, \sigma^{n-1}, \mathbf{u}^{n-1}, \eta^{n-1}\right), \left(\widetilde{\sigma}^n, \sigma^n, \mathbf{u}^n, \eta^n\right)$, we compute $\left(\widetilde{\sigma}^{n+1}, \sigma^{n+1}, \rho^{n+1}, \mathbf{u}^{n+1}, p^{n+1}, \eta^{n+1}\right)$ via the following steps:

Step 1. Find a solution $\widetilde{\sigma}^{n+1}$ by solving a transport equation

$$\frac{3\widetilde{\sigma}^{n+1}-4\widetilde{\sigma}^n+\widetilde{\sigma}^{n-1}}{2\delta t} + \nabla\cdot(\widetilde{\sigma}^{n+1}\mathbf{u}^{\star,n}) = 0, \tag{3.32}$$

Step 2. Compute ρ^{n+1}, σ^{n+1} from

$$\rho^{n+1} = \left(\widetilde{\sigma}^{n+1}\right)^2 + \rho_m, \quad \sigma^{n+1} = \sqrt{\rho^{n+1}}, \tag{3.33}$$

Step 3. Solve $\mathbf{u}^{n+1}, p^{n+1}, \eta^{n+1}$ from

$$\sigma^{n+1}\frac{3\sigma^{n+1}\mathbf{u}^{n+1} - 4\sigma^n\mathbf{u}^n + \sigma^{n-1}\mathbf{u}^{n-1}}{2\delta t} + (\rho^{n+1}\mathbf{u}^{\star,n}\cdot\nabla)\mathbf{u}^{n+1} + \frac{\mathbf{u}^{n+1}}{2}\nabla\cdot(\rho^{n+1}\mathbf{u}^{\star,n}) + \nabla p^{n+1} - \nu\Delta\mathbf{u}^{n+1} - \frac{\eta^{n+1}}{\sqrt{E_1(\rho^{n+1}) + C_0}}\mathbf{f}^{n+1} = \mathbf{0}, \tag{3.34}$$

$$\nabla\cdot\mathbf{u}^{n+1} = 0, \tag{3.35}$$

$$\frac{3\eta^{n+1} - 4\eta^n + \eta^{n-1}}{2\delta t} = -\frac{1}{2\sqrt{E_1(\rho^{n+1}) + C_0}}\left(\mathbf{f}^{n+1}, \mathbf{u}^{n+1}\right), \tag{3.36}$$

where $\mathbf{u}^{\star,n}$ is any explicit second order extrapolation for $\mathbf{u}(t^{n+1})$, such as $\mathbf{u}^{\star,n} = 2\mathbf{u}^{n+1} - \mathbf{u}^n$.

Theorem 3.2. *The **Scheme 2** satisfies*

$$\rho^{n+1} \geq \rho_m, \tag{3.37}$$

$$\int_\Omega 3\widetilde{\sigma}^{n+1} - 4\widetilde{\sigma}^n + \widetilde{\sigma}^{n-1}\mathrm{d}\mathbf{x} = 0, \tag{3.38}$$

and unconditionally stable in the sense that

$$E^{n+1} - E^n \leq -\delta t\nu\|\nabla u^{n+1}\|^2, \quad \forall\delta t, n \geq 0, \tag{3.39}$$

where

$$E^{n+1} = \frac{1}{4}\left(\|\sigma^{n+1}\mathbf{u}^{n+1}\|^2 + \|2\sigma^{n+1}\mathbf{u}^{n+1} - \sigma^n\mathbf{u}^n\|^2\right) + \frac{1}{2}\left(|\eta^{n+1}|^2 + |2\eta^{n+1} - \eta^n|^2\right).$$

Proof. One can infer from (3.33) the result in (3.37). Taking the inner product of (3.32) with 1, we obtain (3.38).

Taking the inner product of (3.34) with $\mathbf{u}^{n+1}$ and multiplying (3.36) by $2\eta^{n+1}$, summing them up, we obtain (3.39). □

Remark 3.1. The proposed schemes preserve the lower-bound of the density without modifying the structure of the original equations. Moreover, the scheme does not require any additional post-processing techniques. This feature represents a significant advantage of the proposed schemes, as it can simplify the numerical implementation and reduce computational costs. Furthermore, preserving the lower bound of the density will be of critical importance in the error analysis.

3.2. Full discretization with FEM

Let $\Sigma = \{K\}$ be a shape regular quasi-uniform partition of Ω with mesh size h. We define

$$\mathbb{W}_h = \left\{w_h \in L^2(\Omega) : \ w_h|_K \in Q_{k_1}(K), \quad \forall K \in \Sigma\right\},$$
$$\mathbb{V}_h = \left\{\mathbf{v}_h \in C^0(\Omega)^d : \ \mathbf{v}_h|_K \in \left[Q_{k_2}(K)\right]^d, \quad \forall K \in \Sigma; \quad \mathbf{v}_h|_\Gamma = 0\right\},$$
$$\mathbb{Q}_h = \left\{q_h \in L_0^2(\Omega) \cap C^0(\Omega) : \ q_h|_K \in Q_{k_3}(K), \quad \forall K \in \Sigma\right\},$$

where $Q_{k_1}(K), Q_{k_2}(K)$, and $Q_{k_3}(K)$ are spaces of polynomials with degree k_1, k_2, k_3 in each direction. We shall approximate the (velocity, pressure) in $(\mathbb{V}_h, \mathbb{Q}_h)$, which satisfy the inf-sup condition. For the density, we shall use an upwind DG scheme in $\mathbb{W}_h$. To this end, for any function v continuous on each $K \in \Sigma$, we define

$$[\![v]\!] = \left(v^- - v^+\right)\mathbf{n}_e \tag{3.40}$$

to be the jump of the function v on an edge $e = \partial K^- \cap \partial K^+$, with $v^\pm$ denoting the trace of v from $K^\pm$ onto e, and $\mathbf{n}_e$ denoting the normal vector on e pointing towards K^+. This definition is independent of the choice of K^- and K^+, as switching K^- and K^+ does not change $[\![v]\!]$.

Let $\mathrm{RT}^1(K)$ denote $H(\mathrm{div}, \Omega)$-conforming Raviart-Thomas finite element spaces of order 1, namely,

$$\mathrm{RT}^1(\Sigma) := \left\{\mathbf{v} \in H(\mathrm{div}, \Omega) : \ \mathbf{v}|_K \in Q_1(K)^d + \mathbf{x}Q_1(K), \forall K \in \Sigma\right\}. \tag{3.41}$$

Additionally, we introduce the following finite element space:

$$\mathrm{RT}_0^1(\Sigma) := \left\{\mathbf{v}_h \in \mathrm{RT}^1(\Sigma) : \ \nabla\cdot\mathbf{v}_h = 0 \text{ in } \Omega \text{ and } \mathbf{v}_h\cdot\mathbf{n} = 0 \text{ on } \Gamma\right\}. \tag{3.42}$$

Let $P_h^{\mathrm{RT}} : L^2(\Omega)^d \to \mathrm{RT}_0^1(\Sigma)$ denote the L^2-orthogonal projection, defined by

$$\left(\mathbf{v} - P_h^{\mathrm{RT}}\mathbf{v}, \mathbf{z}_h\right) = 0 \quad \forall \mathbf{z}_h \in \mathrm{RT}_0^1(\Sigma), \quad \forall \mathbf{v} \in L^2(\Omega)^d. \tag{3.43}$$

Then, a fully discrete linearized FEM for **Scheme 1** is as follows.

Scheme 1_h **(Full discrete first-order scheme)** Find $\widetilde{\sigma}_h^{n+1}, \rho_h^{n+1}, \sigma_h^{n+1} \in \mathbb{W}_h$, $\mathbf{u}_h^{n+1} \in \mathbb{V}_h$, $p_h^{n+1} \in \mathbb{Q}_h$ and $\eta_h^{n+1} \in \mathbb{R}$ satisfying the equations:

$$\left(\frac{\widetilde{\sigma}_h^{n+1} - \widetilde{\sigma}_h^n}{\delta t}, w_h\right) + \left(\sum_e \left(\widehat{\widetilde{\sigma}}_h^{n+1} P_h^{\mathrm{RT}}\mathbf{u}_h^n, [\![w_h]\!]\right)_e - \left(\widetilde{\sigma}_h^{n+1} P_h^{\mathrm{RT}}\mathbf{u}_h^n, \nabla w_h\right)\right) = 0, \tag{3.44}$$

$$\rho_h^{n+1} = \left(\widetilde{\sigma}_h^{n+1}\right)^2 + \rho_m, \quad \sigma_h^{n+1} = \sqrt{\rho_h^{n+1}}, \tag{3.45}$$

$$\begin{aligned}&\left(\sigma_h^{n+1}\frac{\sigma_h^{n+1}\mathbf{u}_h^{n+1} - \sigma_h^n\mathbf{u}_h^n}{\delta t}, \mathbf{v}_h\right) + \left((\rho_h^{n+1}\mathbf{u}_h^n \cdot \nabla)\mathbf{u}_h^{n+1} + \frac{\mathbf{u}_h^{n+1}}{2}\nabla \cdot (\rho_h^{n+1}\mathbf{u}_h^n), \mathbf{v}_h\right) \\ &\quad + \left(\nabla p_h^{n+1}, \mathbf{v}_h\right) + \nu\left(\nabla\mathbf{u}_h^{n+1}, \nabla\mathbf{v}_h\right) - \frac{\eta_h^{n+1}}{\sqrt{E_1(\rho_h^{n+1}) + C_0}}\left(\mathbf{f}_h^{n+1}, \mathbf{v}_h\right) = \mathbf{0},\end{aligned} \tag{3.46}$$

$$\left(\mathbf{u}_h^{n+1}, \nabla q_h\right) = 0, \tag{3.47}$$

$$\frac{\eta_h^{n+1} - \eta_h^n}{\delta t} = -\frac{1}{2\sqrt{E_1(\rho_h^{n+1}) + C_0}}\left(\mathbf{f}_h^{n+1}, \mathbf{u}_h^{n+1}\right), \tag{3.48}$$

for all test function $(w_h, \mathbf{v}_h, q_h) \in \mathbb{W}_h \times \mathbb{V}_h \times \mathbb{Q}_h$, where $\widehat{\widetilde{\sigma}}_h^{n+1}$ is the numerical flux, which is a single-valued function defined at the cell interfaces and in general depending on the values of the numerical solution $\widetilde{\sigma}_h^{n+1}$ from both sides of the interface. Here, the numerical flux on an edge e can be chosen according to the upwind principle, namely,

$$\widehat{\widetilde{\sigma}}_h^{n+1} = \begin{cases} (\widetilde{\sigma}_h^{n+1})^- & \text{on } e_+^n, \\ (\widetilde{\sigma}_h^{n+1})^+ & \text{on } e_-^n, \end{cases} \tag{3.49}$$

where e_-^n and e_+^n denote the numerical inflow and outflow parts of an edge e at time $t = t^n$, defined by

$$e_-^n := \left\{\mathbf{x} \in e : \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n \cdot \mathbf{n}_e\right)(\mathbf{x}) \leqslant 0\right\} \quad \text{and} \quad e_+^n := \left\{\mathbf{x} \in e : \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n \cdot \mathbf{n}_e\right)(\mathbf{x}) > 0\right\}.$$

The upwind DG method used in the above is effective in controlling numerical oscillations and maintaining the stability of the numerical scheme, even in the presence of strong gradients and discontinuities in the solution. This approach provides a robust and accurate solution method for simulating variable-density Navier-Stokes equations. At the same time, we employ the Raviart-Thomas (RT) space projection method to approximate the velocity field in the transport equation. This approach maintains the accuracy and stability of the numerical solution and eliminates the need for using divergence-free finite elements, which can be computationally expensive and difficult to implement.

Theorem 3.3. ***Scheme*** 1_h *admits a unique solution.* ***Scheme*** 1_h *fulfills the following condition:*

$$\int_\Omega \widetilde{\sigma}_h^{n+1}\mathrm{d}\mathbf{x} = \int_\Omega \widetilde{\sigma}_h^n\mathrm{d}\mathbf{x}. \tag{3.50}$$

Moreover, it is unconditionally energy stable in the sense that

$$E_h^{n+1} - E_h^n \leq -\delta t\nu\|\nabla u_h^{n+1}\|^2, \quad \forall \delta t, n \geq 0, \tag{3.51}$$

where

$$E_h^{n+1} = \frac{1}{2}\|\sigma_h^{n+1}\mathbf{u}_h^{n+1}\|^2 + |\eta_h^{n+1}|^2.$$

Proof. Note that equation (3.44) has a unique solution $\widetilde{\sigma}_h^{n+1}$ if and only if the following homogeneous linear equation only has zero solution $\phi_h = 0$. Find $\phi_h \in \mathbb{W}_h$ such that

$$\left(\frac{\phi_h}{\delta t}, w_h\right) + \sum_e \left(\widehat{\phi}_h P_h^{\mathrm{RT}}\mathbf{u}_h^n, [\![w_h]\!]\right)_e - \left(\phi_h P_h^{\mathrm{RT}}\mathbf{u}_h^n, \nabla w_h\right) = 0, \quad \forall\, w_h \in \mathbb{W}_h.$$

Substituting $w_h = \phi_h$ into the homogeneous linear equation above and using integration by parts, we obtain

$$\frac{1}{\delta t}\|\phi_h\|^2 + \sum_e \left(\widehat{\phi}_h P_h^{\mathrm{RT}}\mathbf{u}_h^n, [\![\phi_h]\!]\right)_e - \sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n, \{\phi_h\}[\![\phi_h]\!]\right)_e = 0,$$

where $\{\phi_h\} = \frac{1}{2}(\phi_h^- + \phi_h^+)$ denotes the average trace of ϕ_h onto an edge e. Here we have used the following identity

$$\begin{aligned}\sum_K \left(\phi_h P_h^{\mathrm{RT}}\mathbf{u}_h^n, \nabla\phi_h\right)_K &= \frac{1}{2}\sum_K \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n, \nabla\phi_h^2\right)_K \\ &= \frac{1}{2}\sum_e \left(\left(-P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, (\phi_h^+)^2\right)_e + \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, (\phi_h^-)^2\right)_e\right) \\ &= \frac{1}{2}\sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, (\phi_h^-)^2 - (\phi_h^+)^2\right)_e \\ &= \sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n, \{\phi_h\}[\![\phi_h]\!]\right)_e.\end{aligned}$$

