Journal of Computational Physics 471 (2022) 111653

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp

An efficient IMEX-DG solver for the compressible Navier-Stokes equations for non-ideal gases

Giuseppe Orlando *, Paolo Francesco Barbante, Luca Bonaventura

MOX, Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, Milano, 20133, Italy

ARTICLE INFO

Article history:
Received 24 May 2022
Received in revised form 14 September 2022
Accepted 20 September 2022
Available online 26 September 2022

Keywords:
Navier-Stokes equations
Compressible flows
Discontinuous Galerkin methods
Implicit methods
ESDIRK methods

ABSTRACT

We propose an efficient, accurate and robust IMEX solver for the compressible Navier-Stokes equations describing non-ideal gases with general cubic equation of state and Stiffened-Gas EOS. The method is based on an *h*-adaptive Discontinuous Galerkin spatial discretization and on an Additive Runge Kutta IMEX method for time discretization. It is specifically tailored for low Mach number applications and allows to simulate low Mach regimes at a significantly reduced computational cost, while maintaining full second order accuracy also for higher Mach number regimes. The method has been implemented in the framework of the *deal.II* numerical library, whose adaptive mesh refinement capabilities are employed to enhance efficiency. Refinement indicators appropriate for real gas phenomena have been introduced. A number of numerical experiments on classical benchmarks for compressible flows and their extension to real gases demonstrate the properties of the proposed method.

1. Introduction

The efficient numerical solution of the compressible Navier-Stokes equations poses several major computational challenges. In particular, for flow regimes characterized by low Mach number and moderate Reynolds number values, severe time step restrictions may be required by standard explicit time discretization methods. The use of implicit and semi-implicit methods has a long tradition in low Mach number flows, see for example the seminal papers [15,16,48], as well as many other contributions in the literature on numerical weather prediction, see e.g. [7,24,26,27,36,47,49,55] and the reviews in [52,9]. Other contributions have been proposed in the literature on more classical computational fluid dynamics, see e.g. [6,5,11,13,17,41,54]. Many of these contributions focus exclusively on the equations of motion of an ideal gas and their extension to real gases is not necessarily straightforward. Stability concerns are even more critical in these particular regimes for spatial discretizations based on the Discontinuous Galerkin (DG) method (see e.g. [25,32] for a general presentation of this method), which is the spatial discretization used in many of the above referenced papers. In this work we propose a discretization approach for the equations of non-ideal compressible gas dynamics which allows to employ a generic cubic equation of state (EOS), as well as other simpler models such as the stiffened gas equation. The use of a general cubic EOS presents numerical challenges that allows in principle to handle more general cases of tabulated or analytical EOS. We discuss in detail the implications of this more physically realistic model of gas behavior for the numerical solution procedure, generalizing and extending previous proposals in this direction such as [17]. The functional dependency of the EOS on

* Corresponding author.
E-mail addresses: giuseppe.orlando@polimi.it (G. Orlando), paolo.barbante@polimi.it (P.F. Barbante), luca.bonaventura@polimi.it (L. Bonaventura).

https://doi.org/10.1016/j.jcp.2022.111653

each thermodynamic variable has been carefully analyzed, in order to identify a practical procedure to handle it within an implicit solver. We use an accurate and flexible discontinuous DG space discretization and a second order implicit-explicit (IMEX) time discretization, see e.g. [33,45,10], combined to obtain an efficient method for compressible flow of real gases at low to moderate Mach numbers. In this way, we aim to derive a method that can then be easily extended to handle multiphase compressible flows, where a number of coupling and forcing terms arise that cannot be dealt with efficiently by straightforward application of conventional solvers. More specifically, we extend to second order in time the approach of [15,17], coupling implicitly the energy equation to the momentum equation, while treating the continuity equation in an explicit fashion. Our treatment also provides an outline of how a generic IMEX method based on a Diagonally Implicit Runge Kutta can be extended along the same lines. Notice that, with respect to the IMEX approach proposed for the Euler equations in [57], the technique presented here does not require to introduce reference solutions, does not introduce inconsistencies in the splitting with respect to a reference solution and only requires the solution of linear system of a size equal to that of the number of discrete degrees of freedom needed to describe a scalar variable, as in [17]. A conceptually similar approach has been used in [36,49] for the discretization employed in the IFS-FVM atmospheric model. In order to obtain a formulation that is efficient also in presence of non negligible viscous terms, we resort to an operator splitting approach, see e.g. [39]. As commonly done in numerical models for atmospheric physics, we split the hyperbolic part of the problem, which is treated by an IMEX extension of the method proposed in [17], from the diffusive terms, which are treated implicitly. Second order accuracy can then be obtained by the Strang splitting approach [39,53].

For the spatial discretization, we rely on the DG approach implemented in the numerical library *deal.II* [3], which is a very convenient environment to develop a reliable and easily accessible tool for large scale industrial applications, as we have already shown in [44] for the case of the incompressible Navier-Stokes equations. This software also provides h-refinement capabilities that are exploited by the proposed method. For the specific case of real gases, novel physically based refinement criteria have been proposed and tested, which allow to track accurately convection phenomena also for more general equations of state. The numerical experiments reported below show the ability of the proposed scheme and of its adaptive implementation to perform accurate simulations in a range of different settings appropriate to describe non-ideal gas dynamics.

The model equations and their non-dimensional formulation are reviewed in Section 2. The time discretization approach is outlined and discussed in Section 4. The spatial discretization is presented in Section 5. The validation of the proposed method and its application to a number of significant benchmarks is reported in Section 6. Some conclusions and perspectives for future work are presented in Section 7.

2. The compressible Navier-Stokes equations

Let $\Omega \subset \mathbb{R}^d, 2 \leq d \leq 3$ be a connected open bounded set with a sufficiently smooth boundary $\partial\Omega$ and denote by $\mathbf{x}$ the spatial coordinates and by t the temporal coordinate. We consider the classical unsteady compressible Navier-Stokes equations, written in flux form as:

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) &= 0 \\ \frac{\partial(\rho\mathbf{u})}{\partial t} + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla p &= \nabla\cdot\boldsymbol{\tau} + \rho\mathbf{f} \\ \frac{\partial(\rho E)}{\partial t} + \nabla\cdot[(\rho E + p)\mathbf{u}] &= \nabla\cdot(\boldsymbol{\tau}\mathbf{u} - \mathbf{q}) + \rho\mathbf{f}\cdot\mathbf{u} \end{aligned} \tag{1}$$

for $\mathbf{x} \in \Omega$, $t \in [0, T_f]$, along with suitable initial and boundary conditions to be discussed later. Here T_f is the final time, ρ is the density, $\mathbf{u}$ is the fluid velocity, p is the pressure, $\mathbf{q}$ denotes the heat flux and $\mathbf{f}$ represents volumetric forces. ρE is the total energy, which can be rewritten as $\rho E = \rho e + \rho k$, where e is the internal energy and $k = \|\mathbf{u}\|^2/2$ is the kinetic energy. At this stage, no more specific assumptions are made on the fluid and the choices of equations of state will be specified in the following. We also introduce the specific enthalpy $h = e + p/\rho$ and remark that one can also rewrite the energy flux as

$$(\rho E + p)\,\mathbf{u} = \left(e + k + \frac{p}{\rho}\right)\rho\mathbf{u} = (h + k)\,\rho\mathbf{u}.$$

We assume that $\mathbf{q} = -\kappa\nabla T$, where T denotes the absolute temperature and κ the thermal conductivity. Furthermore, we assume that the linear stress constitutive equation holds and we neglect the bulk viscosity and we denote the shear viscosity as μ, so that

$$\boldsymbol{\tau} = \mu\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right) - \frac{2\mu}{3}(\nabla\cdot\mathbf{u})\mathbf{I}.$$

For the sake of simplicity, we also assume constant values for both μ and κ. This choice can be justified by considering that we aim to simulate regimes with limited variations of temperature and density and, moreover, we are mainly interested in time scales where diffusive effects play a less relevant role. The equations can then be rewritten as

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0\\
&\frac{\partial(\rho\mathbf{u})}{\partial t} + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \nabla p = \mu\,\nabla\cdot\left[\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right) - \frac{2}{3}(\nabla\cdot\mathbf{u})\mathbf{I}\right] + \rho\mathbf{f}\\
&\frac{\partial(\rho E)}{\partial t} + \nabla\cdot[(h+k)\rho\mathbf{u}] = \mu\,\nabla\cdot\left[\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right)\mathbf{u} - \frac{2}{3}(\nabla\cdot\mathbf{u})\mathbf{u}\right] + \kappa\Delta T + \rho\mathbf{f}\cdot\mathbf{u}.
\end{aligned} \tag{2}$$

We now introduce reference scaling values $\mathcal{L}, \mathcal{T}, \mathcal{U}$ for the length, time and velocity, respectively, as well as reference values $\mathcal{P}, \mathcal{R}, \Theta, \mathcal{E}, \mathcal{I}$ for pressure, density, temperature, total energy and internal energy, respectively. We assume unit Strouhal number $St = \mathcal{L}/\mathcal{U}\mathcal{T} \approx 1$. This choice is the standard one in the case of advection dominated problems, see e.g. [34,41]. We also assume that the enthalpy scales like $\mathcal{I} + \mathcal{P}/\mathcal{R}$ and that

$$\mathcal{I} \approx \frac{\mathcal{P}}{\mathcal{R}} \qquad \mathcal{E} \approx \mathcal{I} + \mathcal{U}^2.$$

The model equations can then be written in non-dimensional form as

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0\\
&\frac{\partial \rho\mathbf{u}}{\partial t} + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \frac{\mathcal{P}}{\mathcal{R}\mathcal{U}^2}\nabla p = \frac{\mu}{\mathcal{R}\mathcal{U}\mathcal{L}}\,\nabla\cdot\left[\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right) - \frac{2}{3}(\nabla\cdot\mathbf{u})\mathbf{I}\right] + \frac{\mathcal{T}}{\mathcal{U}}\rho\mathbf{f}\\
&\frac{\partial \rho E}{\partial t} + \nabla\cdot\left[\left(h\frac{\mathcal{I}+\mathcal{P}/\mathcal{R}}{\mathcal{E}} + k\frac{\mathcal{U}^2}{\mathcal{E}}\right)\rho\mathbf{u}\right] = \frac{\mu\mathcal{U}}{\mathcal{R}\mathcal{E}\mathcal{L}}\,\nabla\cdot\left[\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right)\mathbf{u} - \frac{2}{3}(\nabla\cdot\mathbf{u})\mathbf{u}\right] + \frac{\kappa\Theta}{\mathcal{R}\mathcal{E}\mathcal{U}\mathcal{L}}\Delta T + \frac{\mathcal{L}}{\mathcal{E}}\rho\mathbf{f}\cdot\mathbf{u}.
\end{aligned} \tag{3}$$

Notice that, with a slight abuse of notation, we have kept the same notation for the non-dimensional variables. We then define the Reynolds, Prandtl and Mach numbers as

$$Re = \frac{\mathcal{R}\mathcal{U}\mathcal{L}}{\mu} \quad \kappa = \frac{c_p\mu}{Pr} \quad M^2 = \frac{\mathcal{R}\mathcal{U}^2}{\mathcal{P}},$$

where c_p denotes the specific heat at constant pressure, so that, for low and moderate Mach numbers,

$$\begin{aligned}
&\frac{\mathcal{U}^2}{\mathcal{E}} = \frac{1}{\frac{\mathcal{I}}{\mathcal{U}^2}+1} = O(M^2)\\
&\frac{\mathcal{I}+\mathcal{P}/\mathcal{R}}{\mathcal{E}} = \frac{\frac{\mathcal{I}}{\mathcal{U}^2}+\frac{1}{M^2}}{\frac{\mathcal{I}}{\mathcal{U}^2}+1} = O(1).
\end{aligned} \tag{4}$$

This justifies, in the above mentioned regimes, methods in which an implicit coupling between the pressure gradient and the energy flux is enforced. This strategy has been proposed in the seminal paper [15] and in the more recent works [17,41]. We finally assume that the only acting volumetric force is gravity, so that $\mathbf{f} = -g\mathbf{k}$, where g denotes the acceleration of gravity and $\mathbf{k}$ the upward pointing unit vector in the standard Cartesian reference frame. It follows that

$$\begin{aligned}
&\frac{\mathcal{T}}{\mathcal{U}}\rho g = \frac{g\mathcal{T}\mathcal{U}}{\mathcal{U}^2}\rho = \frac{g\mathcal{L}}{\mathcal{U}^2}\rho = \frac{\rho}{Fr^2} \quad Fr^2 = \frac{\mathcal{U}^2}{g\mathcal{L}}\\
&\frac{\mathcal{L}}{\mathcal{E}}\rho g = \frac{g\rho\mathcal{L}}{\mathcal{I}+\frac{\mathcal{P}}{\mathcal{R}}+\mathcal{U}^2} = \frac{g\rho\mathcal{L}}{\mathcal{U}^2}\frac{1}{\frac{\mathcal{I}}{\mathcal{U}^2}+1+\frac{1}{M^2}} = \frac{\rho}{Fr^2}O(M^2).
\end{aligned} \tag{5}$$

As a result, we will consider the non-dimensional model equations

$$\begin{aligned}
&\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0\\
&\frac{\partial \rho\mathbf{u}}{\partial t} + \nabla\cdot(\rho\mathbf{u}\otimes\mathbf{u}) + \frac{1}{M^2}\nabla p = \frac{1}{Re}\,\nabla\cdot\left[\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right) - \frac{2}{3}(\nabla\cdot\mathbf{u})\mathbf{I}\right] - \frac{\rho}{Fr^2}\mathbf{k}\\
&\frac{\partial \rho E}{\partial t} + \nabla\cdot\left[\left(h + kM^2\right)\rho\mathbf{u}\right] = \frac{M^2}{Re}\,\nabla\cdot\left[\left(\nabla\mathbf{u} + \nabla\mathbf{u}^T\right)\mathbf{u} - \frac{2}{3}(\nabla\cdot\mathbf{u})\mathbf{u}\right] + \frac{1}{PrRe}\Delta T - \rho\frac{M^2}{Fr^2}\mathbf{k}\cdot\mathbf{u},
\end{aligned} \tag{6}$$

where we have taken $c_p\Theta/\mathcal{E} \approx 1$, which can be justified at moderate values of the Mach number. Notice that these non-dimensional equations are very similar to those derived in [41].