It is straightforward to verify the following relations:

$$\widehat{\phi}_h - \{\phi_h\} = \frac{\phi_h^- - \phi_h^+}{2} \text{ on } e_+^n \quad \text{and} \quad \widehat{\phi}_h - \{\phi_h\} = -\frac{\phi_h^- - \phi_h^+}{2} \text{ on } e_-^n. \tag{3.52}$$

Then we have

$$\begin{aligned}\sum_e \left(\widehat{\phi}_h P_h^{\mathrm{RT}}\mathbf{u}_h, [\![\phi_h]\!]\right)_e &= \sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h, \left(\{\phi_h\} + \frac{\phi_h^- - \phi_h^+}{2}\right)[\![\phi_h]\!]\right)_{e_+^n} + \left(P_h^{\mathrm{RT}}\mathbf{u}_h, \left(\{\phi_h\} - \frac{\phi_h^- - \phi_h^+}{2}\right)[\![\phi_h]\!]\right)_{e_-^n} \\ &= \sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h, \{\phi_h\}[\![\phi_h]\!]\right)_e + \left(P_h^{\mathrm{RT}}\mathbf{u}_h\cdot\mathbf{n}_e, \frac{(\phi_h^- - \phi_h^+)^2}{2}\right)_{e_+^n} - \left(P_h^{\mathrm{RT}}\mathbf{u}_h\cdot\mathbf{n}_e, \frac{(\phi_h^- - \phi_h^+)^2}{2}\right)_{e_-^n}.\end{aligned}$$

Therefore,

$$\frac{1}{\delta t}\|\phi_h\|^2 + \sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, \frac{(\phi_h^- - \phi_h^+)^2}{2}\right)_{e_+^n} - \sum_e \left(P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, \frac{(\phi_h^- - \phi_h^+)^2}{2}\right)_{e_-^n} = 0.$$

Since $\left(P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, (\phi_h^- - \phi_h^+)^2\right)_{e_+^n} \geqslant 0$ and $\left(P_h^{\mathrm{RT}}\mathbf{u}_h^n\cdot\mathbf{n}_e, (\phi_h^- - \phi_h^+)^2\right)_{e_-^n} \leqslant 0$, it follows that $\|\phi_h\| = 0$. This proves the existence and uniqueness of solutions to (3.44).

Since $\rho_h^{n+1} = \left(\widetilde{\sigma}_h^{n+1}\right)^2 + \rho_m > 0$ and $\sigma_h^{n+1} = \sqrt{\rho_h^{n+1}} > 0$, we can easily show that the homogeneous linear equations corresponding to (3.46)–(3.48) only have zero solutions. This proves the existence and uniqueness of solutions to (3.46)–(3.48).

Substituting $w_h = 1$ into (3.44), one can obtain result (3.50).

Substituting $\mathbf{v}_h = \mathbf{u}_h^{n+1}$ and $q_h = p_h^{n+1}$ into (3.46) and (3.47) respectively, and multiplying (3.48) with $2\eta_h^{n+1}$, one can derive the following relation:

$$\frac{1}{2\delta t}\left(\|\sigma_h^{n+1}\mathbf{u}_h^{n+1}\|^2 - \|\sigma_h^n\mathbf{u}_h^n\|^2 + \|\sigma_h^{n+1}\mathbf{u}_h^{n+1} - \sigma_h^n\mathbf{u}_h^n\|^2\right) + \nu\|\nabla\mathbf{u}_h^{n+1}\|^2 + \frac{1}{\delta t}\left(|\eta_h^{n+1}|^2 - |\eta_h^n|^2 + |\eta_h^{n+1} - \eta_h^n|^2\right) = 0,$$

which implies the energy inequality in (3.51). □

Similarly, we can construct full discrete version **Scheme** 2_h for the second-order scheme **Scheme** 2, and prove similar results as stated above for the fully discrete second-order scheme. For the sake of brevity, we omit the detail here.

The above scheme can be efficiently implemented as follows.

First, given $\mathbf{u}_h^n$, we can obtain $\Psi_h = P_h^{\mathrm{RT}}\mathbf{u}_h^n$ by solving the following mixed problem where $\phi_h \in \mathrm{Q}_{\mathrm{dG}}^1(K)$ is a Lagrange multiplier:

$$\left(\Psi_h, \mathbf{z}_h\right) + \left(\phi_h, \nabla\cdot\mathbf{z}_h\right) = \left(\mathbf{u}_h^n, \mathbf{z}_h\right), \quad \forall\mathbf{z}_h \in \mathrm{RT}^1(K), \tag{3.53}$$

$$\left(\nabla\cdot\Psi_h, \psi_h\right) = 0, \quad \forall\psi_h \in \mathrm{Q}_{\mathrm{dG}}^1(K). \tag{3.54}$$

Next, we set

$$\mathbf{u}_h^{n+1} = \mathbf{u}_{1,h}^{n+1} + \eta_h^{n+1}\mathbf{u}_{2,h}^{n+1}, \quad \mathbf{u}_{i,h}^{n+1}|_\Gamma = 0, \tag{3.55}$$

$$p_h^{n+1} = p_{1,h}^{n+1} + \eta_h^{n+1}p_{2,h}^{n+1}, \tag{3.56}$$

in (3.46)-(3.47). We can then determine $\mathbf{u}_{i,h}^{n+1}, p_{i,h}^{n+1}$ $(i = 1, 2)$ from

$$\left(\sigma_h^{n+1}\frac{\sigma_h^{n+1}\mathbf{u}_{1,h}^{n+1} - \sigma_h^n\mathbf{u}_h^n}{\delta t}, \mathbf{v}_h\right) + \left((\rho_h^{n+1}\mathbf{u}_h^n\cdot\nabla)\mathbf{u}_{1,h}^{n+1} + \frac{\mathbf{u}_{1,h}^{n+1}}{2}\nabla\cdot(\rho_h^{n+1}\mathbf{u}_h^n), \mathbf{v}_h\right) + \left(\nabla p_{1,h}^{n+1}, \mathbf{v}_h\right) + \nu\left(\nabla\mathbf{u}_{1,h}^{n+1}, \nabla\mathbf{v}_h\right) = \mathbf{0}, \tag{3.57}$$

$$\left(\sigma_h^{n+1}\frac{\sigma_h^{n+1}\mathbf{u}_{2,h}^{n+1}}{\delta t}, \mathbf{v}_h\right) + \left((\rho_h^{n+1}\mathbf{u}_h^n\cdot\nabla)\mathbf{u}_{2,h}^{n+1} + \frac{\mathbf{u}_{2,h}^{n+1}}{2}\nabla\cdot(\rho_h^{n+1}\mathbf{u}_h^n), \mathbf{v}_h\right) \tag{3.58}$$

$$+\left(\nabla p_{2,h}^{n+1},\mathbf{v}_h\right)+\nu\left(\nabla\mathbf{u}_{2,h}^{n+1},\nabla\mathbf{v}_h\right)-\frac{1}{\sqrt{E_1(\rho_h^{n+1})+C_0}}\left(\mathbf{f}_h^{n+1},\mathbf{v}_h\right)=\mathbf{0},$$

$$\left(\mathbf{u}_{i,h}^{n+1},\nabla q_h\right)=0. \tag{3.59}$$

The above system is a mixed formulation for $\mathbf{u}_{i,h}^{n+1},p_{i,h}^{n+1}$ $(i=1,2)$ whose wellposedness can be established in the same way as (3.46)–(3.48).

Once $\mathbf{u}_{i,h}^{n+1}(i=1,2)$ are known, we can plug (3.55) in (3.48) to determine explicitly η_h^{n+1} from

$$\left(1+\frac{\delta t}{2\sqrt{E_1(\rho^{n+1})+C_0}}\left(\mathbf{f}_h^{n+1},\mathbf{u}_{2,h}^{n+1}\right)\right)\eta_h^{n+1}=\eta_h^n-\frac{\delta t}{2\sqrt{E_1(\rho^{n+1})+C_0}}\left(\mathbf{f}_h^{n+1},\mathbf{u}_{1,h}^{n+1}\right). \tag{3.60}$$

We can replace the term $\frac{\delta t}{2\sqrt{E_1(\rho^{n+1})+C_0}}\left(\mathbf{f}_h^{n+1},\mathbf{u}_{2,h}^{n+1}\right)$ in the above by setting $\mathbf{v}_h=\mathbf{u}_{2,h}^{n+1}$ in (3.58) to obtain

$$\left(1+\frac{1}{2}\|\sigma_h^{n+1}\mathbf{u}_{2,h}^{n+1}\|^2+\frac{1}{2}\delta t\nu\|\nabla\mathbf{u}_2^{n+1}\|^2\right)\eta_h^{n+1}=\eta_h^n-\frac{\delta t}{2\sqrt{E_1(\rho^{n+1})+C_0}}\left(\mathbf{f}_h^{n+1},\mathbf{u}_{1,h}^{n+1}\right). \tag{3.61}$$

Hence, we can determine η_h^{n+1} uniquely from the above. Finally, we can obtain $\mathbf{u}_h^{n+1}$ and p_h^{n+1} from (3.55)-(3.56).

To summarize, given $\widetilde{\sigma}_h^n,\sigma_h^n,\mathbf{u}_h^n,\eta_h^n$, **Scheme** 1_h can be implemented as follows:

1. Solve $P_h^{\mathrm{RT}}\mathbf{u}_h^n$ from (3.53)-(3.54);
2. Solve $\widetilde{\sigma}_h^{n+1}$ from (3.44);
3. Compute ρ_h^{n+1} and σ_h^{n+1} from (3.45);
4. Solve $\mathbf{u}_{i,h}^{n+1},p_{i,h}^{n+1}(i=1,2)$ from (3.57)-(3.59);
5. Compute η_h^{n+1} from (3.61);
6. Update $\mathbf{u}_h^{n+1},p_h^{n+1}$ by (3.55)-(3.56).

Remark 3.2. For the scheme **Scheme** 1_h, setting $w_h=\widetilde{\sigma}_h^{n+1}$, we can obtain $\|\widetilde{\sigma}_h^{n+1}\|^2-\|\widetilde{\sigma}_h^n\|^2=\|\widetilde{\sigma}_h^{n+1}-\widetilde{\sigma}_h^n\|^2=O(\delta t^2)$, i.e. $\int_\Omega\rho_h^{n+1}\mathrm{d}x-\int_\Omega\rho_h^n\mathrm{d}x=O(\delta t^2)$, which indicates that the mass is only conserved up to second-order accuracy. Inspired by [26], we can obtain a numerical scheme that precisely satisfies mass conservation. Denote $\bar{\sigma}=\frac{1}{|\Omega|}\int_\Omega\widetilde{\sigma}\mathrm{d}x$. We can equivalently rewrite (2.19) as

$$\widetilde{\sigma}_t+\mathbf{u}\cdot\nabla\widetilde{\sigma}+\alpha(\nabla\cdot\mathbf{u})\left(\widetilde{\sigma}-\bar{\sigma}\right)=0,$$

where α is a constant. Therefore, we can construct a first-order fully discrete numerical scheme as follows:

$$\left(\frac{\widetilde{\sigma}_h^{n+1}-\widetilde{\sigma}_h^n}{\delta t},w_h\right)+\left(\mathbf{u}_h^n\cdot\nabla\widetilde{\sigma}_h^{n+1},w_h\right)+\alpha\left(\left(\nabla\cdot\mathbf{u}_h^n\right)\left(\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\right),w_h\right)=0. \tag{3.62}$$

By setting $w_h=1$ in (3.62) and noting that $\nabla\bar{\sigma}=0$, one can then obtain

$$\begin{aligned}\int_\Omega\widetilde{\sigma}_h^{n+1}\mathrm{d}x&=\int_\Omega\widetilde{\sigma}_h^n\mathrm{d}x-\delta t\left(\mathbf{u}_h^n,\nabla(\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1})\right)-\delta t\alpha\left(\nabla\cdot\mathbf{u}_h^n,\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\right)\\&=\int_\Omega\widetilde{\sigma}_h^n\mathrm{d}x+\delta t(1-\alpha)\left(\nabla\cdot\mathbf{u}_h^n,\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\right).\end{aligned}$$

Since $\int_\Omega\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\mathrm{d}x=0$, it follows that $\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\in\mathbb{Q}_h$ provided $k_1\le k_3$, where we recall that k_1,k_2,k_3 are the degrees of polynomials of $Q_{k_1}(K),Q_{k_2}(K)$, respectively. By the weak divergence-free condition, we have

$$\left(\nabla\cdot\mathbf{u}_h^n,\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\right)=0.$$

Therefore, if $\alpha=1$ or $k_1\le k_3$, the numerical scheme ensures conservation for $\widetilde{\sigma}_h^{n+1}$, i.e. $\int_\Omega\widetilde{\sigma}_h^{n+1}\mathrm{d}x=\int_\Omega\widetilde{\sigma}_h^n\mathrm{d}x$.