3. The equation of state

The above equations must be complemented by an equation of state (EOS) for the compressible fluid. A classical choice is that of an ideal gas; in the non-dimensional variables introduced above the equation that links together pressure, density and internal energy is given by

$$p = (\gamma - 1)\left(\rho E - \frac{1}{2}M^2 \rho \mathbf{u}\cdot\mathbf{u}\right), \tag{7}$$

with γ denoting the specific heats ratio. The above relation is valid only in case of constant γ [56]. An example of non-ideal gas equation of state is given by the general cubic equation of state, whose equation linking together internal energy, density and temperature, according to [56], is given in dimensional form by:

$$e = e^{\#}(T) + \frac{a(T) - T\frac{da}{dT}}{b} U(\rho, b, r_1, r_2), \tag{8}$$

where $e^{\#}$ denotes the internal energy of an ideal gas at temperature T, whereas a and b are suitable parameters that characterize the gas behavior. In case $c_v = \frac{de^{\#}}{dT}$ is constant, we can write

$$e = c_v T + \frac{a(T) - T\frac{da}{dT}}{b} U(\rho, b, r_1, r_2). \tag{9}$$

In case the previous hypothesis does not hold, we analogously define $\overline{c_v}(T) = \frac{e^{\#}(T)}{T}$, so that (8) reads as follows

$$e = \overline{c_v}(T)T + \frac{a(T) - T\frac{da}{dT}}{b} U(\rho, b, r_1, r_2). \tag{10}$$

The quantity $\overline{c_v}(T)$ should not be understood as a real specific heat, but only as a convenient way of writing the above EOS. The function U and the constants r_1 and r_2 depend on the specific equation of state. In this work, we consider the van der Waals EOS, for which $r_1 = r_2 = 0$ and

$$U = -b\rho \tag{11}$$

and the Peng-Robinson EOS, for which $r_1 = -1 - \sqrt{2}$, $r_2 = -1 + \sqrt{2}$ and

$$U = \frac{1}{r_1 - r_2}\log\left(\frac{1 - \rho b r_1}{1 - \rho b r_2}\right). \tag{12}$$

The link between pressure, density and temperature for the general cubic EOS in dimensional form can be expressed as follows:

$$p = \frac{\rho R_g T}{1 - \rho b} - \frac{a\rho^2}{(1 - \rho b r_1)(1 - \rho b r_2)}, \tag{13}$$

with R_g denoting the specific gas constant. Notice that for $a = b = 0$, the expression for the pressure of an ideal gas is retrieved. For the sake of clarity, we introduce the following non-dimensional variables

$$\tilde{R}_g = \frac{\mathcal{R}\Theta}{\mathcal{P}} R_g \qquad \tilde{a} = a\frac{\mathcal{R}^2}{\mathcal{P}} \qquad \tilde{b} = \mathcal{R} b, \tag{14}$$

so that (13) can be rewritten in non-dimensional form as:

$$p = \frac{\rho \tilde{R}_g T}{1 - \rho\tilde{b}} - \frac{\tilde{a}\rho^2}{\left(1 - \rho\tilde{b} r_1\right)\left(1 - \rho\tilde{b} r_2\right)}. \tag{15}$$

Finally, we define $\tilde{c}_v(T) = \overline{c_v}\frac{\mathcal{R}\Theta}{\mathcal{P}}$, so that the non-dimensional version of (10) reads as follows:

$$e = \tilde{c}_v(T)T + \frac{\tilde{a}(T) - T\frac{d\tilde{a}}{dT}}{\tilde{b}} U\left(\rho, \tilde{b}, r_1, r_2\right). \tag{16}$$

The last example of non-ideal gas considered is represented by the Stiffened Gas equation of state (SG-EOS) [37], which is interesting for its convexity property and is given in dimensional variables by:

$$p = (\gamma - 1)\left(\rho E - \frac{1}{2}\rho\mathbf{u}\cdot\mathbf{u} - \rho q_\infty\right) - \gamma\pi_\infty, \tag{17}$$

where q_∞ and π_∞ are parameters that determine the gas characteristics. Notice that, for this equation of state, the parameters have to be taken constant [37]. We define

$$\tilde{q}_\infty = \frac{\mathcal{R}}{\mathcal{P}} q \qquad \tilde{\pi}_\infty = \frac{\pi}{\mathcal{P}}, \tag{18}$$

so that (17) reads in terms of non-dimensional variables as follows:

$$p = (\gamma - 1)\left(\rho E - \frac{1}{2} M^2 \rho \mathbf{u}\cdot\mathbf{u} - \rho \tilde{q}_\infty\right) - \gamma \tilde{\pi}_\infty. \tag{19}$$

Finally, the link between pressure, density and temperature for the SG-EOS can be written as:

$$T = \frac{p + \pi_\infty}{\rho\,(\gamma - 1)\, c_v}. \tag{20}$$

We define $\tilde{c}_v = c_v \frac{\mathcal{R}\Theta}{\mathcal{P}}$, so that the non-dimensional version of (20) is given by:

$$T = \frac{p + \tilde{\pi}_\infty}{\rho\,(\gamma - 1)\, \tilde{c}_v}. \tag{21}$$

More accurate and general equations of state are available in literature [51], [38], but the above choices, in particular the cubic EOS, are suitable for the regimes of interest, involve non trivial non-linearities and provide a sufficient level of complexity for the validation of the proposed numerical scheme.

An important parameter to determine the regime in which real gas effects are relevant is the so-called compressibility factor. In terms of dimensional variables, it is given by

$$z = \frac{p}{\rho R_g T}. \tag{22}$$

When $z \approx 1$, the gas can be treated as an ideal one, while the ideal gas law is no longer valid for values of z very different from 1.

3.1. Analysis of isentropic processes for the general cubic EOS

We recall here the definition of the potential temperature θ for an ideal gas, which is commonly used in applications to atmospheric flows

$$\theta = \frac{T}{\Pi}, \tag{23}$$

with $\Pi = \left(\frac{p}{p_0}\right)^{\frac{\gamma-1}{\gamma}}$ denoting the so-called Exner pressure. Here, p_0 denotes a reference pressure value and, in this work, we consider $p_0 = 10^5$ Pa. For isentropic processes, the initial condition is typically given as a perturbation with respect to a constant background potential temperature. Therefore, as discussed in Section 6, the gradient of the potential temperature is a good candidate to drive the mesh adaptation procedure. Our goal is to employ adaptive mesh refinement also in the case of real gases, so as to enhance the computational efficiency, but for non-ideal gases the definition of a potential temperature or of quantities with similar properties is not trivial. We propose here a quantity with a simple definition, stemming from the analysis of isentropic processes, that is valid for the general cubic equation of state in the case $\frac{da}{dT} = 0$ and $\frac{de^\#}{dT} = c_v = const$. For the sake of simplicity, in order to avoid the influence of reference quantities, we report the computations using dimensional variables. Let us recall the first law of thermodynamics

$$de = Tds - pdv = Tds + \frac{p}{\rho^2} d\rho, \tag{24}$$

where s denotes the specific entropy and v is the inverse of the density. Dividing both sides in the previous equation by T we obtain

$$\frac{1}{T} de = ds + \frac{p}{\rho^2 T} d\rho \tag{25}$$

which in an isentropic process reduces to

$$\frac{1}{T} de - \frac{p}{\rho^2 T} d\rho = 0. \tag{26}$$

Under the specific assumptions made, we get

$$\frac{c_v}{T}dT + \frac{a}{b}\frac{1}{T}\frac{\partial U}{\partial \rho}d\rho = 0. \tag{27}$$

The EOS can be rewritten in the following form [56]

$$T = \left[p + \frac{a\rho^2}{(1-\rho b r_1)(1-\rho b r_2)}\right]\frac{(1-\rho b)}{\rho R_g}. \tag{28}$$

If we substitute (28) into (27), we obtain

$$\frac{c_v}{T}dT + \left(\frac{a}{b}\frac{\partial U}{\partial \rho}\rho - \frac{p}{\rho}\right)\frac{R}{\rho(1-\rho b)}\frac{(1-\rho b r_1)(1-\rho b r_2)}{p(1-\rho b r_1)(1-\rho b r_2)+a\rho^2}d\rho = 0. \tag{29}$$

In the case of van der Waals EOS, $U = -b\rho$ and $\frac{\partial U}{\partial \rho} = -b$, whereas in the case of Peng-Robinson EOS one has $U = \frac{1}{r_1 - r_2}\log\left(\frac{1-\rho b r_1}{1-\rho b r_2}\right)$ and $\frac{\partial U}{\partial \rho} = -\frac{b}{(1-\rho b r_1)(1-\rho b r_2)}$. Since, for van der Waals EOS $r_1 = r_2 = 0$, the expression

$$\frac{\partial U}{\partial \rho} = -\frac{b}{(1-\rho b r_1)(1-\rho b r_2)} \tag{30}$$

can be applied for both van der Waals and Peng-Robinson EOS. Hence, (29) reduces to

$$\frac{c_v}{T}dT - \frac{R_g}{\rho(1-\rho b)}d\rho = 0, \tag{31}$$

which can then be integrated to yield

$$c_v \log(T) - 2R_g \operatorname{atanh}(2\rho b - 1) = const \tag{32}$$

or, equivalently,

$$\log(T) - 2\frac{R_g}{c_v}\operatorname{atanh}(2\rho b - 1) = const. \tag{33}$$

From (31), it is immediate to verify that, in the non-dimensional case, we obtain

$$\log(T^*) - 2\frac{\tilde{R}_g}{\tilde{c}_v}\operatorname{atanh}\left(2\rho^*\tilde{b} - 1\right) = const, \tag{34}$$

where the symbol $*$ denotes non-dimensional variables. In the more general case $\frac{da}{dT} \neq 0$, it can be shown that [42]

$$\frac{p}{\rho^{\gamma_{p\rho}}} = const, \tag{35}$$

where

$$\gamma_{p\rho} = \frac{c^2}{M^2}\frac{\rho}{p}. \tag{36}$$

The evaluation of this quantity is less straightforward than that of (34), since it involves the computation of non trivial derivatives, see the discussions in Appendix B and [42]. Both the conserved quantities will be employed in Section 6 for adaptive simulations.

4. The time discretization strategy

In the low Mach number limit, terms proportional to $1/M^2$ in (6) yield stiff components of the resulting semidiscretized ODE system. Therefore, following as remarked above [15,17], it is appropriate to couple implicitly the energy equation to the momentum equation, while the continuity equation can be discretized in a fully explicit fashion. While this would be sufficient to yield an efficient time discretization approach for the purely hyperbolic system associated to (6) in absence of gravity, in regimes for which

$$Pr \approx O(1), \quad Fr \ll 1$$

thermal diffusivity and gravity terms would also have to be treated implicitly for the time discretization methods to be efficient. Straightforward application of any monolithic solver would then yield large algebraic systems with multiple couplings between discrete DOF associated to different physical variables. To avoid this, we resort to an operator splitting approach, see e.g. [39], as commonly done in numerical models for atmospheric physics. More specifically, after spatial discretization, all diffusive terms on the right hand side of (6) are split from the hyperbolic part on the left hand side. The hyperbolic part

Table 1
Butcher tableaux of the explicit ARK2 method.