By denoting $\widetilde{\sigma}_h^{n+1/2}=\frac{\widetilde{\sigma}_h^{n+1}+\widetilde{\sigma}_h^n}{2}$, setting $w_h=\widetilde{\sigma}_h^{n+1/2}-\bar{\sigma}_h^{n+1/2}$ in (3.62), and using $(\boldsymbol{u}\cdot\nabla\rho,\rho)=-\frac{1}{2}((\nabla\cdot\boldsymbol{u})\rho,\rho)$, we obtain

$$\begin{aligned}\left(\widetilde{\sigma}_h^{n+1}-\widetilde{\sigma}_h^n,\widetilde{\sigma}_h^{n+1/2}-\bar{\sigma}_h^{n+1/2}\right)&=-\delta t\left(\mathbf{u}_h^n\cdot\nabla\widetilde{\sigma}_h^{n+1},\widetilde{\sigma}_h^{n+1/2}-\bar{\sigma}_h^{n+1/2}\right)-\delta t\alpha\left(\left(\nabla\cdot\mathbf{u}_h^n\right)\left(\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\right),\widetilde{\sigma}_h^{n+1/2}-\bar{\sigma}_h^{n+1/2}\right)\\&=\delta t\left(\frac{1}{2}-\alpha\right)\left(\left(\nabla\cdot\mathbf{u}_h^n\right)\left(\widetilde{\sigma}_h^{n+1}-\bar{\sigma}_h^{n+1}\right),\widetilde{\sigma}_h^{n+1/2}-\bar{\sigma}_h^{n+1/2}\right).\end{aligned}$$

If $\bar{\sigma}_h^{n+1/2}=\frac{1}{|\Omega|}\int_\Omega\widetilde{\sigma}_h^{n+1/2}\mathrm{d}x$, then

Table 1
The temporal discretization error and convergence rates for **scheme** 1_h at $T = 1$.

δt	$\|e_p\|_{L^2}$	order	$\|e_{\mathbf{u}}\|_{L^2}$	order	$\|e_p\|_{L^2}$	order
1/8	4.79E-02	-	7.21E-03	-	4.80E-02	-
1/16	2.40E-02	0.98	3.58E-03	1.01	2.46E-02	0.97
1/32	1.23E-02	0.99	1.78E-03	1.01	1.24E-02	0.98
1/64	6.16E-03	0.99	8.86E-04	1.00	6.25E-03	0.99
1/128	3.08E-03	1.00	4.42E-04	1.00	3.13E-03	1.00

$$\begin{aligned}\left(\widetilde{\sigma}_h^{n+1}-\widetilde{\sigma}_h^{n},\widetilde{\sigma}_h^{n+1/2}-\bar{\sigma}_h^{n+1/2}\right)&=\frac{1}{2}\|\widetilde{\sigma}_h^{n+1}\|^2-\left(\widetilde{\sigma}_h^{n+1},\bar{\sigma}_h^{n+1/2}\right)-\left(\frac{1}{2}\|\widetilde{\sigma}_h^{n}\|^2-\left(\widetilde{\sigma}_h^{n},\bar{\sigma}_h^{n+1/2}\right)\right)\\&=\frac{1}{2}\left(\|\widetilde{\sigma}_h^{n+1}\|^2-\left(\int_\Omega\widetilde{\sigma}_h^{n+1}\mathrm{d}x\right)^2\right)-\frac{1}{2}\left(\|\widetilde{\sigma}_h^{n}\|^2-\left(\int_\Omega\widetilde{\sigma}_h^{n}\mathrm{d}x\right)^2\right)\end{aligned}$$

In summary, if $\alpha=\frac{1}{2}$ and $k_1\le k_3$, we can obtain $\|\widetilde{\sigma}_h^{n+1}\|^2=\|\widetilde{\sigma}_h^{n}\|^2$, which implies mass conservation, i.e., $\int_\Omega\rho_h^{n+1}\mathrm{d}x=\int_\Omega\rho_h^{n}\mathrm{d}x$.

Although the numerical scheme (3.62) accurately satisfies mass conservation, it introduces an additional term, $\alpha(\nabla\cdot\mathbf{u})(\widetilde{\sigma}-\bar{\sigma})$, which increases the error and may lead to inaccuracies, particularly in rapidly changing flows. Additionally, this scheme constrains the polynomial degree of the finite element space for the $\widetilde{\sigma}$ to $k_1\le k_3$. To capture more detailed fluid dynamics by using a higher-order approximation space for the $\widetilde{\sigma}$, the polynomial degree of the finite element space of the pressure must be increased. This, in turn, raises the computational cost of the scheme. Balancing the need for higher accuracy and lower computational cost remains a significant challenge, especially in simulations involving complex and highly dynamic flow fields.

4. Numerical simulations

We start by presenting numerical examples to verify the convergence of the proposed **Scheme** 1_h and **Scheme** 2_h, followed by numerical simulations of some benchmark simulations, such as the air bubble rising in water and Rayleigh-Taylor instability, to demonstrate the robustness and energy stability of the schemes. These simulations are challenging due to the complex and dynamic nature of the fluid interfaces and the occurrence of discontinuities. We compare our simulation results with those obtained from other numerical methods to validate the accuracy and effectiveness of our proposed schemes.

4.1. Accuracy test

In order to test the numerical schemes constructed as above, we consider variable density Navier–Stokes equations with the exact solution

$$\begin{aligned}&\rho(x,y,t)=2+x\cos(\sin(t))+y\sin(\sin(t)),\\&\mathbf{u}(x,y,t)=\begin{pmatrix}-y\cos(t)\\x\cos(t)\end{pmatrix},\\&p(x,y,t)=\sin(x)\sin(y)\sin(t),\end{aligned}\tag{4.63}$$

in $\Omega=(-1,1)\times(-1,1)$, and the corresponding forcing term is

$$\mathbf{f}=\begin{pmatrix}\rho(x,y,t)\left(y\sin(t)-x\cos^2(t)\right)+\cos(x)\sin(y)\sin(t)\\-\rho(x,y,t)\left(x\sin(t)+y\cos^2(t)\right)+\sin(x)\cos(y)\sin(t)\end{pmatrix}.\tag{4.64}$$

Here we set $\nu=1$. Denote

$$e_\rho^{n+1}=\rho\left(t^{n+1}\right)-\rho_h^{n+1},\quad e_{\mathbf{u}}^{n+1}=\mathbf{u}\left(t^{n+1}\right)-\mathbf{u}_h^{n+1},\quad e_p^{n+1}=p\left(t^{n+1}\right)-p_h^{n+1}.$$

We use $\left(Q^2,Q^2,Q^1\right)$ discretization for the density, velocity, and pressure fields in our numerical simulations. We first use a mesh size of $h=1/128$ so that the spatial error is negligible compared with the time discretization error for the time steps considered. The L^2-norm errors at $T=1$ for **Scheme** 1_h and **Scheme** 2_h are presented in Table 1 and Table 2, respectively. To compute the convergence rate in space, we set $\delta t=\frac{1}{8}h^{3/2}$ so that the time discretization error is negligible compared with spatial discretization error. We observe in Table 3 that the expected convergence rates are achieved for all cases, confirming the accuracy of the proposed schemes.

In Fig. 1, we present the mass error and convergence rates for **scheme** 1_h and **scheme** 2_h at $T=1$ on the left subplot, and the evolution of mass error for both schemes with a time step of $\delta t=1E-4$ on the right subplot. This demonstrates that, although exact mass conservation cannot be guaranteed, both **scheme** 1_h and **scheme** 2_h are capable of preserving mass conservation to the first and second order, respectively. Moreover, with a time step of $\delta t=1E-4$, the mass conservation accuracy for the first-order scheme can achieve a precision of $1E-5$, whereas for the second-order scheme, the precision reaches $1E-10$. This signifies a substantial degree of adherence to mass conservation principles by both schemes, underlining their capacity to uphold mass conservation with distinct levels of precision corresponding to their respective orders.

Table 2
The temporal discretization error and convergence rates for **scheme** 2_h at $T = 1$.

δt	$\|e_\rho\|_{L^2}$	order	$\|e_{\mathbf{u}}\|_{L^2}$	order	$\|e_p\|_{L^2}$	order
1/8	6.26E-03	-	6.48E-04	-	6.09E-03	-
1/16	1.77E-03	1.82	1.72E-04	1.92	1.47E-03	2.05
1/32	4.68E-04	1.92	4.40E-05	1.96	3.59E-04	2.03
1/64	1.20E-04	1.96	1.11E-05	1.98	8.90E-05	2.01
1/128	3.04E-05	1.98	2.80E-06	1.99	2.22E-05	2.00

Table 3
The spatial discretization error and convergence rates for **scheme** 2_h with $\delta t = \frac{1}{8}h^{3/2}$ at $T = 0.1$.

h	$\|e_\rho\|_{L^2}$	order	$\|e_{\mathbf{u}}\|_{L^2}$	order	$\|e_p\|_{L^2}$	order
1/16	4.57E-06	-	1.10E-07	-	5.55E-05	-
1/32	5.66E-07	3.01	9.93E-09	3.48	1.18E-05	2.23
1/64	7.07E-08	3.00	1.16E-09	3.10	2.84E-06	2.06
1/128	8.84E-09	3.00	1.45E-10	3.00	7.03E-07	2.01
1/256	1.11E-09	3.00	1.83E-11	2.98	1.75E-07	2.00

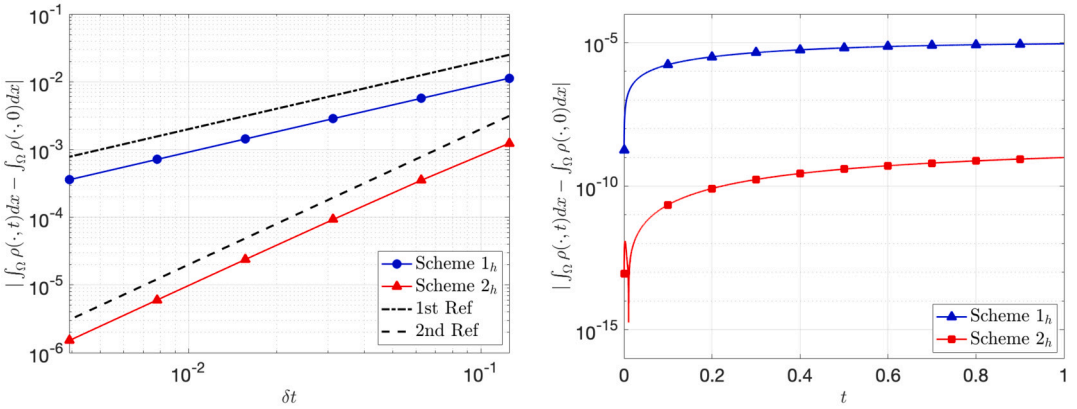

Fig. 1. Left: mass error and convergence rates for **scheme** 1_h and **scheme** 2_h at $T = 1$. Right: the evolution of mass error for **scheme** 1_h and **scheme** 2_h with $\delta t = 1E-4$.

Table 4
Physical parameters for rising bubble test.

Parameter	Air	Water	Unit (MKS)
Density (ρ)	1.161	995.65	kg/m^3
Viscosity (ν)	0.0000186	0.0007977	kg/ms

4.2. Rising bubbles

In this subsection, we conduct a simulation of a bubble rising in water to validate the performance of our numerical scheme. This problem has been extensively studied by many researchers in a two-dimensional rectangular domain. The physical parameters used in our simulation are listed in Table 4, and they are the same as those used in [28].

We use the finite element space of (Q^2, Q^2, Q^1) for $(\rho, \mathbf{u}, p)$ in the simulation. Since air and water have different viscosities, we replace the viscous term $-\nu\Delta\mathbf{u}$ with $-\nabla\cdot(\nu(\rho)\nabla\mathbf{u})$ where

$$\nu(\rho) = \frac{\rho - \frac{\rho_{air}+\rho_{water}}{2}}{\rho_{air}-\rho_{water}}\left(\nu_{air}-\nu_{water}\right) + \frac{\nu_{air}+\nu_{water}}{2}. \tag{4.65}$$

To approximate the initial discontinuous density at the air-water interface, we use the following initial discontinuous density

$$\rho(\mathbf{x}, t=0) = \rho_{\text{air}} + \left(\frac{\rho_{\text{water}}-\rho_{\text{air}}}{2}\right)\times\left(1+\tanh\left(\frac{c-0.0025}{0.00025}\right)\right), \tag{4.66}$$

where $c = \sqrt{x_1^2 + (x_2 - 0.0075)^2}$ is the distance from the center of the bubble to a point.

The initial velocity is set to be zero, and the computational domain is $[-0.01m, 0.01m]\times[0, 0.03m]$. Initially, an air bubble with a radius of $0.0025m$ is placed in the lower middle of the domain filled with water. In the experiment, the gravitational term is defined as $\mathbf{f} = (0, \rho g)^T$, where $g = -9.80665m/s^2$ represents the acceleration due to gravity.