0	0		
χ	χ	0	
1	$1-a_{32}$	a_{32}	0
	$\frac{1}{2}-\frac{\chi}{4}$	$\frac{1}{2}-\frac{\chi}{4}$	$\frac{\chi}{2}$

Table 2
Butcher tableaux of the implicit ARK2 method.

0	0		
χ	$\frac{\chi}{2}$	$\frac{\chi}{2}$	
1	$\frac{1}{2\sqrt{2}}$	$\frac{1}{2\sqrt{2}}$	$1-\frac{1}{\sqrt{2}}$
	$\frac{1}{2}-\frac{\chi}{4}$	$\frac{1}{2}-\frac{\chi}{4}$	$\frac{\chi}{2}$

is treated in a similar fashion to what outlined in [17], while the diffusive terms are treated implicitly. For simplicity, the gravity terms will be treated explicitly in this first attempt and only a basic, first order splitting will be described, which can be easily improved to second order accuracy by the Strang splitting approach [39,53].

For the time discretization, an IMplicit EXplicit (IMEX) Additive Runge Kutta method (ARK) [33] method will be used. These methods are useful for time dependent problems that can be formulated as $\mathbf{y}' = \mathbf{f}_S(\mathbf{y}, t) + \mathbf{f}_{NS}(\mathbf{y}, t)$, where the S and NS subscripts denote the stiff and non-stiff components of the system, to which the implicit and explicit companion methods are applied, respectively. If $\mathbf{v}^n \approx y(t^n)$, the generic $s-$ stage IMEX-ARK method can be defined as

$$\begin{aligned}\mathbf{v}^{(n,l)} = \mathbf{v}^n \; + \; \Delta t \sum_{m=1}^{s-1} \Big(& a_{lm}\mathbf{f}_{NS}(\mathbf{v}^{(n,m)}, t + c_m \Delta t) \\ & + \; \tilde{a}_{lm}\mathbf{f}_S(\mathbf{v}^{(n,m)}, t + c_m \Delta t) \Big) + \Delta t \, \tilde{a}_{ll} \, \mathbf{f}_S(\mathbf{v}^{(n,l)}, t + c_l \Delta t),\end{aligned} \tag{37}$$

where $l = 1, \ldots, s$. After computation of the intermediate stages, $\mathbf{v}^{n+1}$ is computed as

$$\mathbf{v}^{n+1} = \mathbf{v}^n + \Delta t \sum_{l=1}^{s} b_l \left[\mathbf{f}_{NS}(\mathbf{v}^{(n,l)}, t + c_l \Delta t) + \mathbf{f}_S(\mathbf{v}^{(n,l)}, t + c_l \Delta t) \right]. \tag{38}$$

IMEX-ARK methods are represented compactly by the following two Butcher tableaux [14].

$$\begin{array}{c|c} c & A \\ \hline & b^T \end{array} \qquad \begin{array}{c|c} \tilde{c} & \tilde{A} \\ \hline & \tilde{b}^T \end{array}$$

with $A = \{A_{ij}\}, b = \{b_i\}, c = \{c_i\}, \tilde{A} = \{\tilde{a}_{ij}\}, \tilde{b} = \{\tilde{b}_i\}$ and $\tilde{c} = \{\tilde{c}_i\}$. Coefficients $a_{lm}, \tilde{a}_{lm}, c_l$ and b_l are determined so that the method is consistent of a given order. In particular, in addition to the order conditions specific to each sub-method, the coefficients should respect coupling conditions. Here, we consider a variant of the IMEX method proposed in [26], whose coefficients are presented in the Butcher tableaux reported in Tables 1 and 2 for the explicit and implicit method, respectively, where $\chi = 2 - \sqrt{2}$. The coefficients of the explicit method were proposed in [26], while the implicit method, also employed in the same paper, coincides indeed for the above choice of χ with the TR-BDF2 method proposed in [4,31]. The TR-BDF2 has been shown to be stiffly accurate in [31] and has been very successfully employed in simulations of low Mach number flows with gravity in [55]. The corresponding IMEX method with this implicit part was also used successfully in analogous applications in [26]. A solver based on this implicit method for the incompressible Navier Stokes equations formulated in pseudo-compressible fashion has been proposed in [44], in which applications to cases with extremely small Mach numbers are presented, thus providing ample guarantees the robustness of the proposed approach in the low Mach number limit. Notice that, as discussed in [26], the choice of the coefficients

$$a_{32} = \frac{7 - 2\chi}{6} \quad 1 - a_{32} = \frac{2\chi - 1}{6}$$

in the third stage of the explicit part of the method is arbitrary. In [26], the above value of a_{32} was chosen with the aim of maximizing the stability region of the method. However, if a stability and absolute monotonicity analysis is carried out, as discussed in detail in Appendix A, it can be seen that different choices might be more advantageous, in order to improve the monotonicity of the method without compromising its stability. In particular, the value of $a_{32} = 1/2$ appears to be a more appropriate choice, as also demonstrated by the numerical experiments reported in Section 6.

Finally, even though we focus here on this specific second order method, the same strategy we outline is applicable to a generic DIRK method. In particular, higher order methods could be considered for coupling to high order spatial discretization, even though the effective overall accuracy would be limited by the splitting procedure if gravity and viscous terms are present.

4.1. Discretization of hyperbolic and forcing terms

We now describe the application of this IMEX method and of the splitting approach outlined above to equations (6). Notice that, for simplicity, we first present the time semi-discretization only, while maintaining the continuous form of (6) with respect to the spatial variables. The detailed description of the algebraic problems resulting from the full space and time discretization according to the method outlined here will be presented in Section 5.

For each time step, we first consider the discretization of the hyperbolic and forcing terms. For the first stage of the method, one simply has

$$\rho^{(n,1)} = \rho^n \quad \mathbf{u}^{(n,1)} = \mathbf{u}^n \quad E^{(n,1)} = E^n.$$

For the second stage, we can write formally

$$\begin{aligned} \rho^{(n,2)} &= \rho^n - a_{21}\Delta t\,\nabla\cdot\left(\rho^n\mathbf{u}^n\right) \\ \rho^{(n,2)}\mathbf{u}^{(n,2)} &+ \tilde{a}_{22}\frac{\Delta t}{M^2}\nabla p^{(n,2)} = \mathbf{m}^{(n,2)} \\ \rho^{(n,2)}E^{(n,2)} &+ \tilde{a}_{22}\Delta t\,\nabla\cdot\left(h^{(n,2)}\rho^{(n,2)}\mathbf{u}^{(n,2)}\right) = \hat{e}^{(n,2)}, \end{aligned} \tag{39}$$

where we have set

$$\begin{aligned} \mathbf{m}^{(n,2)} &= \rho^n\mathbf{u}^n \\ &\quad - a_{21}\Delta t\,\nabla\cdot\left(\rho^n\mathbf{u}^n\otimes\mathbf{u}^n\right) - \tilde{a}_{21}\frac{\Delta t}{M^2}\nabla p^n - a_{21}\frac{\Delta t}{Fr^2}\rho^n\mathbf{k} \end{aligned} \tag{40}$$

$$\begin{aligned} \hat{e}^{(n,2)} &= \rho^n E^n - \tilde{a}_{21}\Delta t\,\nabla\cdot\left(h^n\rho^n\mathbf{u}^n\right) - a_{21}\Delta t M^2\,\nabla\cdot\left(k^n\rho^n\mathbf{u}^n\right) \\ &\quad - a_{21}\frac{\Delta t M^2}{Fr^2}\rho^n\mathbf{k}\cdot\mathbf{u}^n. \end{aligned}$$

Notice that, substituting formally $\rho^{(n,2)}\mathbf{u}^{(n,2)}$ into the energy equation and taking into account the definitions $\rho E = \rho e + M^2\rho k$ and $h = e + p/\rho$, the above system can be solved by computing the solution of

$$\begin{aligned} &\rho^{(n,2)}[e(p^{(n,2)},\rho^{(n,2)}) + M^2k^{(n,2)}] \\ &- \tilde{a}_{22}^2\frac{\Delta t^2}{M^2}\nabla\cdot\left[\left(e(p^{(n,2)},\rho^{(n,2)}) + \frac{p^{(n,2)}}{\rho^{(n,2)}}\right)\nabla p^{(n,2)}\right] \\ &+ \tilde{a}_{22}\Delta t\,\nabla\cdot\left[\left(e(p^{(n,2)},\rho^{(n,2)}) + \frac{p^{(n,2)}}{\rho^{(n,2)}}\right)\mathbf{m}^{(n,2)}\right] = \hat{e}^{(n,2)} \end{aligned} \tag{41}$$

in terms of $p^{(n,2)}$ according to the fixed point procedure described in [17]. More specifically, setting $\xi^{(0)} = p^{(n,2)}, k^{(n,2,0)} = k^{(n,1)}$, one solves for $l = 1,\ldots,L$ the equation

$$\begin{aligned} \rho^{(n,2)}e(\xi^{(l+1)},\rho^{(n,2)}) &- \tilde{a}_{22}^2\frac{\Delta t^2}{M^2}\nabla\cdot\left[\left(e(\xi^{(l)},\rho^{(n,2)}) + \frac{\xi^{(l)}}{\rho^{(n,2)}}\right)\nabla\xi^{(l+1)}\right] \\ &= \hat{e}^{(n,2)} - M^2\rho^{(n,2)}k^{(n,2,l)} \\ &\quad - \tilde{a}_{22}\Delta t\,\nabla\cdot\left[\left(e(\xi^{(l)},\rho^{(n,2)}) + \frac{\xi^{(l)}}{\rho^{(n,2)}}\right)\mathbf{m}^{(n,2)}\right] \end{aligned} \tag{42}$$

and updates the velocity as

$$\mathbf{u}^{(n,2,l+1)} + \frac{\tilde{a}_{22}\Delta t}{\rho^{(n,2)}M^2}\nabla\xi^{(l+1)} = \mathbf{m}^{(n,2)}.$$

In the case of SG-EOS, $\rho^{(n,2)} e\left(\xi^{(l+1)}, \rho^{(n,2)}\right)$ contains a term that only depends on the density, as evident from Equation (19) and, therefore, it has to be properly considered in the right-hand side of (42). On the other hand, the general cubic EOS (16) contains products of quantities depending on temperature and on density. For the sake of simplicity, in order to avoid the solution of a nonlinear equation for each quadrature node, in these cases we keep the temperature at the value in the previous iteration of the fixed point procedure, so as to obtain:

$$\begin{aligned}
&\frac{\tilde{c}_v\left(T\left(\xi^{(l)}, \rho^{(n,2)}\right)\right)}{\tilde{R}_g} \xi^{(l+1)} \left(1-\rho^{(n,2)}\tilde{b}\right) - \\
&\tilde{a}_{22}^2 \frac{\Delta t^2}{M^2} \nabla\cdot\left[\left(e\left(\xi^{(l)}, \rho^{(n,2)}\right) + \frac{\xi^{(l)}}{\rho^{(n,2)}}\right) \nabla \xi^{(l+1)}\right] = \\
&\hat{e}^{(n,2)} - M^2 \rho^{(n,2)} k^{(n,2,l)} - \\
&\frac{\tilde{c}_v\left(T\left(\xi^{(l)}, \rho^{(n,2)}\right)\right)}{\tilde{R}_g} \frac{\tilde{a}\left(T\left(\xi^{(l)}, \rho^{(n,2)}\right)\right)\left(\rho^{(n,2)}\right)^2}{\left(1-\rho^{(n,2)}\tilde{b} r_1\right)\left(1-\rho^{(n,2)}\tilde{b} r_2\right)} \left(1-\rho^{(n,2)}\tilde{b}\right) - \\
&\frac{\rho^{(n,2)}}{\tilde{b}}\left[\tilde{a}\left(T\left(\xi^{(l)}, \rho^{(n,2)}\right)\right) - T\left(\xi^{(l)}, \rho^{(n,2)}\right) \frac{d\tilde{a}}{dT}\left(\xi^{(l)}, \rho^{(n,2)}\right)\right] U\left(\rho^{(n,2)}\right) - \\
&\tilde{a}_{22} \Delta t \nabla\cdot\left[\left(e\left(\xi^{(l)}, \rho^{(n,2)}\right) + \frac{\xi^{(l)}}{\rho^{(n,2)}}\right) \mathbf{m}^{(n,2)}\right].
\end{aligned} \tag{43}$$

The same considerations as in [17] apply concerning the favorable properties of the weakly nonlinear system resulting from the discrete form of (42). Once the iterations have been completed, one sets $\mathbf{u}^{(n,2)} = \mathbf{u}^{(n,2,L+1)}$ and $E^{(n,2)}$ accordingly. For the third stage, one proceeds in an analogous way, according to the scheme prescribed in Table 2.

4.2. Discretization of viscous terms

Let us consider now the diffusive part of the Navier-Stokes equations that, as already mentioned in Section 1, will be treated with an operator splitting technique. For the sake of clarity, we denote with $\sim$ the quantities computed in this part of the scheme; hence, we define $\tilde{\mathbf{u}}^{(n,1)}$ and $\tilde{E}^{(n,1)}$ as the quantities obtained by applying (38) and we proceed to the discretization of the viscous terms, which is carried out by the implicit part of the IMEX method previously described:

$$\begin{aligned}
&\rho^{n+1}\tilde{\mathbf{u}}^{(n,2)} - \tilde{a}_{22}\frac{\Delta t}{Re}\nabla\cdot\left[\left(\nabla\tilde{\mathbf{u}} + \nabla\tilde{\mathbf{u}}^T\right) - \frac{2}{3}(\nabla\cdot\tilde{\mathbf{u}})\mathbf{I}\right]^{(n,2)} = \tilde{\mathbf{m}}^{(n,2)} \\
&\rho^{n+1}\tilde{E}^{(n,2)} - \tilde{a}_{22}\frac{\Delta t M^2}{Re}\nabla\cdot\left[\left(\nabla\tilde{\mathbf{u}} + \nabla\tilde{\mathbf{u}}^T\right)\tilde{\mathbf{u}} - \frac{2}{3}(\nabla\cdot\tilde{\mathbf{u}})\tilde{\mathbf{u}}\right]^{(n,2)} \\
&\qquad - \tilde{a}_{22}\frac{\Delta t}{PrRe}\Delta\tilde{T}^{(n,2)} = \tilde{e}^{(n,2)},
\end{aligned} \tag{44}$$

where we have set

$$\begin{aligned}
\tilde{\mathbf{m}}^{(n,2)} &= \rho^{n+1}\tilde{\mathbf{u}}^{(n,1)} + \tilde{a}_{21}\frac{\Delta t}{Re}\nabla\cdot\left[\left(\nabla\tilde{\mathbf{u}} + \nabla\tilde{\mathbf{u}}^T\right) - \frac{2}{3}(\nabla\cdot\tilde{\mathbf{u}})\mathbf{I}\right]^{(n,1)} \\
\tilde{e}^{(n,2)} &= \rho^{n+1}\tilde{E}^{(n,1)} \\
&\quad + \tilde{a}_{21}\frac{\Delta t M^2}{Re}\nabla\cdot\left[\left(\nabla\tilde{\mathbf{u}} + \nabla\tilde{\mathbf{u}}^T\right)\tilde{\mathbf{u}} - \frac{2}{3}(\nabla\cdot\tilde{\mathbf{u}})\tilde{\mathbf{u}}\right]^{(n,1)} \\
&\quad + \tilde{a}_{21}\frac{\Delta t}{PrRe}\Delta\tilde{T}^{(n,1)}.
\end{aligned}$$

Notice that the momentum equation in (44) is decoupled from the energy equation and can be solved independently, so that in a subsequent step the equation for $\tilde{E}^{(n,2)}$ can be solved using temperature as an unknown. It is worth to mention that, in case $\frac{d\tilde{a}}{dT} \neq 0$ or $\frac{d\tilde{c}_v}{dT} \neq 0$, for the cubic EOS, we end up with a non-linear equation. The following fixed point procedure is considered: setting $\xi^{(0)} = \tilde{T}^{(n,1)}$, one solves for $l = 1, \ldots, L$

$$\tilde{c}_v\left(\xi^{(l)}\right)\xi^{(l+1)} + \frac{\tilde{a}\left(\xi^{(l)}\right) - \xi^{(l+1)}\frac{d\tilde{a}}{dT}\left(\xi^{(l)}\right)}{\tilde{b}}$$

$$-\tilde{a}_{22}\frac{\Delta t M^2}{Re}\nabla\cdot\left[\left(\nabla\tilde{\mathbf{u}}+\nabla\tilde{\mathbf{u}}^T\right)\tilde{\mathbf{u}}-\frac{2}{3}(\nabla\cdot\tilde{\mathbf{u}})\tilde{\mathbf{u}}\right]^{(n,2)}$$
$$-\tilde{a}_{22}\frac{\Delta t}{PrRe}\Delta\tilde{\xi}^{(l+1)}=\tilde{e}^{(n,2)}. \tag{45}$$

Again, the third stage can be expressed in a similar manner, according to the scheme in Table 2. Finally, since we are considering a stiffly accurate method for the discretization of the viscous terms and, therefore, the output of the last stage is actually equal to the update solution [12], one sets

$$\mathbf{u}^{n+1}=\tilde{\mathbf{u}}^{(n,3)}\quad E^{n+1}=\tilde{E}^{(n,3)}$$

and the computation of the n-th time step is completed.

5. The spatial discretization strategy

We consider a decomposition of the domain Ω into a family of hexahedra $\mathcal{T}_h$ (quadrilaterals in the two-dimensional case) and denote each element by K. The skeleton $\mathcal{E}$ denotes the set of all element faces and $\mathcal{E}=\mathcal{E}^I\cup\mathcal{E}^B$, where $\mathcal{E}^I$ is the subset of interior faces and $\mathcal{E}^B$ is the subset of boundary faces. Suitable jump and average operators can then be defined as customary for finite element discretizations. A face $\Gamma\in\mathcal{E}^I$ shares two elements that we denote by K^+ with outward unit normal $\mathbf{n}^+$ and K^- with outward unit normal $\mathbf{n}^-$, whereas for a face $\Gamma\in\mathcal{E}^B$ we denote by $\mathbf{n}$ the outward unit normal. For a scalar function φ the jump is defined as

$$[[\varphi]]=\varphi^+\mathbf{n}^++\varphi^-\mathbf{n}^-\quad\text{if }\Gamma\in\mathcal{E}^I\qquad[[\varphi]]=\varphi\mathbf{n}\quad\text{if }\Gamma\in\mathcal{E}^B.$$

The average is defined as

$$\{\{\varphi\}\}=\frac{1}{2}\left(\varphi^++\varphi^-\right)\quad\text{if }\Gamma\in\mathcal{E}^I\qquad\{\{\varphi\}\}=\varphi\quad\text{if }\Gamma\in\mathcal{E}^B.$$

Similar definitions apply for a vector function $\boldsymbol{\varphi}$:

$$[[\boldsymbol{\varphi}]]=\boldsymbol{\varphi}^+\cdot\mathbf{n}^++\boldsymbol{\varphi}^-\cdot\mathbf{n}^-\quad\text{if }\Gamma\in\mathcal{E}^I\qquad[[\boldsymbol{\varphi}]]=\boldsymbol{\varphi}\cdot\mathbf{n}\quad\text{if }\Gamma\in\mathcal{E}^B$$
$$\{\{\boldsymbol{\varphi}\}\}=\frac{1}{2}\left(\boldsymbol{\varphi}^++\boldsymbol{\varphi}^-\right)\quad\text{if }\Gamma\in\mathcal{E}^I\qquad\{\{\boldsymbol{\varphi}\}\}=\boldsymbol{\varphi}\quad\text{if }\Gamma\in\mathcal{E}^B.$$

For vector functions, it is also useful to define a tensor jump as:

$$\langle\langle\boldsymbol{\varphi}\rangle\rangle=\boldsymbol{\varphi}^+\otimes\mathbf{n}^++\boldsymbol{\varphi}^-\otimes\mathbf{n}^-\quad\text{if }\Gamma\in\mathcal{E}^I\qquad\langle\langle\boldsymbol{\varphi}\rangle\rangle=\boldsymbol{\varphi}\otimes\mathbf{n}\quad\text{if }\Gamma\in\mathcal{E}^B.$$

We also introduce the following finite element spaces

$$Q_r=\left\{v\in L^2(\Omega):v|_K\in\mathbb{Q}_r\quad\forall K\in\mathcal{T}_h\right\}\quad\mathbf{V}_r=[Q_r]^d,$$

where $\mathbb{Q}_r$ is the space of polynomials of degree r in each coordinate direction. We then denote by $\boldsymbol{\varphi}_i(\mathbf{x})$ the basis functions for the space $\mathbf{V}_r$ and by $\psi_i(\mathbf{x})$ the basis functions for the space Q_r, the finite element spaces chosen for the discretization of the velocity and of the pressure (as well as the density), respectively.

$$\mathbf{u}\approx\sum_{j=1}^{\dim(\mathbf{V}_r)}u_j(t)\boldsymbol{\varphi}_j(\mathbf{x})\qquad p\approx\sum_{j=1}^{\dim(Q_r)}p_j(t)\psi_j(\mathbf{x}).$$

The spatial discretization coincides with that described in [1] and implemented in the *deal.II* library, so that it does not introduce any particular novelty. The shape functions correspond to the products of Lagrange interpolation polynomials for the support points of $(r+1)$-order Gauss-Lobatto quadrature rule in each coordinate direction. Given these definitions, the weak formulation for the momentum equation of the second stage (39) reads as follows:

$$\sum_K\int_K\rho^{(n,2)}\mathbf{u}^{(n,2)}\cdot\mathbf{v}d\Omega-\sum_K\int_K\tilde{a}_{22}\frac{\Delta t}{M^2}p^{(n,2)}\nabla\cdot\mathbf{v}d\Omega+\sum_{\Gamma\in\mathcal{E}}\int_\Gamma\tilde{a}_{22}\frac{\Delta t}{M^2}\left\{\left\{p^{(n,2)}\right\}\right\}[[\mathbf{v}]]d\Sigma$$
$$=\sum_K\int_K\rho^n\mathbf{u}^n\cdot\mathbf{v}d\Omega-\sum_K\int_K a_{21}\frac{\Delta t}{Fr^2}\rho^n\mathbf{k}\cdot\mathbf{v}d\Omega$$
$$+\sum_K\int_K a_{21}\Delta t\left(\rho^n\mathbf{u}^n\otimes\mathbf{u}^n\right):\nabla\mathbf{v}d\Omega+\sum_K\int_K\tilde{a}_{21}\frac{\Delta t}{M^2}p^n\nabla\cdot\mathbf{v}d\Omega$$

$$
\begin{aligned}
&- \sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} a_{21}\Delta t \left\{\left\{\rho^n \mathbf{u}^n \otimes \mathbf{u}^n\right\}\right\} : \langle\langle \mathbf{v}\rangle\rangle \, d\Sigma - \sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} \tilde{a}_{21} \frac{\Delta t}{M^2} \left\{\left\{p^n\right\}\right\} [[\mathbf{v}]] d\Sigma \\
&- \sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} a_{21}\Delta t \frac{\lambda^{(n,1)}}{2} \left\langle\left\langle \rho^n \mathbf{u}^n \right\rangle\right\rangle : \langle\langle \mathbf{v}\rangle\rangle \, d\Sigma,
\end{aligned} \tag{46}
$$

where $\lambda^{(n,1)} = \max\left(\left|\mathbf{u}^{n^+}\cdot\mathbf{n}^+\right|, \left|\mathbf{u}^{n^-}\cdot\mathbf{n}^-\right|\right)$. One can notice that centered flux has been employed as numerical flux for the quantities defined implicitly, whereas an upwind flux has been used for the quantities computed explicitly. In view of the implicit coupling between the momentum and the energy equations, we need to derive the algebraic formulation of (46) in order to formally substitute the degrees of freedom of the velocity into the algebraic formulation of the energy equation. We take $\mathbf{v} = \boldsymbol{\varphi}_i$, $i = 1...\dim(\mathbf{V}_r)$ and we exploit the representation introduced above to obtain

$$
\begin{aligned}
&\sum_K \int_K \rho^{(n,2)} \sum_{j=1}^{\dim(\mathbf{V}_r)} u_j^{(n,2)} \boldsymbol{\varphi}_j \cdot \boldsymbol{\varphi}_i d\Omega - \sum_K \int_K \tilde{a}_{22} \frac{\Delta t}{M^2} \sum_{j=1}^{\dim(Q_r)} p_j^{(n,2)} \psi_j \nabla\cdot\boldsymbol{\varphi}_i d\Omega \\
+ &\sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} \tilde{a}_{22} \frac{\Delta t}{M^2} \sum_{j=1}^{\dim(Q_r)} p_j^{(n,2)} \left\{\left\{\psi_j\right\}\right\} \left[\left[\boldsymbol{\varphi}_i\right]\right] d\Sigma = \sum_K \int_K \rho^n \mathbf{u}^n \cdot \boldsymbol{\varphi}_i d\Omega - \sum_K \int_K a_{21} \frac{\Delta t}{Fr^2} \rho^n \mathbf{k}\cdot\boldsymbol{\varphi}_i d\Omega \\
+ &\sum_K \int_K a_{21}\Delta t \left(\rho^n \mathbf{u}^n \otimes \mathbf{u}^n\right) : \nabla\boldsymbol{\varphi}_i d\Omega + \sum_K \int_K \tilde{a}_{21} \frac{\Delta t}{M^2} p^n \nabla\cdot\boldsymbol{\varphi}_i d\Omega \\
- &\sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} a_{21}\Delta t \left\{\left\{\rho^n \mathbf{u}^n \otimes \mathbf{u}^n\right\}\right\} : \left\langle\left\langle \boldsymbol{\varphi}_i \right\rangle\right\rangle d\Sigma - \sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} \tilde{a}_{21} \frac{\Delta t}{M^2} \left\{\left\{p^n\right\}\right\} \left[\left[\boldsymbol{\varphi}_i\right]\right] d\Sigma \\
- &\sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} a_{21}\Delta t \frac{\lambda^{(n,1)}}{2} \left\langle\left\langle \rho^n \mathbf{u}^n \right\rangle\right\rangle : \left\langle\left\langle \boldsymbol{\varphi}_i \right\rangle\right\rangle d\Sigma,
\end{aligned} \tag{47}
$$