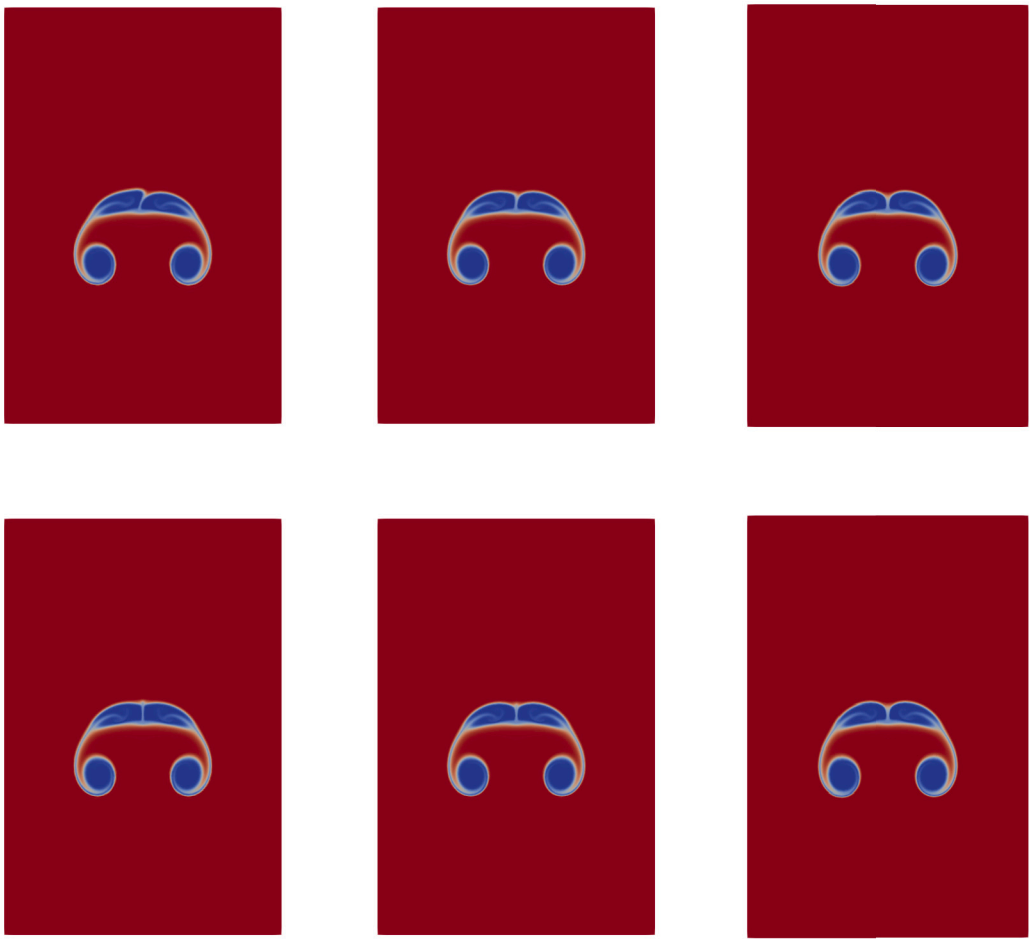

Fig. 2. First row: Numerical solution of air bubble rising in water at $t = 0.08$ obtained by triangle mesh. Second row: numerical solution of air bubble rising in water at $t = 0.08$ obtained by rectangle mesh. Time step: $\delta t = 1E-4$; spatial mesh sizes: first column: $h = 0.01/80$, second column: $h = 0.01/100$, third column: $h = 0.01/150$; CFL-like number: first column: $\delta t/h = 0.8$, second column: $\delta t/h = 1$, third column: $\delta t/h = 1.5$.

First, we use **Scheme** 1_h with a fixed time step of $\delta t = 1E-4$. The results obtained with triangular and rectangular meshes for different spatial mesh sizes are presented in Fig. 2. From the results, several observations can be made: (i) When the spatial mesh size is not sufficiently small, the results obtained with the triangular mesh are asymmetric, as shown in the first two subplots displayed in the first row of Fig. 2; the results obtained with the rectangular mesh are always symmetric, even when the spatial mesh size is not small. (ii) With a fixed time step, decreasing the spatial mesh size leads to a more pronounced breaking of the connection between the left and right sides of the upper part of the bubble. This phenomenon may be related to a CFL-like (Courant-Friedrichs-Lewy-like) number that is not sufficiently small, as can be more clearly observed in the second row of Fig. 2, where here the CFL-like number refers to the ratio of the time step size to the spatial step size, i.e. $\delta t/h$. Overall, the numerical simulations suggest that the rectangular mesh is a more robust choice for solving the variable-density Navier-Stokes equations, as it yields more symmetric and stable results compared to the triangular mesh. Furthermore, the results indicate that the choice of CFL-like number is crucial, as demonstrated in Fig. 3, which depicts the results obtained with different CFL-like numbers: $3.2, 0.8, 0.75, 0.256$, respectively. It can be observed that a smaller CFL-like number results in a clearer interface between the air bubble and water. Moreover, comparing the first subplot to the second subplot of Fig. 3 and the third subplot of second row of Fig. 2 to the third subplot of Fig. 3, it is evident that a smaller CFL-like number leads to a narrower gap between the left and right sides of the upper part of the bubble. These observations highlight the importance of choosing an appropriate CFL-like number to ensure accurate and stable simulations of variable-density Navier-Stokes equations.

The evolution of modified energy and $\xi^{n+1} = \frac{\eta^{n+1}}{\sqrt{E_1(\rho_h^{n+1})+C_0}}$ obtained using different time step sizes in the first row of Fig. 4. It can be observed that the energy decay becomes slower as the time step size decreases. Additionally, for smaller time step size, ξ^{n+1} approaches 1, indicating a more accurate approximation of numerical solution. The modified energy and original energy evolution with $h = 0.01/150$ and $\delta t = 2.5E-4$ are shown in the left subplot of second row of Fig. 4. It is clear that in the initial stage, the modified energy decreases while a slight increase is observed in the original energy. Additionally, the right subplot of the second row in Fig. 4 displays the errors between the modified energy and original energy at $t = 0.005$. These errors, being of first-order, suggest that the differences between the modified energy and original energy are insignificant.

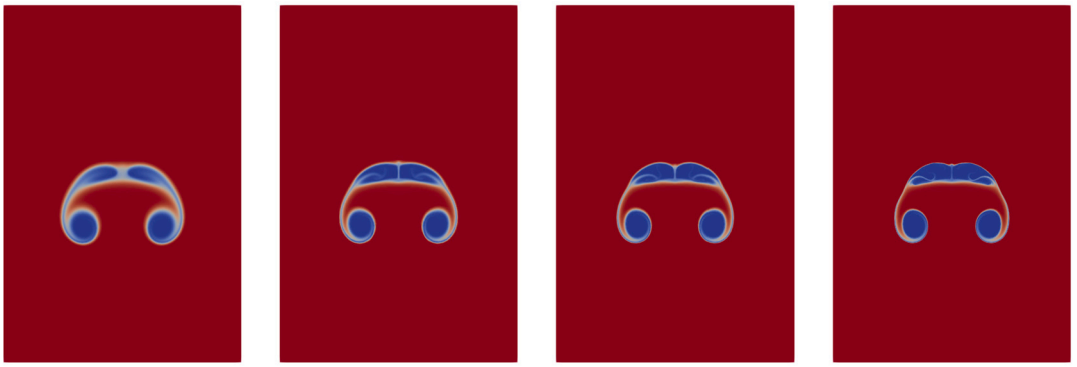

Fig. 3. Numerical solution of air bubble rising in water at $t = 0.08$. First: $\delta t = 4E-4$, $h = 0.01/80$, $\delta t/h = 3.2$; second: $\delta t = 1E-4$, $h = 0.01/80$, $\delta t/h = 0.8$; third: $\delta t = 5E-5$, $h = 0.01/150$, $\delta t/h = 0.75$; fourth: $\delta t = 1E-5$, $h = 0.01/256$, $\delta t/h = 0.256$.

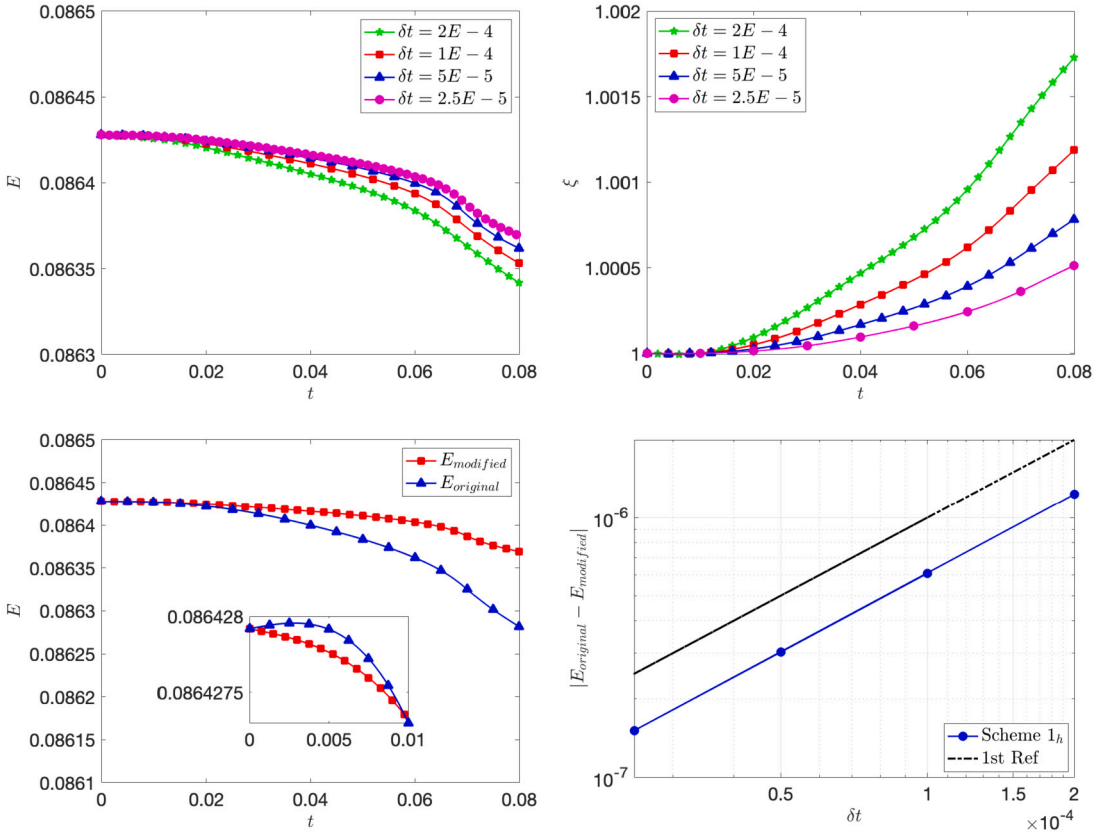

Fig. 4. Rising bubble. First row: the evolution of modified energy (left) and ξ (right) with $h = 0.01/150$. Second row: the evolution of modified energy and original energy with $h = 0.01/150, \delta t = 2.5E-5$ (left); the errors of modified energy and original energy at $t = 0.005$ with $h = 0.01/150$ (right).

Next, we elaborate on the necessity of using certain techniques in **Scheme** 1_h. The spatial discretization of density, when conducted without the implementation of upwind schemes, is represented by the following equation:

$$\left(\frac{\widetilde{\sigma}_h^{n+1} - \widetilde{\sigma}_h^n}{\delta t}, w_h\right) + \left(\sum_e \left(P_h^{\mathrm{RT}} \mathbf{u}_h^n, [\![\widetilde{\sigma}_h^{n+1} w_h]\!]\right)_e - \left(\widetilde{\sigma}_h^{n+1} P_h^{\mathrm{RT}} \mathbf{u}_h^n, \nabla w_h\right)\right) = 0. \tag{4.67}$$

The outcomes of this approach, as illustrated in Fig. 5, indicate that oscillations emerge at the interface by $t = 0.01$. These oscillations worsen by $t = 0.013$, resulting in the inaccurate deformation of the air bubble. This highlights the importance of using upwind schemes. The results of the velocity field for the transport equation without using Raviart-Thomas projection are shown in Fig. 6, where oscillations are observed at the interface of the lower half of the air bubble at $t = 0.03$, and the oscillations become more severe over time. This emphasizes the significance of using Raviart-Thomas projection for the velocity field. In particular, if we use a usual (without bound preserving) scheme

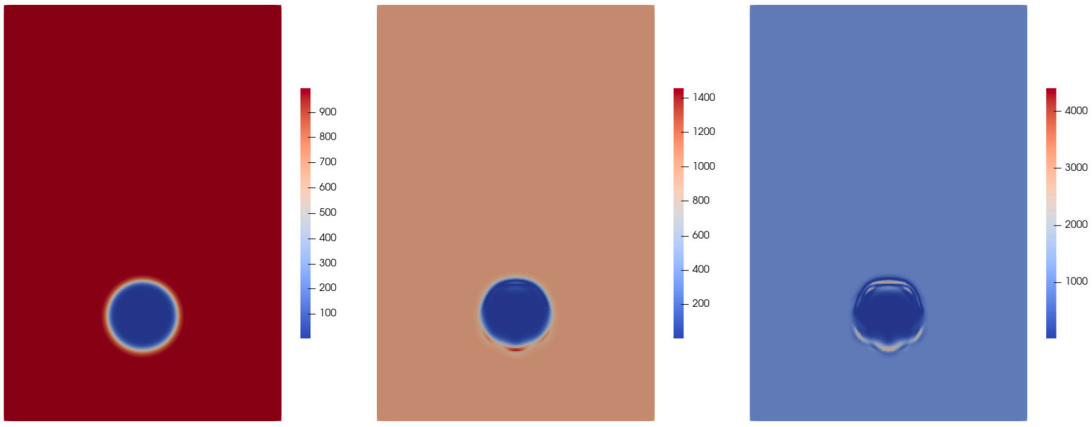

Fig. 5. Numerical solution of air bubble rising in water at $t = 0.005, 0.01, 0.013$ with spatial mesh of $0.01/150$ and time step of $\delta t = 5E-5$ but without upwind.

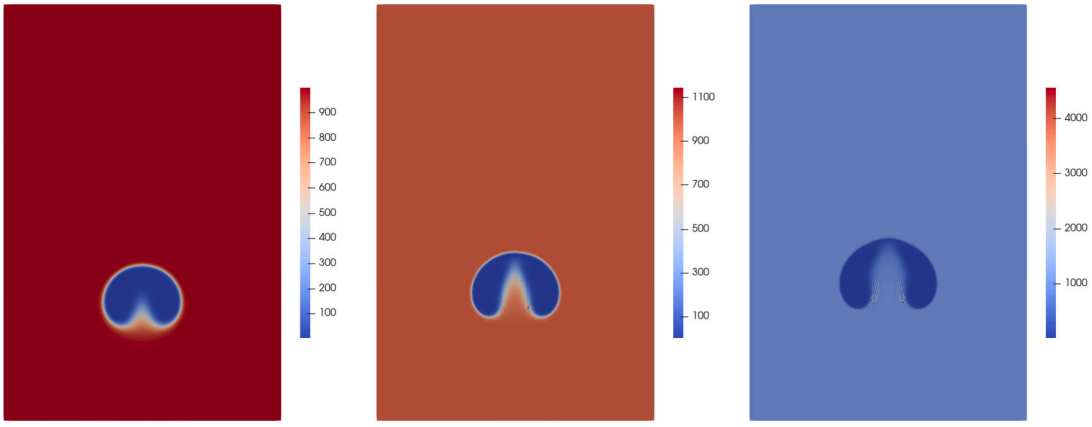

Fig. 6. Numerical solution of air bubble rising in water at $t = 0.02, 0.03, 0.04$ with spatial mesh size of $0.01/150$ and time step of $\delta t = 5E-5$ but without projecting velocity fields to Raviart-Thomas element space in transport equation.