which can be written in compact form as $\mathbf{A}^{(n,2)}\mathbf{U}^{(n,2)} + \mathbf{B}^{(n,2)}\mathbf{P}^{(n,2)} = \mathbf{F}^{(n,2)}$, where we have set

$$
A_{ij}^{(n,2)} = \sum_K \int_K \rho^{(n,2)} \boldsymbol{\varphi}_j \cdot \boldsymbol{\varphi}_i d\Omega \tag{48}
$$

$$
B_{ij}^{(n,2)} = \sum_K \int_K -\tilde{a}_{22} \frac{\Delta t}{M^2} \nabla\cdot\boldsymbol{\varphi}_i \psi_j d\Omega + \sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} \tilde{a}_{22} \frac{\Delta t}{M^2} \left\{\left\{\psi_j\right\}\right\} \left[\left[\boldsymbol{\varphi}_i\right]\right] d\Sigma \tag{49}
$$

with $\mathbf{U}^{(n,2)}$ denoting the vector of the degrees of freedom associated to the velocity field and $\mathbf{P}^{(n,2)}$ denoting the vector of the degrees of freedom associated to the pressure. Consider now the weak formulation for the energy equation of the second stage (39)

$$
\begin{aligned}
&\sum_K \int_K \rho^{(n,2)} E^{(n,2)} w d\Omega - \sum_K \int_K \tilde{a}_{22}\Delta t h^{(n,2)} \rho^{(n,2)} \mathbf{u}^{(n,2)} \cdot \nabla w d\Omega \\
+ &\sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} \tilde{a}_{22}\Delta t \left\{\left\{h^{(n,2)} \rho^{(n,2)} \mathbf{u}^{(n,2)}\right\}\right\} \cdot [[w]] d\Sigma = \sum_K \int_K \rho^n E^n w d\Omega - \sum_K \int_K a_{21} \frac{\Delta t M^2}{Fr^2} \rho^n \mathbf{k}\cdot\mathbf{u}^n w d\Omega \\
+ &\sum_K \int_K a_{21}\Delta t M^2 \left(k^n \rho^n \mathbf{u}^n\right) \cdot \nabla w d\Omega + \sum_K \int_K \tilde{a}_{21}\Delta t \left(h^n \rho^n \mathbf{u}^n\right) \cdot \nabla w d\Omega \\
- &\sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} a_{21}\Delta t M^2 \left\{\left\{k^n \rho^n \mathbf{u}^n\right\}\right\} \cdot [[w]] d\Sigma - \sum_{\Gamma\in\mathcal{E}^I} \int_{\Gamma} \tilde{a}_{21}\Delta t \left\{\left\{h^n \rho^n \mathbf{u}^n\right\}\right\} \cdot [[w]] d\Sigma \\
- &\sum_{\Gamma\in\mathcal{E}} \int_{\Gamma} a_{21}\Delta t \frac{\lambda^{(n,1)}}{2} \left[\left[\rho^n E^n\right]\right] \cdot [[w]] d\Sigma.
\end{aligned} \tag{50}
$$

Notice that, while the fully discrete formulation is presented here for the case of an ideal gas, in the more general case it has to be modified properly as already shown in (43) for the semi discrete formulation. Take $w = \psi_i$ and consider the expansion for $\mathbf{u}^{(n,2)}$ in (50) to get

$$
\begin{aligned}
&\sum_K \int_K \rho^{(n,2)} E^{(n,2)} \psi_i d\Omega - \sum_K \int_K \tilde{a}_{22} \Delta t h^{(n,2)} \rho^{(n,2)} \sum_{j=1}^{\dim(\mathbf{V}_r)} u_j^{(n,2)} \boldsymbol{\varphi}_j \cdot \nabla \psi_i d\Omega \\
+ \quad & \sum_{\Gamma \in \mathcal{E}} \int_\Gamma \tilde{a}_{22} \Delta t \sum_{j=1}^{\dim(\mathbf{V}_r)} u_j^{(n,2)} \left\{ \left\{ h^{(n,2)} \rho^{(n,2)} \boldsymbol{\varphi}_j \right\} \right\} \cdot [[\psi_i]] d\Sigma \\
= \quad & \sum_K \int_K \rho^n E^n \psi_i d\Omega - \sum_K \int_K a_{21} \frac{\Delta t M^2}{Fr^2} \rho^n \mathbf{k} \cdot \mathbf{u}^n \psi_i d\Omega \\
+ \quad & \sum_K \int_K a_{21} \Delta t M^2 \left(k^n \rho^n \mathbf{u}^n \right) \cdot \nabla \psi_i d\Omega + \sum_K \int_K \tilde{a}_{21} \Delta t \left(h^n \rho^n \mathbf{u}^n \right) \cdot \nabla \psi_i d\Omega \\
- \quad & \sum_{\Gamma \in \mathcal{E}} \int_\Gamma a_{21} \Delta t M^2 \left\{ \left\{ k^n \rho^n \mathbf{u}^n \right\} \right\} \cdot [[\psi_i]] d\Sigma - \sum_{\Gamma \in \mathcal{E}} \int_\Gamma \tilde{a}_{21} \Delta t \left\{ \left\{ h^n \rho^n \mathbf{u}^n \right\} \right\} \cdot [[\psi_i]] d\Sigma \\
- \quad & \sum_{\Gamma \in \mathcal{E}} \int_\Gamma a_{21} \Delta t \frac{\lambda^{(n,1)}}{2} \left[\left[\rho^n E^n \right] \right] \cdot [[\psi_i]] d\Sigma,
\end{aligned} \tag{51}
$$

which can be expressed in compact form as $\mathbf{C}^{(n,2)} \mathbf{U}^{(n,2)} = \mathbf{G}^{(n,2)}$, where we have set

$$
C_{ij}^{(n,2)} = \sum_K \int_K -\tilde{a}_{22} \Delta t h^{(n,2)} \rho^{(n,2)} \boldsymbol{\varphi}_j \cdot \nabla \psi_i d\Omega + \sum_{\Gamma \in \mathcal{E}} \int_\Gamma \tilde{a}_{22} \Delta t \left\{ \left\{ h^{(n,2)} \rho^{(n,2)} \boldsymbol{\varphi}_j \right\} \right\} \cdot [[\psi_i]] d\Sigma.
$$

Formally we can then derive $\mathbf{U}^{(n,2)} = (\mathbf{A}^{(n,2)})^{-1} \left(\mathbf{F}^{(n,2)} - \mathbf{B}^{(n,2)} \mathbf{P}^{(n,2)} \right)$ and obtain the following relation

$$
\mathbf{C}^{(n,2)} (\mathbf{A}^{(n,2)})^{-1} \left(\mathbf{F}^{(n,2)} - \mathbf{B}^{(n,2)} \mathbf{P}^{(n,2)} \right) = \mathbf{G}^{(n,2)}. \tag{52}
$$

Taking into account that $\rho^{(n,2)} E^{(n,2)} = \rho^{(n,2)} e^{(n,2)} (p^{(n,2)}) + M^2 \rho^{(n,2)} k^{(n,2)}$, we finally obtain

$$
\mathbf{C}^{(n,2)} (\mathbf{A}^{(n,2)})^{-1} \left(\mathbf{F}^{(n,2)} - \mathbf{B}^{(n,2)} \mathbf{P}^{(n,2)} \right) = -\mathbf{D}^{(n,2)} \mathbf{P}^{(n,2)} + \tilde{\mathbf{G}}^{(n,2)} \tag{53}
$$

where we have set

$$
D_{ij}^{(n,2)} = \sum_K \int_K \rho^{(n,2)} e^{(n,2)} (\psi_j) \psi_i d\Omega \tag{54}
$$

and $\tilde{\mathbf{G}}^{(n,2)}$ takes into account all the other terms (the one at previous stage and the kinetic energy). The above system can be solved in terms of $\mathbf{P}^{(n,2)}$ according to the fixed point procedure described in [17]. More specifically, setting $\mathbf{P}^{(n,2,0)} = \mathbf{P}^{(n,1)}, k^{(n,2,0)} = k^{(n,1)}$, for $l = 1, \ldots, L$ one solves the equation

$$
\left(\mathbf{D}^{(n,2,l)} - \mathbf{C}^{(n,2,l)} (\mathbf{A}^{(n,2)})^{-1} \mathbf{B}^{(n,2)} \right) \mathbf{P}^{(n,2,l+1)} = \tilde{\mathbf{G}}^{(n,2,l)} - \mathbf{C}^{(n,2,l)} (\mathbf{A}^{(n,2)})^{-1} \mathbf{F}^{(n,2,l)} \tag{55}
$$

and updates the velocity solving $\mathbf{A}^{(n,2)} \mathbf{U}^{(n,2,l+1)} = \mathbf{F}^{(n,2,l)} - \mathbf{B}^{(n,2)} \mathbf{P}^{(n,2,l+1)}$. The third stage can be described in a similar manner. One sets then

$$
\rho^{n+1} = \rho^{(n,3)} \quad \tilde{\mathbf{u}}^{(n,1)} = \mathbf{u}^{(n,3)} \quad \tilde{E}^{(n,1)} = E^{(n,3)}
$$

and proceeds to the implicit discretization of the viscous terms, which is carried out by the implicit part of the IMEX method described above through a Symmetric Interior Penalty (SIP) approach which was introduced in [2] (see also [44]). After integrating by parts on each mesh element, two stabilization terms are then added: a symmetrizing term corresponding to fluxes obtained after integration by parts and a penalty term imposing the weak continuity of the numerical solution [44]. More in detail, following [19], we set for each face Γ of a cell K

$$
\sigma_{\Gamma,K} = (r+1)^2 \frac{\operatorname{diam}(\Gamma)}{\operatorname{diam}(K)}
$$

and we define the penalization constant of the SIP method as

$$
\overline{C} = \frac{1}{2} \left(\sigma_{\Gamma,K^+} + \sigma_{\Gamma,K^-} \right)
$$

if $\Gamma \in \mathcal{E}^I$ and $\overline{C} = \sigma_{\Gamma,K}$ if $\Gamma \in \mathcal{E}^B$, where we remind that r denotes the polynomial degree of the finite element spaces. For the sake of completeness, we report here the weak formulation for the bilinear form of the momentum balance in (44):

$$
\begin{aligned}
B\left(\tilde{\mathbf{u}}, \mathbf{v}\right) = & \sum_K \int_K \rho^{n+1} \tilde{\mathbf{u}} \cdot \mathbf{v} d\Omega + \tilde{a}_{22} \frac{\Delta t}{Re} \sum_K \int_K \left[\nabla \tilde{\mathbf{u}} + \nabla \tilde{\mathbf{u}}^T - \frac{2}{3} \left(\nabla \cdot \tilde{\mathbf{u}}\right) \mathbf{I} \right] : \nabla \mathbf{v} d\Omega \\
& - \tilde{a}_{22} \frac{\Delta t}{Re} \sum_{\Gamma \in \mathcal{E}} \int_\Gamma \left\{ \left\{ \nabla \tilde{\mathbf{u}} + \nabla \tilde{\mathbf{u}}^T - \frac{2}{3} \left(\nabla \cdot \tilde{\mathbf{u}}\right) \mathbf{I} \right\} \right\} : \langle \langle \mathbf{v} \rangle \rangle \, d\Sigma \\
& - \tilde{a}_{22} \frac{\Delta t}{Re} \sum_{\Gamma \in \mathcal{E}} \int_\Gamma \langle \langle \tilde{\mathbf{u}} \rangle \rangle : \left\{ \left\{ \nabla \mathbf{v} + \nabla \mathbf{v}^T - \frac{2}{3} \left(\nabla \cdot \mathbf{v}\right) \mathbf{I} \right\} \right\} d\Sigma + \tilde{a}_{22} \frac{\Delta t}{Re} \sum_{\Gamma \in \mathcal{E}} \int_\Gamma \overline{C} \langle \langle \tilde{\mathbf{u}} \rangle \rangle : \langle \langle \mathbf{v} \rangle \rangle \, d\Sigma.
\end{aligned} \tag{56}
$$

The remaining formulations are obtained in an analogous manner. We would like to stress that the method outlined above does not require to introduce reference solutions, does not introduce inconsistencies in the splitting and only requires the solution of linear systems of a size equal to that of the number of discrete degrees of freedom needed to describe a scalar variable, as in [17]. This contrasts with other low Mach approaches based on IMEX methods, such as e.g. the technique proposed for the Euler equations in [57].