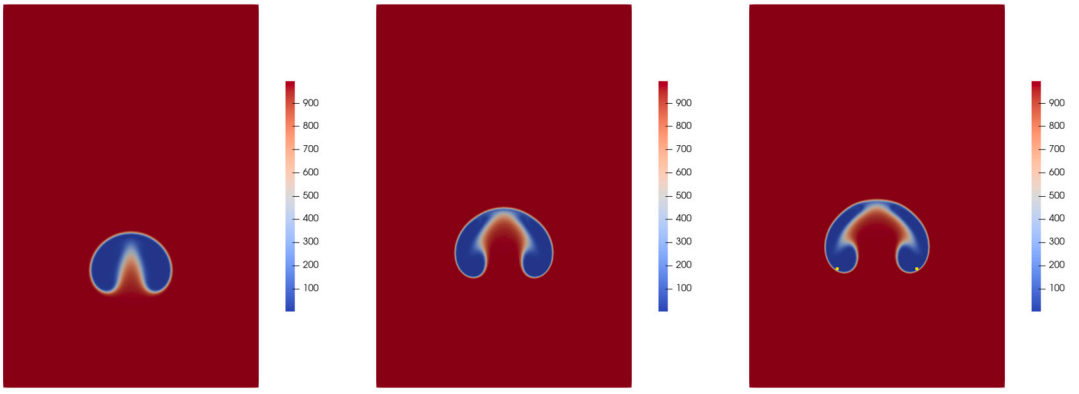

Fig. 7. Numerical solution of air bubble rising in water at $t = 0.03, 0.05, 0.057$ with spatial mesh size of $0.01/150$ and time step of $\delta t = 5E-5$ but without preserving the lower bound of density. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

$$\left(\frac{\rho_h^{n+1}-\rho_h^n}{\delta t}, w_h\right) + \left(\sum_e \left(\widehat{\rho}_h^{n+1} P_h^{\mathrm{RT}} \mathbf{u}_h^n, [\![w_h]\!]\right)_e - \left(\rho_h^{n+1} P_h^{\mathrm{RT}} \mathbf{u}_h^n, \nabla w_h\right)\right) = 0, \tag{4.68}$$

instead of scheme (3.44)-(3.45), the lower bound of density can not be preserved as demonstrated in Fig. 7. At $t = 0.057$, the density field exhibits negative values with a minimum of $\rho_{min} = -0.0158$, as indicated by the yellow data point in the third subplot, and consequently the simulation breaks down.

In Fig. 8, we present snapshots of the air bubble at nine different times ranging from 0 to 0.8 s. These results are in agreement with those reported in [28]. However, minor discrepancies between our results and theirs may be attributed to the fact that their

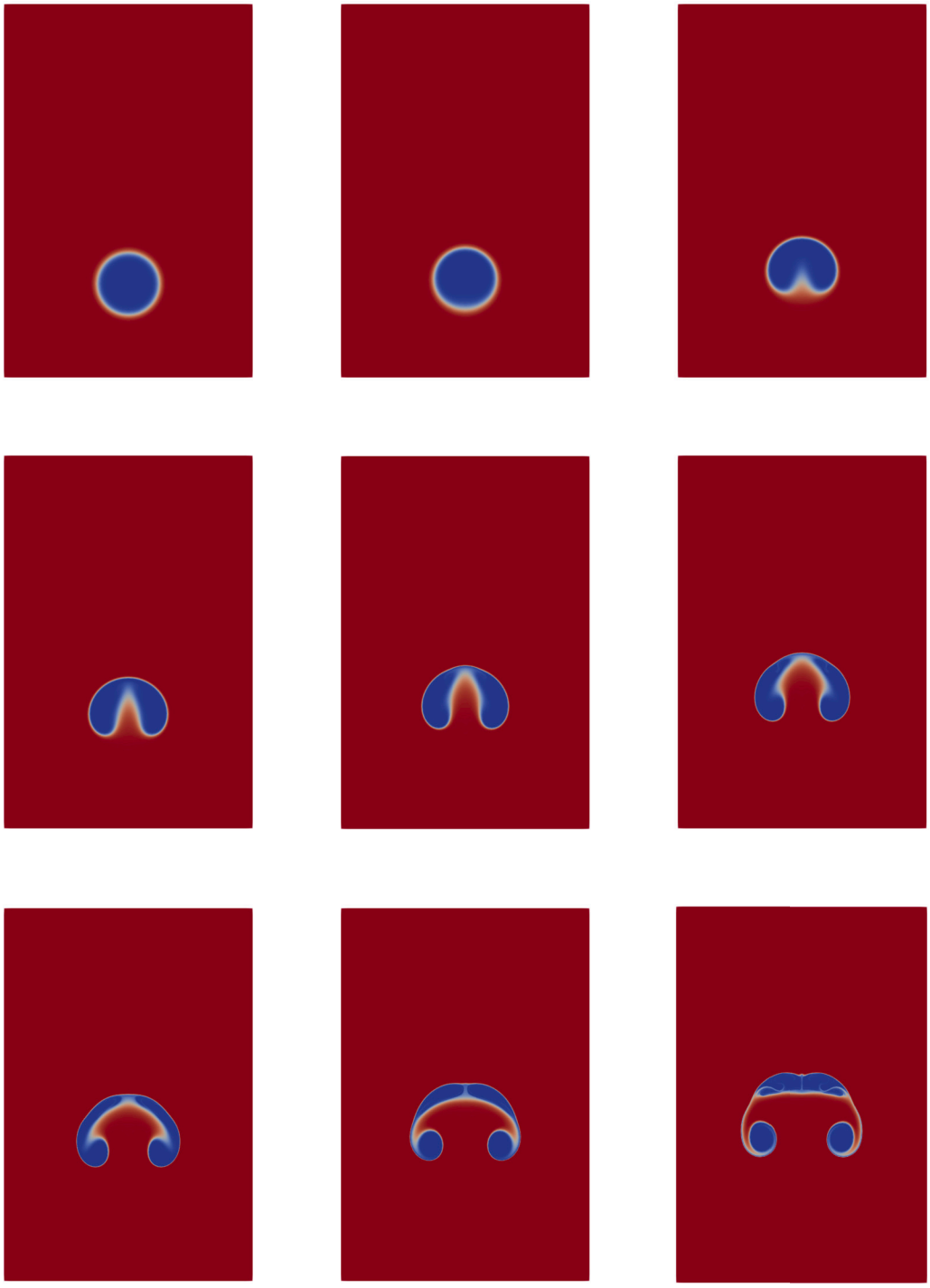

Fig. 8. Numerical solution of air bubble rising in water at $t = 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08$. $\delta t = 1E-5$, $h = 0.01/256$, $\delta t/h = 0.256$.

computations employed a larger CFL-like number, which could result in a larger gap between the left and right sides of the upper part of the bubble.

We also performed simulation for the same problem using **Scheme** 2_h, the results are shown in Fig. 9. It can be observed that, compared to the first-order scheme, a much larger time step can be used to make the gaps on both sides of the bubble's upper part almost invisible.

Simulating the rising bubble in water is a widely studied problem in fluid dynamics, and it serves as an important benchmark for validating the accuracy and stability of numerical methods. By comparing our results with those obtained from other numerical methods, we can confirm the effectiveness and reliability of the proposed numerical schemes.

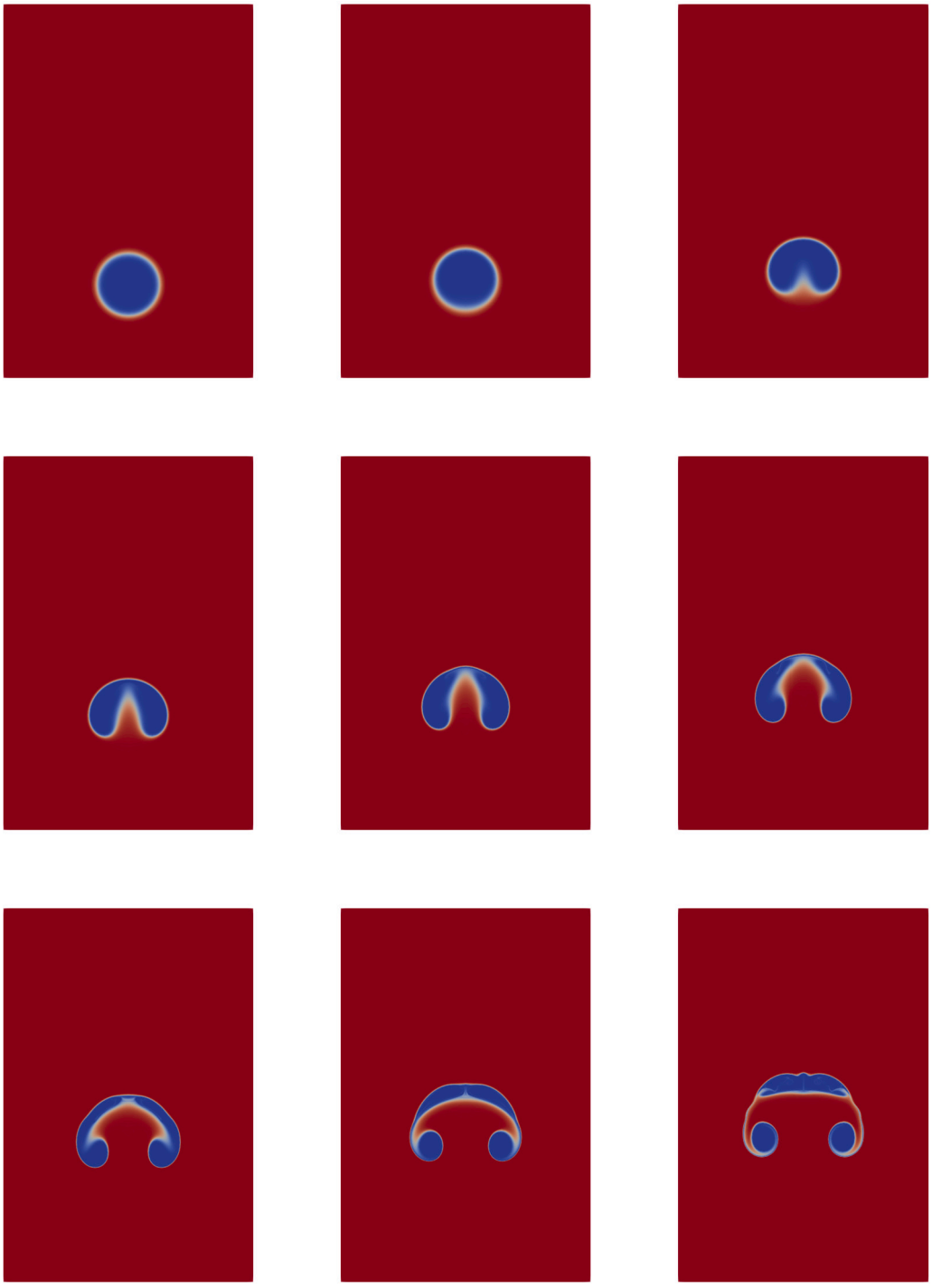

Fig. 9. Numerical solution of air bubble rising in water at $t = 0, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08$ with $\delta t = 1.25E-4$, $h = 0.01/256$ and $\delta t/h = 3.2$.

4.3. Rayleigh-Taylor instability

We further demonstrate the method's performance on another benchmark problem, specifically the development of a Rayleigh-Taylor instability in the viscous regime. This problem involves two layers of fluid initially at rest in the rectangular domain $\Omega = (-1/2, 1/2) \times (-2, 2)$, where the transition between the two fluids is regularized using the following equation:

$$\rho(x, y, t = 0) = \frac{\rho_m + \rho_M}{2} + \frac{\rho_M - \rho_m}{2} \tanh\left(\frac{y - s(x)}{0.01}\right), \tag{4.69}$$

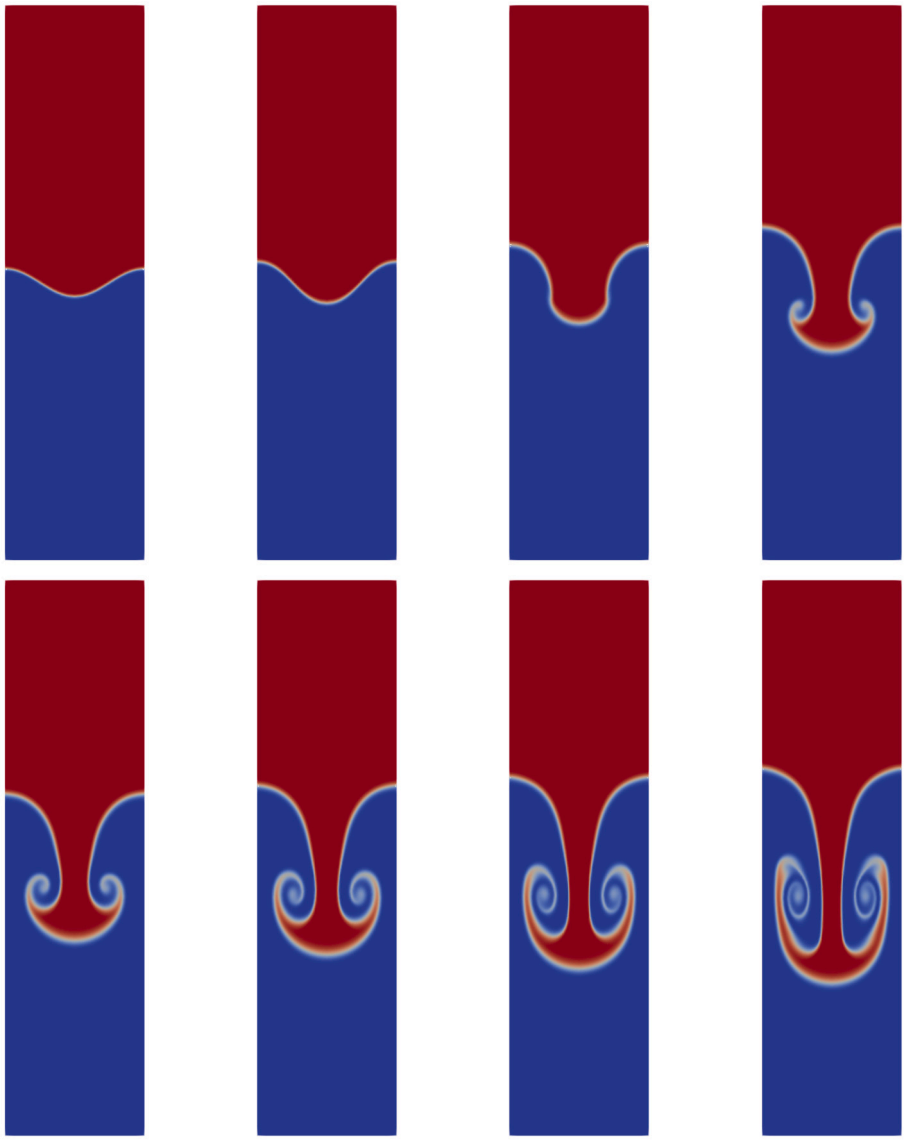

Fig. 10. Numerical solution of Rayleigh-Taylor instability with $A_t = 0.5, \nu = 0.001$ at $t = 0$, 0.2, 0.4, 0.6, 0.7, 0.8, 0.9, 1. Time step: $\delta t = 0.01$, spatial mesh: $h = 1/60$, $\delta t/h = 0.6$.

here, $s(x) = -0.1\cos(2\pi x)$ represents the initial position of the perturbed interface. With the heavy fluid above, the density of the two fluids is 3 and 1, respectively, resulting in an Atwood number of 0.5 according to Tryggvason's definition [32]:

$$A_t = \frac{\rho_M^0 - \rho_m^0}{\rho_M^0 + \rho_m^0}, \tag{4.70}$$

where $\rho_M^0 := \max_{\mathbf{x}\in\Omega} \rho_0(\mathbf{x})$. The no-slip condition is enforced at the top and bottom walls, while periodic condition is imposed on the two vertical sides. In the experiment, the force due to gravity is represented by $\mathbf{f} = (0, \rho g)^T$, with $g = -9.80665 m/s^2$, reflecting the standard acceleration due to Earth's gravity. **Scheme** 1_h is used in the simulation below.

We perform simulations with a Reynolds number of $\nu = 0.001$. The results obtained with different time and space step sizes are shown in Fig. 10, Fig. 11 and Fig. 12. In Fig. 10, we show the simulation results obtained with a time step size of 0.01 and a spatial resolution of $1/60$. It can be seen that the simulated interface between the two fluids is relatively smooth, and the evolution of the instability is captured with reasonable accuracy. In Fig. 11 and Fig. 12, we reduce the time step size and the

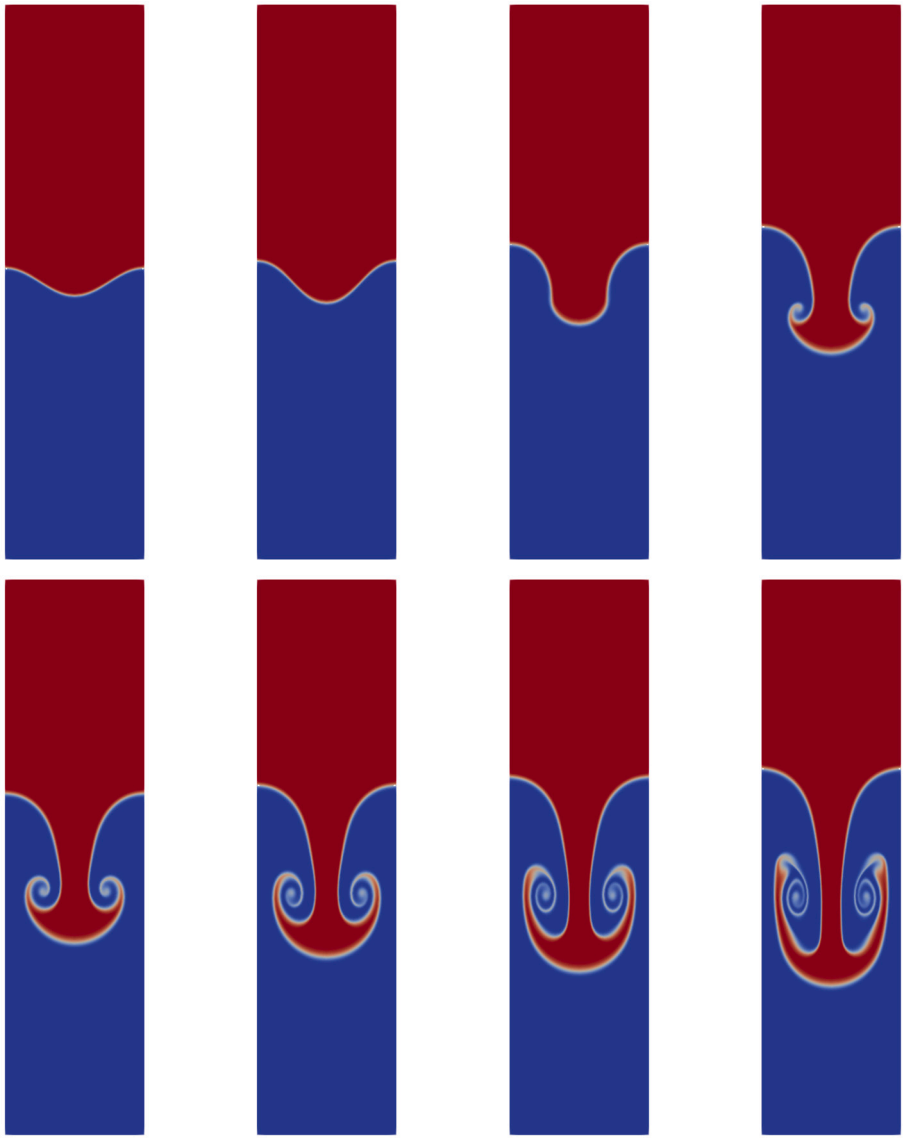

Fig. 11. Numerical solution of Rayleigh-Taylor instability with $A_t = 0.5, \nu = 0.001$ at $t = 0$, 0.2, 0.4, 0.6, 0.7, 0.8, 0.9, 1. Time step: $\delta t = 0.005$, spatial mesh: $h = 1/100$, $\delta t/h = 0.5$.

spatial mesh size. These figures reveal that as the time and space step sizes become smaller, the interface between the two fluids becomes clearer, and more details of the evolution can be observed. This indicates that the proposed method is capable of accurately resolving the complex dynamics of the Rayleigh-Taylor instability in the viscous regime, and the results obtained are reliable.

The evolution of the modified energy and ξ^{n+1} under different time and space mesh sizes is shown in first row of Fig. 13. It can be observed that as the time step size becomes smaller, the energy decay becomes slower and ξ^{n+1} approaches 1 more closely. The observed trends in the modified energy and ξ^{n+1} support the conclusion that decreasing the time step size can lead to a more accurate and detailed representation of the flow field. In the left subplot of the second row of Fig. 13, the evolution of modified energy and original energy with $h = 1/200$ and $\delta t = 0.001$ is illustrated. It is evident that during the initial stage, the modified energy decreases while the original energy exhibits a slight increase. Furthermore, the errors of modified energy and original energy at $t = 0.1$ are presented in the right subplot of the second row of Fig. 13. The first-order error indicates that the discrepancies between the modified energy and original energy are minor.

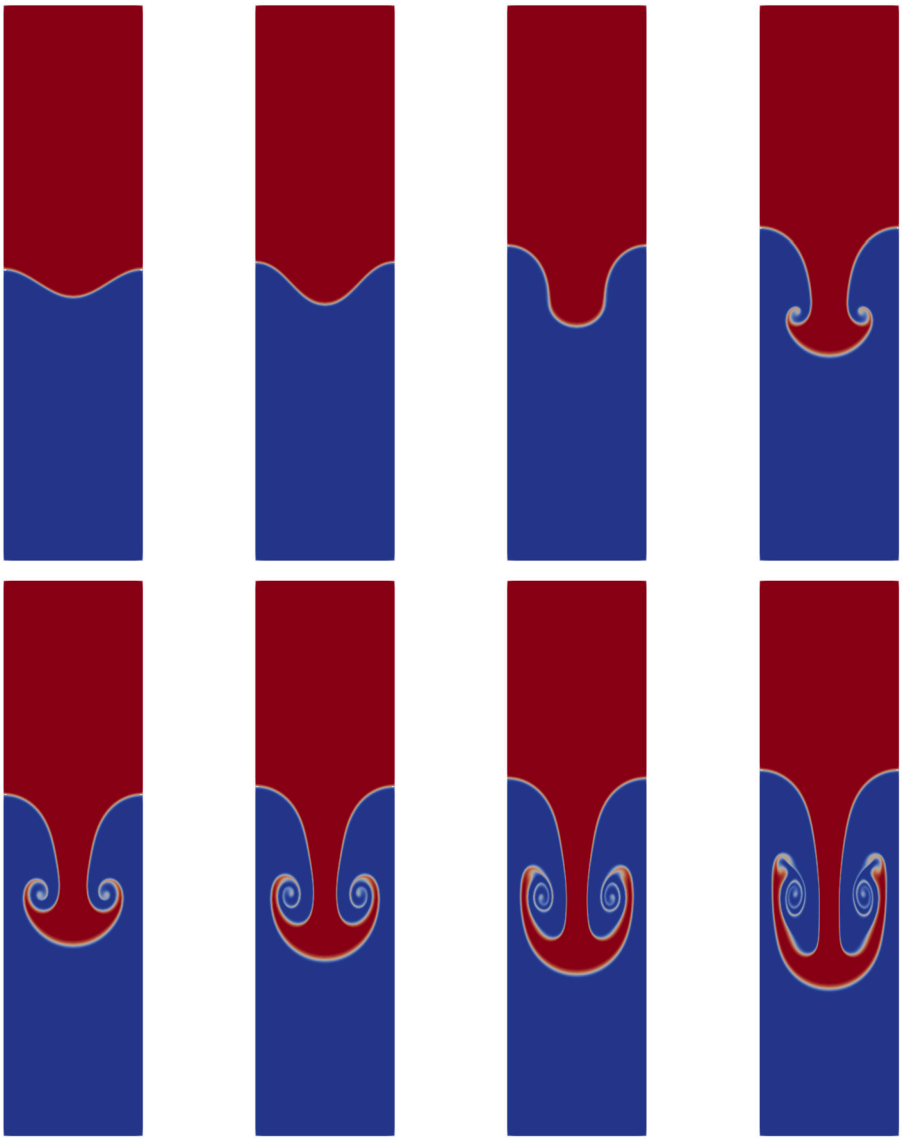

Fig. 12. Numerical solution of Rayleigh-Taylor instability with $A_t = 0.5, \nu = 0.001$ at $t = 0, 0.2, 0.4, 0.6, 0.7, 0.8, 0.9, 1$. Time step: $\delta t = 0.0025$, spatial mesh: $h = 1/200$, $\delta t/h = 0.5$.

To further assess the robustness of the algorithm at high Reynolds numbers, we solved the same problem using $\nu = 2E-4$. The results are presented in Fig. 14. By comparing the results in Fig. 12 and Fig. 14, it can be observed that as the Reynolds number increases, more details of the vortices become visible.

Fig. 15 depicts the time evolution of the density of a fluid with larger Atwood number 0.75, where the densities of the two fluids are 7 and 1. Based on the results presented in Fig. 12 and Fig. 15, we can observe similarities in the flow structure and global characteristics. Additionally, we have observed that the heavy fluid falls faster when the low Atwood number becomes larger.