6. Numerical tests

The numerical scheme outlined in the previous Sections has been validated in a number of benchmarks. We set $\mathcal{H} = \min\{\text{diam}(K) | K \in \mathcal{T}_h\}$ and we define two Courant numbers, one based on the speed of sound denoted by C, the so-called acoustic Courant number, and one based on the local velocity of the flow, the so-called advective Courant number, denoted by C_u:

$$
C = \frac{1}{M} rc \Delta t / \mathcal{H}, \quad C_u = ru \Delta t / \mathcal{H}, \tag{57}
$$

where c is the magnitude of the speed of sound and u is the magnitude of the flow velocity. Notice that the definitions in (57) depend on the polynomial degree r. The factor $\frac{1}{M}$ is due to the scaling of the speed of sound, as reported in [41], for an ideal gas, and proven in Appendix B in the one-dimensional case for a general equation of state. For the tests using the ideal gas law, the value $\gamma = 1.4$ for the specific heat ratio is employed, unless differently stated. The fixed point loops are stopped at the l-th iteration such that the relative difference for the pressure is below 10^{-10}, namely

$$
\frac{\left\| \xi^{(l)} - \xi^{(l-1)} \right\|_\infty}{\left\| \xi^{(l)} \right\|_\infty} < 10^{-10}.
$$

6.1. Isentropic vortex

As a first benchmark, we consider for an ideal gas the two dimensional inviscid isentropic vortex also studied in [54,57]. For this test, an analytic solution is available, that can be used to assess the convergence properties of the scheme. The initial conditions are given as a perturbation of a reference state

$$
\rho(\mathbf{x}, 0) = \rho_\infty + \delta\rho \qquad \mathbf{u}(\mathbf{x}, 0) = \mathbf{u}_\infty + \delta\mathbf{u} \qquad p(\mathbf{x}, 0) = p_\infty + \delta p
$$

and the exact solution is a propagation of the initial condition at the background velocity

$$
\rho(\mathbf{x}, t) = \rho(\mathbf{x} - \mathbf{u}_\infty t, 0) \qquad \mathbf{u}(\mathbf{x}, t) = \mathbf{u}(\mathbf{x} - \mathbf{u}_\infty t, 0) \qquad p(\mathbf{x}, t) = p(\mathbf{x} - \mathbf{u}_\infty t, 0).
$$

The typical perturbation is defined as

$$
\widetilde{\delta T} = \frac{1 - \gamma}{8 \gamma \pi^2} \tilde{\beta}^2 e^{1 - \tilde{r}^2}, \tag{58}
$$

with $\tilde{r}^2 = (x - x_0)^2 + (y - y_0)^2$ denoting the radial coordinate and $\tilde{\beta}$ being the vortex strength. As explained in [57], however, in order to emphasize the role of the Mach number M, we define

$$
\delta T = \frac{1 - \gamma}{8 \gamma \pi^2} M^2 \tilde{\beta}^2 e^{1 - \tilde{r}^2} \tag{59}
$$

and we set

$$
\rho(\mathbf{x}, 0) = (1 + \delta T)^{\frac{1}{\gamma - 1}} \quad p(\mathbf{x}, 0) = M^2 (1 + \delta T)^{\frac{\gamma}{\gamma - 1}}. \tag{60}
$$

Table 3
Convergence test for the inviscid isentropic vortex at $C \approx 0.01$, $C_u \approx 0.01$ with $r = 1$ and $a_{32} = \frac{7-2\chi}{6}$ for the explicit part. Relative errors for the density, the velocity and the pressure in L^2 norm. N_{el} denotes the number of elements along each direction.

N_{el}	L^2 rel. error ρ	L^2 rate ρ	L^2 rel. error $\mathbf{u}$	L^2 rate $\mathbf{u}$	L^2 rel. error p	L^2 rate p
10	$2.00 \cdot 10^{-3}$		$1.19 \cdot 10^{-2}$		$2.79 \cdot 10^{-3}$	
20	$7.86 \cdot 10^{-4}$	1.35	$3.86 \cdot 10^{-3}$	1.62	$1.11 \cdot 10^{-3}$	1.33
40	$2.55 \cdot 10^{-4}$	1.62	$1.07 \cdot 10^{-3}$	1.84	$3.61 \cdot 10^{-4}$	1.62
80	$7.15 \cdot 10^{-5}$	1.83	$2.67 \cdot 10^{-4}$	2.00	$1.01 \cdot 10^{-4}$	1.84

Table 4
Convergence test for the inviscid isentropic vortex at $C \approx 0.01$, $C_u \approx 0.01$ with $r = 2$ and $a_{32} = \frac{7-2\chi}{6}$ for the explicit part. Relative errors for the density, the velocity and the pressure in L^2 norm. N_{el} denotes the number of elements along each direction.

N_{el}	L^2 rel. error ρ	L^2 rate ρ	L^2 rel. error $\mathbf{u}$	L^2 rate $\mathbf{u}$	L^2 rel. error p	L^2 rate p
10	$6.38 \cdot 10^{-4}$		$2.61 \cdot 10^{-3}$		$9.08 \cdot 10^{-4}$	
20	$1.18 \cdot 10^{-4}$	2.43	$3.54 \cdot 10^{-4}$	2.88	$1.64 \cdot 10^{-4}$	2.47
40	$1.81 \cdot 10^{-5}$	2.70	$4.16 \cdot 10^{-5}$	3.09	$2.53 \cdot 10^{-5}$	2.70
80	$2.96 \cdot 10^{-6}$	2.61	$5.18 \cdot 10^{-6}$	3.00	$4.13 \cdot 10^{-6}$	2.60

Table 5
Convergence test for the inviscid isentropic vortex at $C \approx 0.01$, $C_u \approx 0.01$ with $r = 1$ and $a_{32} = 0.5$ for the explicit part. Relative errors for the density, the velocity and the pressure in L^2 norm. N_{el} denotes the number of elements along each direction.

N_{el}	L^2 rel. error ρ	L^2 rate ρ	L^2 rel. error $\mathbf{u}$	L^2 rate $\mathbf{u}$	L^2 rel. error p	L^2 rate p
10	$2.00 \cdot 10^{-3}$		$1.19 \cdot 10^{-2}$		$2.79 \cdot 10^{-3}$	
20	$7.86 \cdot 10^{-4}$	1.35	$3.86 \cdot 10^{-3}$	1.62	$1.11 \cdot 10^{-3}$	1.33
40	$2.55 \cdot 10^{-4}$	1.62	$1.07 \cdot 10^{-3}$	1.85	$3.61 \cdot 10^{-4}$	1.62
80	$7.15 \cdot 10^{-5}$	1.83	$2.67 \cdot 10^{-4}$	2.00	$1.00 \cdot 10^{-4}$	1.85

For what concerns the velocity the typical perturbation is defined as

$$\widetilde{\delta \mathbf{u}} = \tilde{\beta} \begin{pmatrix} -(y - y_0) \\ (x - x_0) \end{pmatrix} \frac{e^{\frac{1}{2}(1-r^2)}}{2\pi} \tag{61}$$

where x_0 and y_0 are the coordinates of the vortex center and also in this case we rescale it using M

$$\delta \mathbf{u} = \tilde{\beta} M \begin{pmatrix} -(y - y_0) \\ (x - x_0) \end{pmatrix} \frac{e^{\frac{1}{2}(1-\tilde{r}^2)}}{2\pi} \tag{62}$$

We apply the same reasoning also to the background velocity and therefore we define $\mathbf{u}_\infty = M\tilde{\mathbf{u}}_\infty$ with $\tilde{\mathbf{u}}_\infty = [10, 10]^T$. To avoid problems related to the definition of boundary conditions, we choose a sufficiently large domain $\Omega = (-10, 10)^2$ and periodic boundary conditions and we set $\rho_\infty = 1$, $p_\infty = 1$, $x_0 = y_0 = 0$, $\tilde{\beta} = 10$, the final time $T_f = 1$ and $M = 0.1$. Notice that we refrain from investigating the properties of the method in the very low Mach number regime for this test, since this entails an almost constant solution. The numerical experiments have been carried out on Cartesian meshes of square elements with N_{el} elements in each coordinate direction, choosing for each spatial resolution time steps so that the Courant numbers remained constant (hyperbolic scaling).

We first consider the original IMEX-ARK scheme with $a_{32} = \frac{7-2\chi}{6}$ for the explicit part. We observe that, in general, convergence rates of at least $r + \frac{1}{2}$ are observed for $r = 1$ and for $r = 2$ as reported in Tables 3 and 4.

Analogous results are shown in Tables 5 and 6 for the modified scheme with $a_{32} = 0.5$, chosen, as discussed in Appendix A, in order to increase the region of absolute monotonicity without affecting too much stability.

In further numerical experiments, we have observed that the lack of absolute monotonicity strongly affects the computation of density and, as a consequence, the stability of the whole numerical scheme. For Courant number around $C \approx 0.3$ the original method becomes unstable, while the modified scheme with $a_{32} = 0.5$ is still able to recover the expected convergence rates at least in the $r = 1$ case, as evident from Table 7, while for $r = 2$ reported in Table 8 we observe a small degradation of the convergence rates due to increasing influence of the dominant second order time discretization error. In order to be able to run at slightly longer time steps we have then chosen to use the $a_{32} = 0.5$ value for the IMEX scheme for the rest of the numerical simulations carried out in this paper. We notice also that, for both schemes, the results compare well with the analogous results presented in [54] and with those obtained in [57] with a higher order IMEX method.

For validation purposes, we have also tested in this case the h-adaptive version of the method. The local refinement criterion is based on the gradient of the density. More specifically, we define for each element K the quantity

Table 6
Convergence test for the inviscid isentropic vortex at $C \approx 0.01$, $C_u \approx 0.01$ with $r = 2$ and $a_{32} = 0.5$ for the explicit part. Relative errors for the density, the velocity and the pressure in L^2 norm. N_{el} denotes the number of elements along each direction.

N_{el}	L^2 rel. error ρ	L^2 rate ρ	L^2 rel. error $\mathbf{u}$	L^2 rate $\mathbf{u}$	L^2 rel. error p	L^2 rate p
10	$6.38 \cdot 10^{-4}$		$2.61 \cdot 10^{-3}$		$9.08 \cdot 10^{-4}$	
20	$1.18 \cdot 10^{-4}$	2.43	$3.54 \cdot 10^{-4}$	2.88	$1.64 \cdot 10^{-4}$	2.47
40	$1.81 \cdot 10^{-5}$	2.70	$4.16 \cdot 10^{-5}$	3.09	$2.53 \cdot 10^{-5}$	2.70
80	$2.96 \cdot 10^{-6}$	2.61	$5.18 \cdot 10^{-6}$	3.00	$4.13 \cdot 10^{-6}$	2.60

Table 7
Convergence test for the inviscid isentropic vortex at $C \approx 0.3$, $C_u \approx 0.3$ with $r = 1$ and $a_{32} = 0.5$ for the explicit part. Relative errors for the density, the velocity and the pressure in L^2 norm. N_{el} denotes the number of elements along each direction.

N_{el}	L^2 rel. error ρ	L^2 rate ρ	L^2 rel. error $\mathbf{u}$	L^2 rate $\mathbf{u}$	L^2 rel. error p	L^2 rate p
10	$2.32 \cdot 10^{-3}$		$1.19 \cdot 10^{-2}$		$2.78 \cdot 10^{-3}$	
20	$7.63 \cdot 10^{-4}$	1.60	$3.88 \cdot 10^{-3}$	1.62	$1.06 \cdot 10^{-3}$	1.39
40	$2.43 \cdot 10^{-4}$	1.65	$1.08 \cdot 10^{-3}$	1.85	$3.41 \cdot 10^{-4}$	1.64
80	$6.84 \cdot 10^{-5}$	1.83	$2.69 \cdot 10^{-4}$	2.01	$9.55 \cdot 10^{-5}$	1.84

Table 8
Convergence test for the inviscid isentropic vortex at $C \approx 0.3$, $C_u \approx 0.3$ with $r = 2$ and $a_{32} = 0.5$ for the explicit part. Relative errors for the density, the velocity and the pressure in L^2 norm. N_{el} denotes the number of elements along each direction.