To illustrate the capability of our model to describe the evolution of multiple fluid densities, we employ the initial density distribution defined by:

$$\rho(x, y, t = 0) = \frac{1}{2}\tanh\left(\frac{y - 0.5 - s(x)}{0.01}\right) + \frac{1}{2}\tanh\left(\frac{y + 0.5 - s(x)}{0.01}\right) + 2,$$

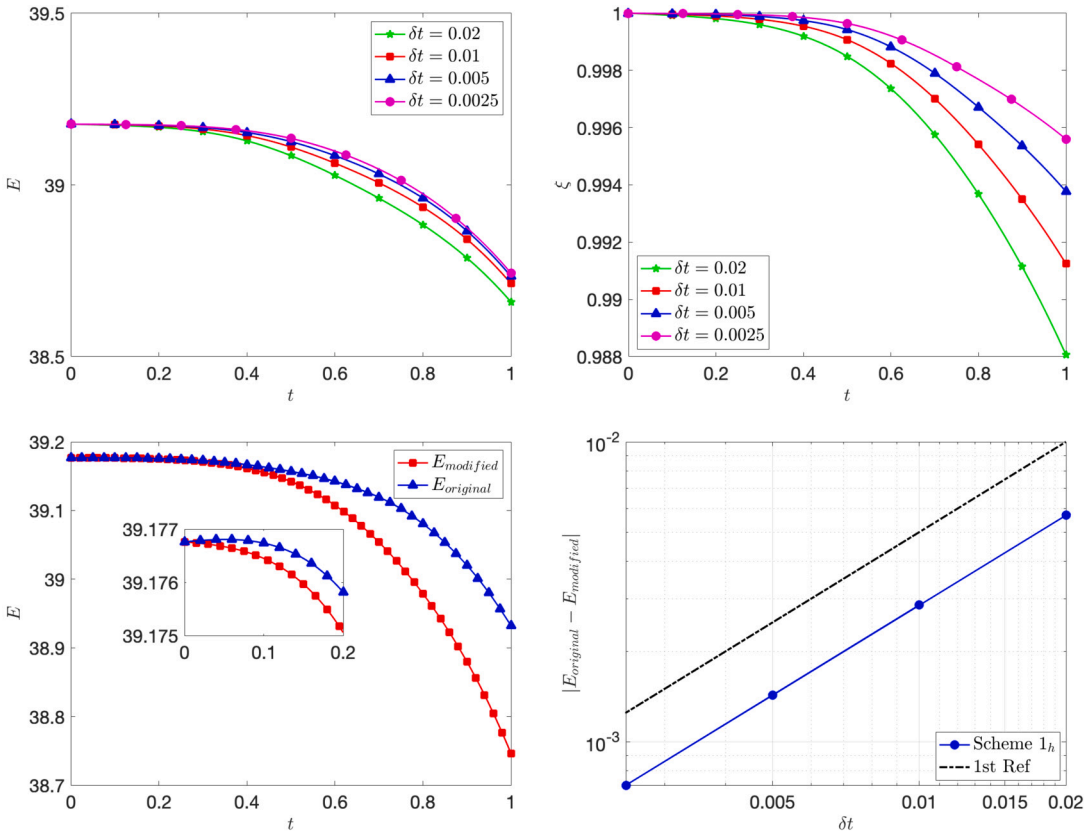

Fig. 13. Rayleigh-Taylor instability. First row: the evolution of modified energy (left) and ξ (right) with $h = 1/200$. Second row: the evolution of modified energy and original energy with $h = 1/200, \delta t = 0.001$ (left); the errors of modified energy and original energy at $t = 0.1$ with $h = 1/200$ (right).

where the fluid densities from top to bottom are set at 3, 2, and 1, respectively. The outcomes are summarized in Fig. 16. We performed the simulation using various time step sizes and spatial mesh sizes, and although not displayed here due to space constraints, it can be observed that smaller time and spatial steps yield more accurate fluid interfaces, confirming the reliability of our numerical results. Throughout the evolution, several phenomena can be observed: (i) As time progresses, the heavier fluid descends into the lighter fluid, and the interactions between fluids of differing densities lead to increasingly complex vortex structures; (ii) The fluid with a density of 2 forms vortices more rapidly than the fluid with a density of 3; (iii) After a certain point, the shape of the density 3 fluid begins to be influenced by the shape of the density 2 fluid. These findings highlight the scheme's ability in accurately capturing the subtle dynamics of fluid interfaces under varying density conditions.

4.4. Falling bubble

This numerical example investigates the evolution process of a heavy droplet passing through a light fluid and impacting the planar surface of a denser fluid. The equations are rendered dimensionless using the following reference values: lower bound ρ_m for density ρ, d for length, and $\sqrt{d/g}$ for time, which sets the reference velocity as $\sqrt{dg}$. In these dimensionless equations, the Reynolds number is defined by $Re = \frac{\rho_m d^{3/2} g^{1/2}}{\nu}$, and the term for gravity is represented as $\frac{\mathbf{f}}{\rho_m g}$, i.e. $\mathbf{f} = (0, -\frac{\rho}{\rho_m})^T$. The computational domain is $\Omega = [0, d] \times [0, 2d]$ with $d = 1$, and the initial density interface takes the form

$$\rho(x, y, t = 0) = \frac{99}{2} \tanh(\frac{0.2 - r}{0.01}) + \frac{99}{2} \tanh(\frac{1 - y}{0.01}) + 100,$$

where $r = \sqrt{(x - 0.5)^2 + (y - 1.75)^2}$.

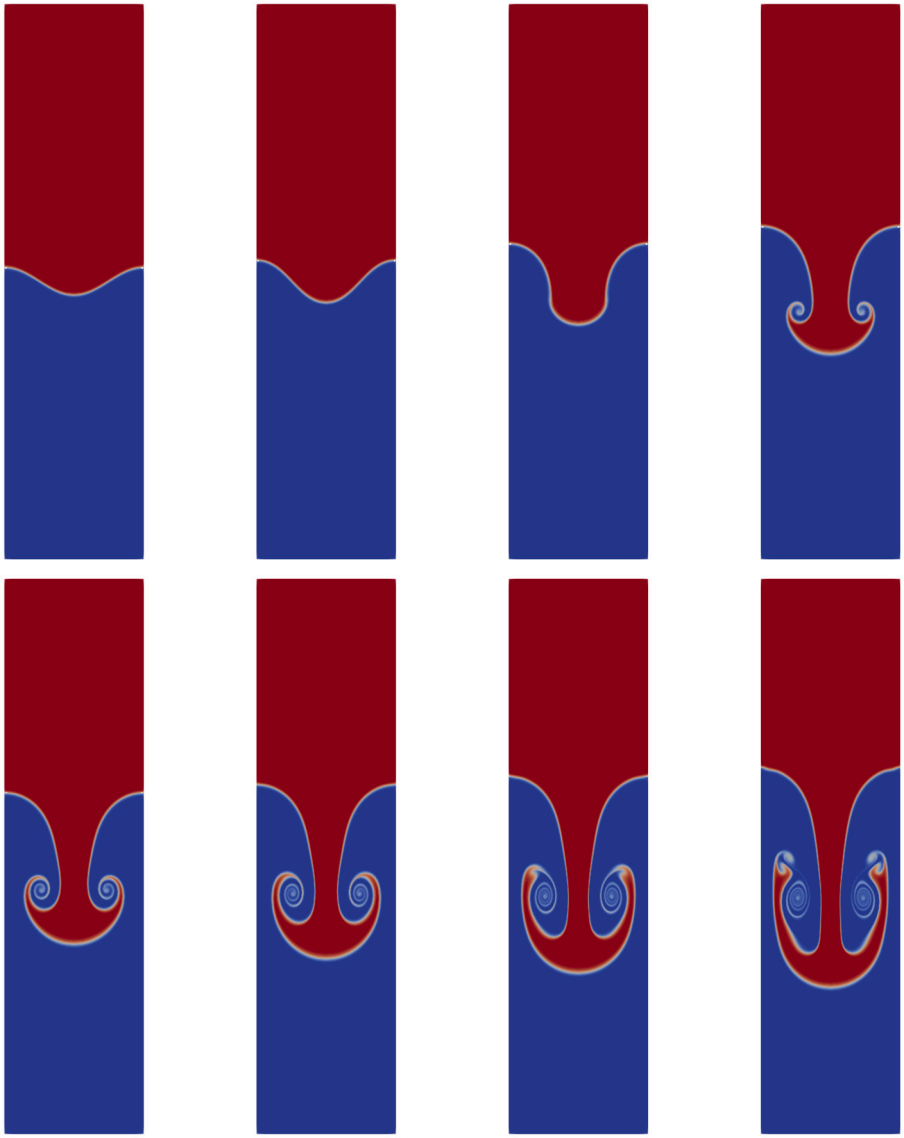

Fig. 14. Numerical solution of Rayleigh-Taylor instability with $A_t = 0.5, \nu = 2E-4$ at $t = 0, 0.2, 0.4, 0.6, 0.7, 0.8, 0.9, 1$. Time step: $\delta t = 0.0025$, spatial mesh: $h = 1/200$, $\delta t/h = 0.5$.

In the experiment, the fluid's viscosity is assumed to be constant throughout the domain, with a Reynolds number of $Re = 200$. The computational domain is discretized into 100×200 cells. The outcomes are depicted in Fig. 17. These results are similar to those in [18]. The visualizations illustrate the droplet descending through a lighter fluid and coalescing with the fluid beneath. During descent, the droplet undergoes minor shape alterations due to the absence of surface tension. Upon impact with the interface, it merges with the denser fluid below, inducing surface waves.

5. Conclusion

We have proposed first- and second-order semi and fully discrete schemes for the Navier-Stokes equations with variable density. These schemes preserve lower density bounds without altering the structure of the transport equation or requiring additional operations. For spatial discretization, we have proposed to use an upwind discontinuous Galerkin method for the density, and continuous finite element spaces for the velocity and pressure. By projecting the velocity in the transport equation to Raviart-Thomas space, the proposed schemes eliminate the requirement of using divergence-free finite elements for the velocity equations, resulting in lower

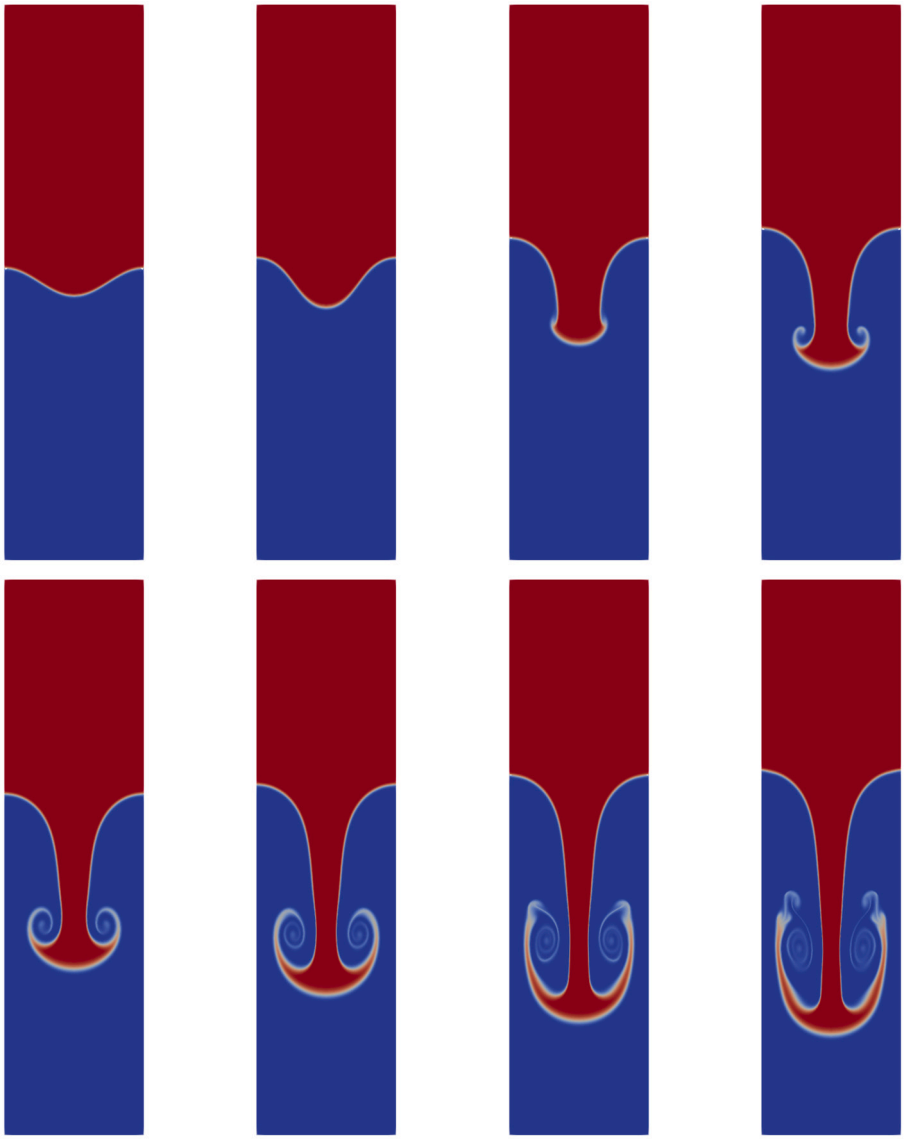

Fig. 15. Numerical solution of Rayleigh-Taylor instability with $A_t = 0.75, \nu = 0.001$ at $t = 0$, 0.2, 0.4, 0.5, 0.6, 0.7, 0.8, 0.85. Time step: $\delta t = 0.0025$, spatial mesh: $h = 1/200$, $\delta t/h = 0.5$.

computational cost and easier implementation. We have established energy stability of these semi and fully discrete schemes in the presence of a gravity term.

To demonstrate the accuracy and efficiency of our proposed schemes, we have conducted convergence tests as well as the simulation of three benchmark problems: an air bubble rising in water, the evolution of Rayleigh-Taylor instability and falling bubble. The proposed schemes demonstrate excellent performance in accurately capturing the complex dynamics of these problems.

CRediT authorship contribution statement

Buyang Li: Writing – review & editing, Methodology, Conceptualization. **Jie Shen:** Writing – review & editing, Methodology, Conceptualization. **Zongze Yang:** Writing – review & editing, Methodology, Conceptualization. **Yanrong Zhang:** Writing – original draft, Writing – review & editing, Methodology, Conceptualization.