N_{el}	L^2 rel. error ρ	L^2 rate ρ	L^2 rel. error $\mathbf{u}$	L^2 rate $\mathbf{u}$	L^2 rel. error p	L^2 rate p
10	$6.43 \cdot 10^{-4}$		$2.71 \cdot 10^{-3}$		$9.15 \cdot 10^{-4}$	
20	$1.28 \cdot 10^{-4}$	2.33	$3.89 \cdot 10^{-4}$	2.80	$1.68 \cdot 10^{-4}$	2.44
40	$2.10 \cdot 10^{-5}$	2.61	$5.78 \cdot 10^{-5}$	2.75	$2.70 \cdot 10^{-5}$	2.64
80	$4.08 \cdot 10^{-6}$	2.36	$1.13 \cdot 10^{-5}$	2.35	$4.97 \cdot 10^{-6}$	2.44

Table 9
Adaptive simulations of the inviscid isentropic vortex at different resolutions with a maximum $C \approx 0.3$, $C_u \approx 0.3$, relative errors for the density, the velocity and the pressure in L^2 norm with $r = 1$. N_{cells} denotes the total number of cells.

N_{cells}	L^2 rel. error ρ	L^2 rel. error $\mathbf{u}$	L^2 rel. error p
241	$2.04 \cdot 10^{-3}$	$1.19 \cdot 10^{-2}$	$2.78 \cdot 10^{-3}$
541	$7.31 \cdot 10^{-4}$	$3.50 \cdot 10^{-3}$	$1.03 \cdot 10^{-3}$
1951	$2.09 \cdot 10^{-4}$	$9.23 \cdot 10^{-4}$	$2.92 \cdot 10^{-4}$
7537	$6.07 \cdot 10^{-5}$	$2.42 \cdot 10^{-4}$	$8.51 \cdot 10^{-5}$

$$\eta_K = \max_{i \in \mathcal{N}_K} |\nabla \rho|_i \tag{63}$$

that acts as local refinement indicator, where $\mathcal{N}_K$ denotes the set of nodes over the element K. Table 9 shows the relative errors for all the quantities on a sequence of adaptive simulations keeping the maximum Courant numbers fixed. The expected results are obtained, even though the relative errors are not significantly reduced with respect to Table 7 in view of the smoothness of the solution. Hence, the following results have to merely intended as a verification of the correctness of the h-adaptive version of the scheme. Fig. 1 shows instead the density and the adapted mesh at $t = T_f$, from which it can be seen that the refinement criterion is able to track the vortex correctly.

6.2. 2D Lid-driven cavity

We consider now the classical 2D lid-driven cavity test case. The computational domain is the box $\Omega = (0, 1) \times (0, 1)$ which is initialized with a density $\rho = 1$ and a velocity $\mathbf{u} = \mathbf{0}$. The flow is driven by the upper boundary, whose velocity is set to $\mathbf{u} = (1, 0)^T$, while on the other three boundaries a no-slip condition is imposed. We set $Re = 100$ and $M^2 = 10^{-5}$ and we consider $r = 2$ as polynomial degree. The advantage of the proposed scheme is that the allowed time step is more than 100 times larger than that of a fully explicit scheme. Indeed, the time-step chosen is such that the maximum advective Courant number C_u is around 0.12, while the maximum Courant number C is around 49. The streamlines are shown in Fig. 2 and highlight the formation of the main recirculation pattern. A comparison of the horizontal component of the velocity along the vertical middle line and of the vertical component of the velocity along the horizontal middle line with the reference solutions in [23,54] is also presented and we note that our results fit very well both the reference solutions.

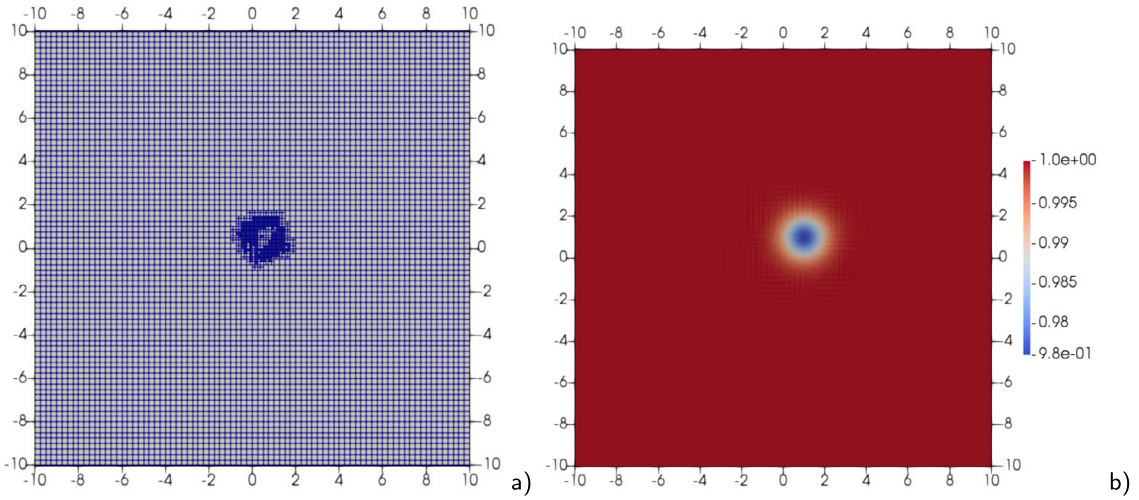

Fig. 1. Adaptive simulation of the inviscid isentropic vortex benchmark: a) computational mesh at $t = T_f$, b) contour plot of the density at $t = T_f$. (For interpretation of the colors in the figure(s), the reader is referred to the web version of this article.)

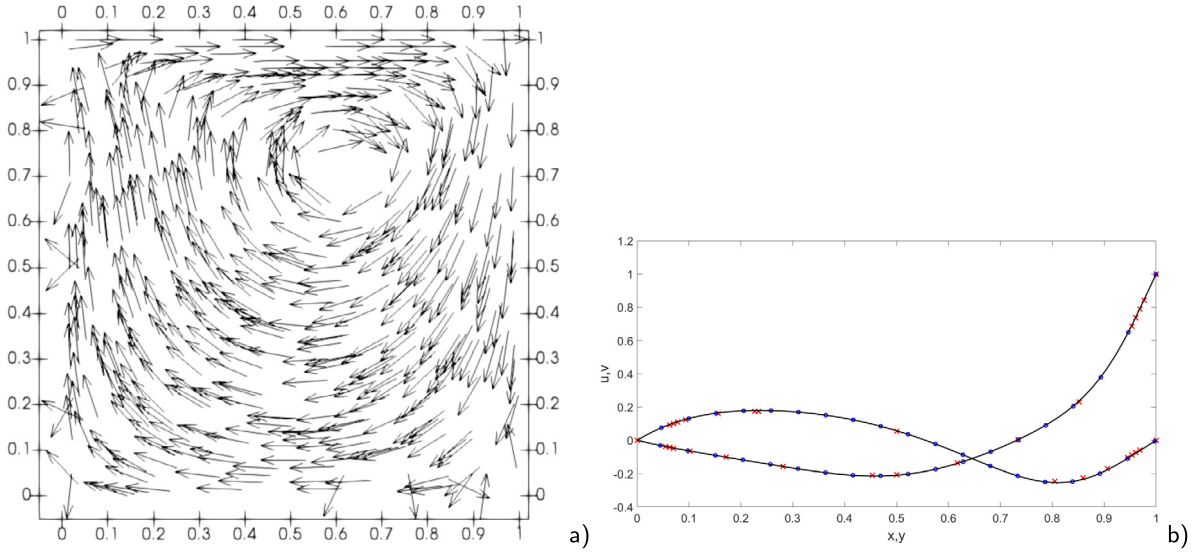

Fig. 2. Computational results for the 2D lid-driven cavity with $k = 2$, a) streamlines, b) comparison with the solutions in [23] and in [54]. Blue dots denote the results in [23], red crosses the results in [54] and black line our numerical results.

6.3. Sod shock tube problem

Even though the proposed method is particularly well suited for low Mach number flows, we have also tested its behavior in a situation in which shock waves occur. For this purpose, we have considered the classical Sod shock tube problem proposed in [50] and also discussed in [17]. It consists of a right-moving shock wave, an intermediate contact discontinuity and a left-moving rarefaction fan. In this higher Mach number regime, as done also in [17], for further stabilization in presence of shocks and discontinuities, the numerical flux employed for the quantities computed explicitly in the above weak formulations is the classical Local Lax-Friedrichs flux (LLF), instead of the upwind flux, defined by setting

$$\lambda^{(n,1)} = \max\left\{\left\|\mathbf{u}^{(n,1)^+}\right\| + \frac{c^{(n,1)^+}}{M}, \left\|\mathbf{u}^{(n,1)^-}\right\| + \frac{c^{(n,1)^-}}{M}\right\} \quad \lambda^{(n,2)} = \max\left\{\left\|\mathbf{u}^{(n,2)^+}\right\| + \frac{c^{(n,2)^+}}{M}, \left\|\mathbf{u}^{(n,2)^-}\right\| + \frac{c^{(n,2)^-}}{M}\right\}.$$

The presence of discontinuities requires the use of a monotonic scheme to avoid undershoots and overshoots. It is well known that using Q_0 finite elements in combination with LLF and an explicit time integration method that complies with the monotonicity constraints discussed in [21,30,29] guarantees the monotonicity of the solution. Hence, a way to obtain monotonic results is to project the numerical solution onto the Q_0 subspace for each element in which a suitable jump

Table 10
Initial left and right states for Sod shock tube problem. x_d denotes the position of the initial discontinuity.

ρ_L	u_L	p_L	ρ_R	u_R	p_R	x_d
1	0	1	0.125	0	0.1	0

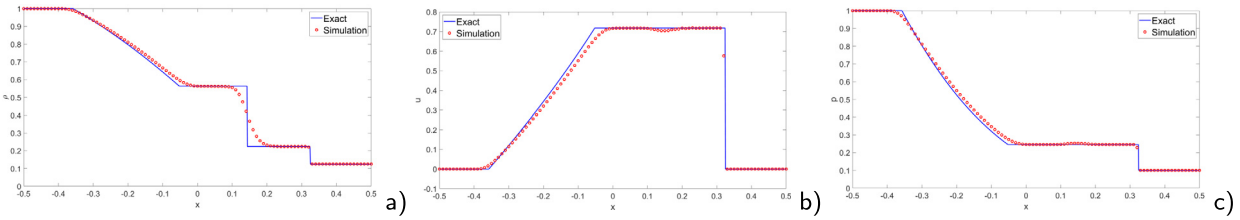

Fig. 3. Sod shock tube problem with van der Waals EOS at $t = 0.2$, comparison with exact solution, a) density, b) velocity, c) pressure.

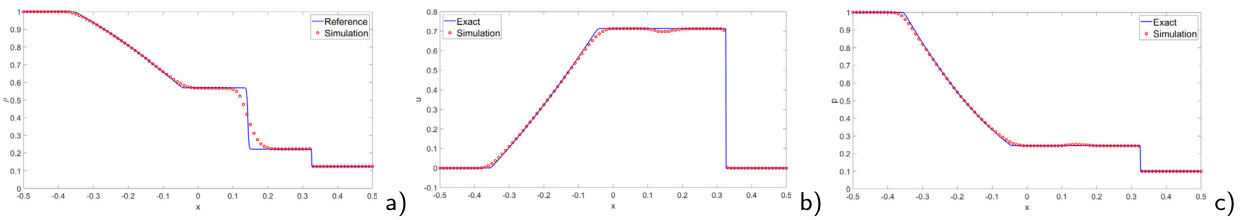

Fig. 4. Sod shock tube problem with Peng-Robinson EOS at $t = 0.2$, comparison with reference solution, a) density, b) velocity, c) pressure.

indicator exceeds a certain threshold. Similar projections onto low order components of the solution are also used in several monotonization approaches, see e.g. [18]. However, since in the proposed scheme only the density is treated in a full explicit fashion, in order to avoid an excessive complication in the structure of the resulting method we choose to apply this Q_0 projection strategy only for the density variable, without introducing monotonization for the velocity and the pressure. While we are aware that this is not sufficient to guarantee full monotonicity, the derivation of a fully monotonic IMEX scheme goes beyond the scope of this work and we do not investigate this issue further here. Therefore, the results in this Section are to be interpreted merely as a first stress test of the proposed scheme at higher Mach number values. In future work, we plan to investigate the behavior of the monotonization approach proposed in [43] for full explicit schemes, when applied only to the density. We use a smoothness indicator based on the jump of the density across two faces. More in detail, we define for each element K the quantity

$$\eta_K = \sum_{\Gamma \in \mathcal{E}_K} \left\| \rho^+ - \rho^- \right\|^2_{2,\Gamma}$$

where $\mathcal{E}_K$ denotes the set of all faces belonging to cell K and $\|\cdot\|$ represents the standard L^2 norm on Γ. The chosen threshold in this case is equal to 10^{-8}. Table 10 defines the initial conditions and position of the initial discontinuity. We consider the domain $\Omega = (-0.5, 0.5)$ and a one-dimensional mesh composed by 500 elements with a time step $\Delta t = 10^{-4}$, chosen in such a way that the maximum Courant number is $C \approx 0.09$, while the maximum advective Courant number is $C_u \approx 0.06$. Following [17], we have first considered the van der Waals EOS with $\tilde{a} = \tilde{b} = 0.5$. Fig. 3 shows the results for the density, the velocity and the pressure at $t = 0.2$ compared with the exact solution. One can easily notice that the shocks are located at the right position and that the values in the wake of the shocks are correct. The technique applied to guarantee the monotonicity introduces however an excessive amount of numerical diffusion in correspondence of contact wave, which is not sharply resolved, in contrast to what happens for the shock and the rarefaction waves. On the other hand, if one decreases the value of the threshold, an oscillating solution is obtained. While far from optimal, these results highlight however the robustness of the proposed approach also in the higher Mach number case.