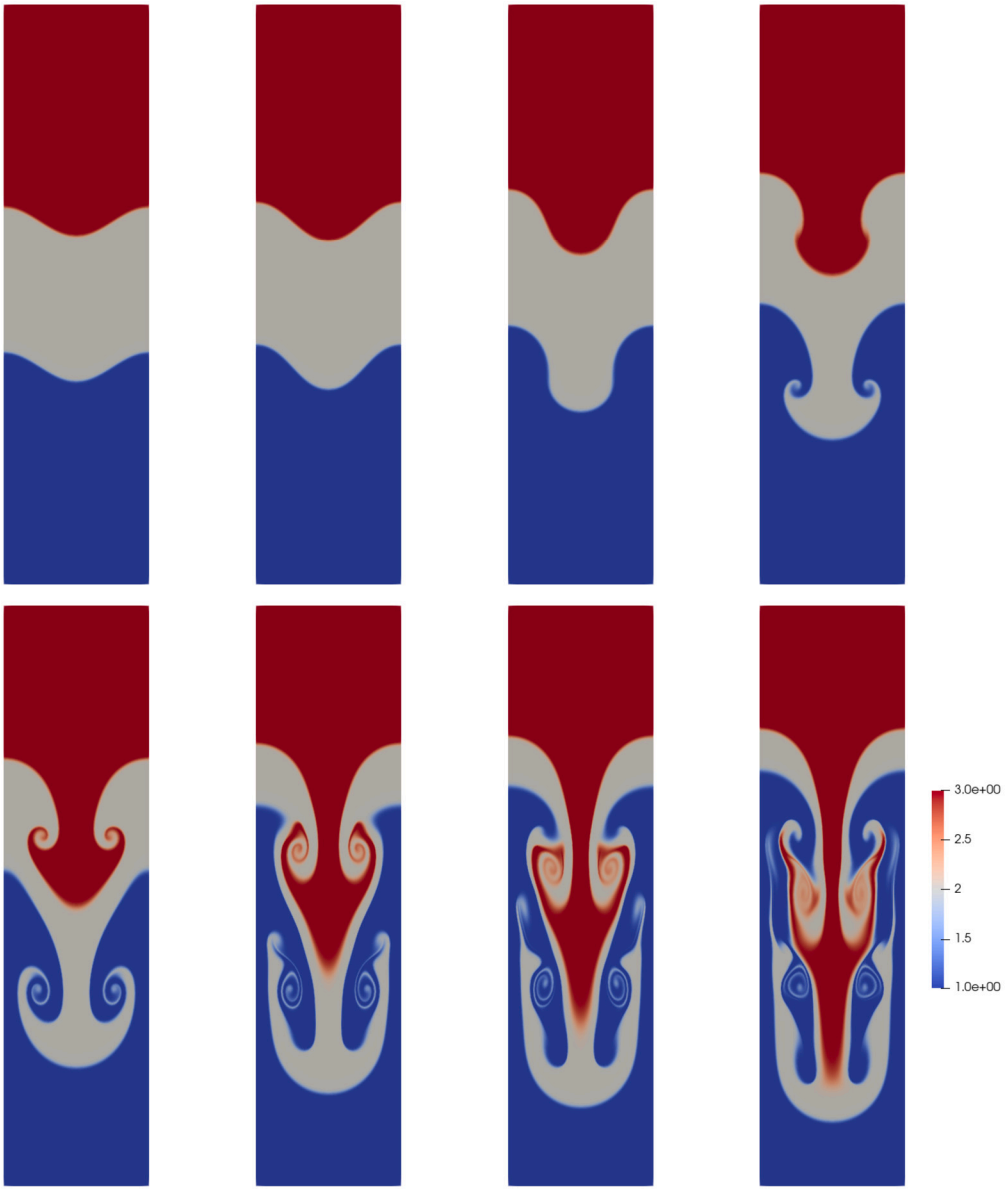

Fig. 16. Numerical solution of Rayleigh-Taylor instability with densities set at $3, 2, 1$ from top to bottom and $\nu = 0.001$ at $t = 0$, 0.25, 0.5, 0.75, 1, 1.25, 1.375, 1.5. Time step: $\delta t = 0.0025$, spatial mesh: $h = 1/200$, $\delta t/h = 0.5$.

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

No data was used for the research described in the article.

Acknowledgement

This work is supported by the National Natural Science Foundation of China under Grant No. 12231003 and No. 12371409.

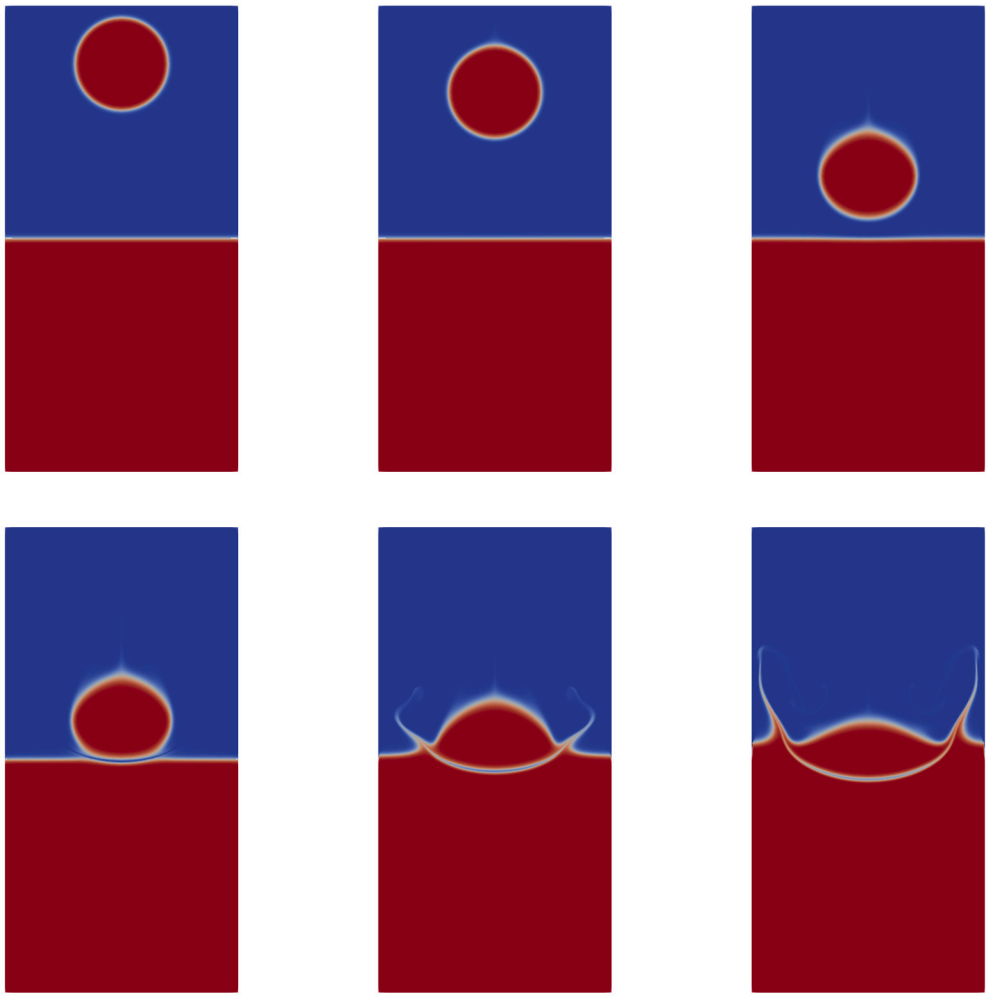

Fig. 17. Numerical solution of falling bubble at $t = 0, 0.5, 1, 1.1, 1.2, 1.3$. Time step: $\delta t = 0.001$, spatial mesh: $h = 1/200$, $\delta t/h = 0.2$.

References

[1] Ann S. Almgren, John B. Bell, Phillip Colella, Louis H. Howell, Michael L. Welcome, A conservative adaptive projection method for the variable density incompressible Navier-Stokes equations, J. Comput. Phys. 142 (1) (1998) 1–46.

[2] Rong An, Error analysis of a new fractional-step method for the incompressible Navier-Stokes equations with variable density, J. Sci. Comput. 84 (1) (2020) 3, 21.

[3] John B. Bell, Daniel L. Marcus, A second-order projection method for variable-density flows, J. Comput. Phys. 101 (2) (1992) 334–348.

[4] Wentao Cai, Buyang Li, Ying Li, Error analysis of a fully discrete finite element method for variable density incompressible flows in two dimensions, ESAIM: Math. Model. Numer. Anal. 55 (suppl.) (2021) S103–S147.

[5] Caterina Calgaro, Emmanuel Creusé, Thierry Goudon, An hybrid finite volume-finite element method for variable density incompressible flows, J. Comput. Phys. 227 (9) (2008) 4671–4696.

[6] Hongtao Chen, Jingjing Mao, Jie Shen, Error estimate of gauge-Uzawa methods for incompressible flows with variable density, J. Comput. Appl. Math. 364 (2020) 112321.

[7] Alexandre Joel Chorin, Numerical solution of the Navier-Stokes equations, Math. Comput. 22 (1968) 745–762.

[8] Yann Fraigneau, Jean-Luc Guermond, Luigi Quartapelle, Approximation of variable density incompressible flows by means of finite elements and finite volumes, Comput. Methods Appl. Mech. Eng. 17 (12) (2001) 893–902.

[9] Evan S. Gawlik, François Gay-Balmaz, A conservative finite element method for the incompressible Euler equations with variable density, J. Comput. Phys. 412 (2020) 109439.

[10] J.L. Guermond, P. Minev, Jie Shen, An overview of projection methods for incompressible flows, Comput. Methods Appl. Mech. Eng. 195 (44–47) (2006) 6011–6045.

[11] J.-L. Guermond, L. Quartapelle, A projection FEM for variable density incompressible flows, J. Comput. Phys. 165 (1) (2000) 167–188.

[12] J.-L. Guermond, Abner Salgado, A splitting method for incompressible flows with variable density based on a pressure Poisson equation, J. Comput. Phys. 228 (8) (2009) 2834–2846.

[13] J.-L. Guermond, Abner J. Salgado, Error analysis of a fractional time-stepping technique for incompressible flows with variable density, SIAM J. Numer. Anal. 49 (3) (2011) 917–944.

[14] J.L. Guermond, Jie Shen, A new class of truly consistent splitting schemes for incompressible flows, J. Comput. Phys. 192 (1) (2003) 262–276.

[15] Fukeng Huang, Jie Shen, Stability and error analysis of a class of high-order IMEX schemes for Navier-Stokes equations with periodic boundary conditions, SIAM J. Numer. Anal. 59 (6) (2021) 2926–2954.

[16] Hans Johnston, Jian-Guo Liu, Finite difference schemes for incompressible flow based on local pressure boundary conditions, J. Comput. Phys. 180 (1) (2002) 120–154.
[17] Buyang Li, Weifeng Qiu, Zongze Yang, A convergent post-processed discontinuous Galerkin method for incompressible flow with variable density, J. Sci. Comput. 91 (1) (2022) 2, 28.
[18] Maojun Li, Yongping Cheng, Jie Shen, Xiangxiong Zhang, A bound-preserving high order scheme for variable density incompressible Navier-Stokes equations, J. Comput. Phys. 425 (2021) 109906, 18.
[19] Ning Li, Jilian Wu, Xinlong Feng, Filtered time-stepping method for incompressible Navier-Stokes equations with variable density, J. Comput. Phys. 473 (2023) 111764, 24.
[20] Xiaoli Li, Jie Shen, Error analysis of the SAV-MAC scheme for the Navier-Stokes equations, SIAM J. Numer. Anal. 58 (5) (2020) 2465–2491.
[21] Xiaoli Li, Jie Shen, Zhengguang Liu, New SAV-pressure correction methods for the Navier-Stokes equations: stability and error analysis, Math. Comput. 91 (333) (2021) 141–167.
[22] Ying Li, Liquan Mei, Jiatai Ge, Feng Shi, A new fractional time-stepping method for variable density incompressible flows, J. Comput. Phys. 242 (2013) 124–137.
[23] Lianlei Lin, Zhiguo Yang, Suchuan Dong, Numerical approximation of incompressible Navier-Stokes equations based on an auxiliary energy variable, J. Comput. Phys. 388 (2019) 1–22.
[24] Chun Liu, Noel J. Walkington, Convergence of numerical approximations of the incompressible Navier-Stokes equations with variable density and viscosity, SIAM J. Numer. Anal. 45 (3) (2007) 1287–1304.
[25] Jian-Guo Liu, Jie Liu, Robert L. Pego, Stability and convergence of efficient Navier-Stokes solvers via a commutator estimate, Commun. Pure Appl. Math. 60 (10) (2007) 1443–1487.
[26] Lukas Lundgren, Murtazo Nazarov, A fully conservative and shift-invariant formulation for Galerkin discretizations of incompressible variable density flow, J. Comput. Phys. 510 (2024) 113086.
[27] Elbridge Gerry Puckett, Ann S. Almgren, John B. Bell, Daniel L. Marcus, William J. Rider, A high-order projection method for tracking fluid interfaces in variable density incompressible flows, J. Comput. Phys. 130 (2) (1997) 269–282.
[28] Jae-Hong Pyo, Jie Shen, Gauge-Uzawa methods for incompressible flows with variable density, J. Comput. Phys. 221 (1) (2007) 181–197.
[29] Jie Shen, Jie Xu, Jiang Yang, The scalar auxiliary variable (SAV) approach for gradient flows, J. Comput. Phys. 353 (2018) 407–416.
[30] Jie Shen, Jie Xu, Jiang Yang, A new class of efficient and robust energy stable schemes for gradient flows, SIAM Rev. 61 (3) (2019) 474–506.
[31] R. Témam, Sur l'approximation de la solution des équations de Navier-Stokes par la méthode des pas fractionnaires. II, Arch. Ration. Mech. Anal. 33 (1969) 377–385.
[32] Grétar Tryggvason, Numerical simulations of the Rayleigh-Taylor instability, J. Comput. Phys. 75 (2) (1988) 253–282.
[33] Ke Wu, Fukeng Huang, Jie Shen, A new class of higher-order decoupled schemes for the incompressible Navier-Stokes equations and applications to rotating dynamics, J. Comput. Phys. 458 (2022) 111097, 16.