Finally, we have analyzed the same test case for the Peng-Robinson EOS with $\tilde{a} = \tilde{b} = 0.5$. In this case no analytic solution is available and a reference solution is computed using the third order optimal SSP Runge-Kutta method derived in [28] in combination with Q_0 finite elements and LLF as numerical flux on a mesh with 16000 elements. Notice that, in the case of the van der Waals EOS, this procedure was found to provide a solution overlapping with the exact one. The good agreement between our numerical results and the reference ones is established also for this equation of state, as evident from Fig. 4, again with significant smoothing of the contact discontinuity.

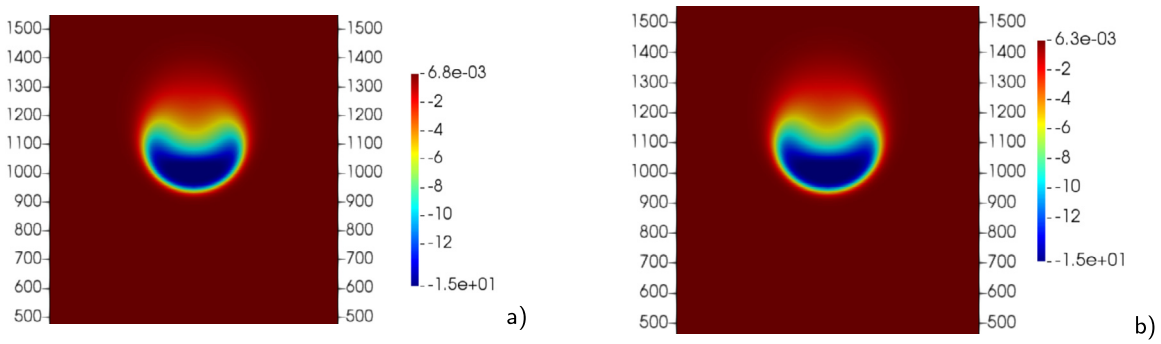

Fig. 5. Cold bubble test case, results at $t = T_f$, a) contour plot of potential temperature perturbation for the reference explicit simulation, b) contour plot of the potential temperature perturbation for the simulation with IMEX scheme.

6.4. Cold bubble

In this Section, we consider a test case proposed in [46,47] for an ideal gas in which the gravity force is active. The computational domain is the rectangle $(0, 1000) \times (0, 2000)$ and the initial condition is represented by a thermal anomaly introduced in an isentropic background atmosphere with constant potential temperature $\theta_0 = 303$. The perturbation of potential temperature θ' defines the initial datum and it is given by

$$\theta' = \begin{cases} A & \text{if } \tilde{r} \le r_0 \\ A \exp\left(-\frac{(\tilde{r}-r_0)^2}{\sigma^2}\right) & \text{if } \tilde{r} > r_0, \end{cases} \tag{64}$$

with $\tilde{r}^2 = (x - x_0)^2 + (y - y_0)^2$ and $x_0 = 500$, $y_0 = 1250$, $r_0 = 50$, $\sigma = 100$ and $A = -15$. Moreover, we set $Fr^2 = \frac{1}{9.81}$, $M^2 = 10^{-5}$ and $T_f = 50$. The expression of the Exner pressure is given by

$$\Pi = 1 - \frac{M^2}{Fr^2} \frac{y}{\tilde{c}_p \theta},$$

with y denoting the vertical coordinate and $\tilde{c}_p = \frac{\gamma}{\gamma - 1} \tilde{R}_g = 1.0045 \cdot 10^{-2}$ denoting the non-dimensional specific heat at constant pressure. Notice that these values are obtained by considering $\mathcal{R} = 1\ \mathrm{kg\,m^{-3}}$, $\Theta = 1$ K and $\mathcal{P} = 10^5$ Pa. Moreover, it is to be remarked that, unlike in [46], no artificial viscosity has been added to stabilize the computation. Wall boundary conditions are imposed at all the boundaries. The time step is taken to be $\Delta t = 0.08$, corresponding to a maximum Courant number $C \approx 5.6$ and a maximum advective Courant number $C_u \approx 0.18$.

For the purpose of a quantitative comparison, a reference solution is computed with an explicit time discretization given by the optimal third order SSP scheme mentioned in Section 6.3. Fig. 5 shows the contour plot of the potential temperature perturbation at $t = T_f$ and one can easily notice that we are able to recover correctly the shape of the reference solution. For a more quantitative point of view, the profile of the density at $y = 1000$ is reported in Fig. 6 and a good agreement between the reference results and those obtained with the IMEX scheme is established. The IMEX scheme allows to employ a time step 40 times larger compared to the fully explicit scheme with a computational saving of around 90%. Three fixed-point iterations were required on average for each IMEX stage.

In order to further enhance the computational efficiency, we employ again the code h-adaptivity capabilities. As mentioned in Section 3.1, we use as refinement indicator the gradient of the potential temperature, since this quantity allows to identify the cold bubble. More specifically, we set

$$\eta_K = \max_{i \in \mathcal{N}_K} |\nabla \theta|_i \tag{65}$$

as local indicator, where $\mathcal{N}_K$ is defined as in (63), and we allow to refine when η_K exceeds 10^{-1} and to coarsen below $6 \cdot 10^{-2}$. The initial computational grid is composed by 50×100 elements and we allow up to two local refinements only, so as to keep the advective Courant number under control and to recover the same maximum resolution employed for the non adaptive mesh simulation. Notice that there is no intrinsic limitation in the maximum number of refinement levels allowed and more refinement levels will be indeed used in the following tests with non-ideal gases. The only constraint is about the necessity of not having neighboring cells with refinement levels differing by more than one. However, a maximum number of allowed local refinements has to be set depending on the chosen time step in order to fulfill the stability of the scheme. As one can easily notice from Fig. 7, the refinement criterion is able to track the bubble and the one-dimensional density profile at $y = 1000$ in Fig. 8 is correctly reproduced. The final mesh consists of 6890 elements instead of the 80000 elements of the full resolution mesh and a further 50% reduction in computational time is achieved. Three fixed-point iterations were

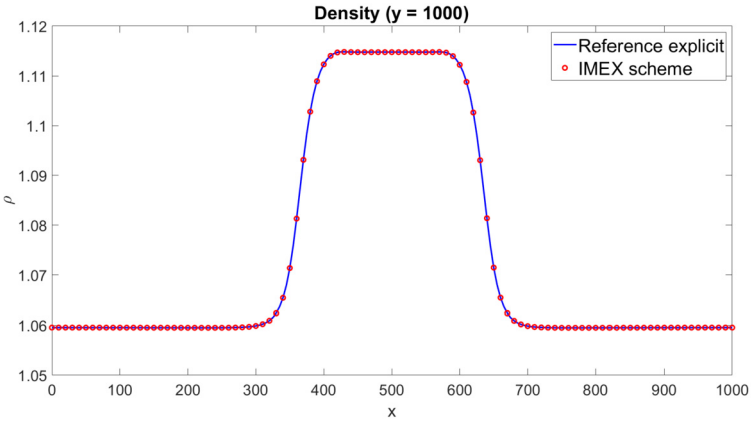

Fig. 6. Cold bubble test case, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the reference explicit simulation, whereas the red dots denote the results for the IMEX scheme.

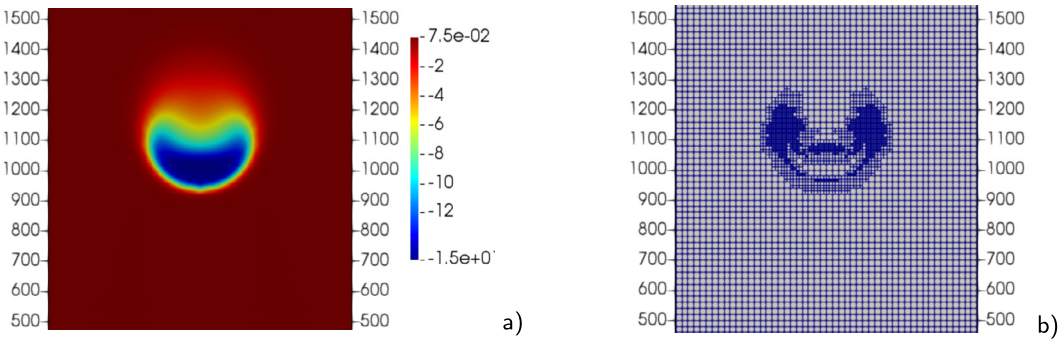

Fig. 7. Cold bubble test case, adaptive simulation, results at $t = T_f$, a) contour plot of potential temperature perturbation, b) computational grid.

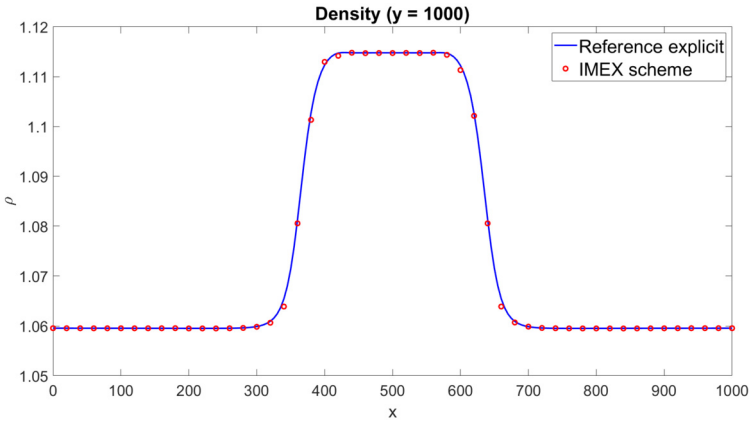

Fig. 8. Cold bubble test case, adaptive simulation, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the reference explicit simulation, whereas the red dots denote the results for the IMEX scheme.

required on average even with the h-adaptive version of the scheme and, therefore, no deterioration in the performances of the fixed-point loop occurred. We noticed instead an increase in the number of iterations required by the GMRES linear solver applied to (55) and to the corresponding third stage. The CPU time required for the mesh adaptation procedure represents less than 1% of the total CPU time.

We repeat now the same test using non-ideal equations of state. We first consider the van der Waals equation with a constant $\tilde{c}_v$ given by $\tilde{c}_v = \frac{\tilde{R}_g}{\gamma - 1} = 7.175 \cdot 10^{-3}$, $\tilde{a} = 5 \cdot 10^{-9}$ and $\tilde{b} = 5 \cdot 10^{-4}$, so that the same specific heat at constant volume with respect to the ideal gas case is obtained and $z \approx 1$. The fluid is initialized using the same pressure and the same density values as in the ideal gas case. Notice that $\frac{d\tilde{a}}{dT} = \frac{d\tilde{c}_v}{dT} = 0$, therefore it is not necessary to compute explicitly

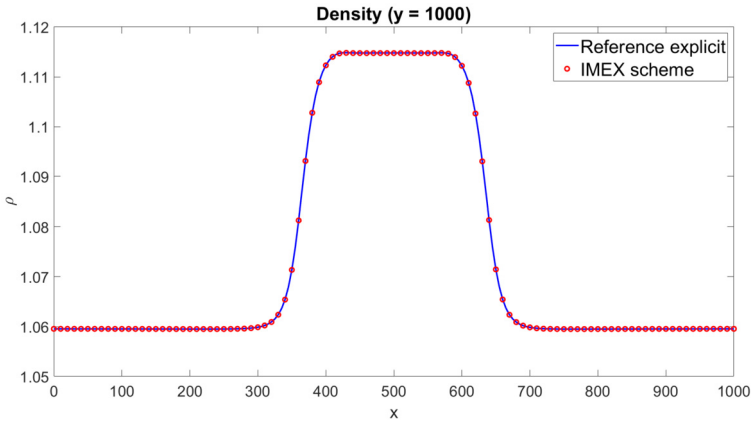

Fig. 9. Cold bubble test case, van der Waals EOS with $\tilde{a} = 5 \cdot 10^{-9}$ and $\tilde{b} = 5 \cdot 10^{-4}$, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the reference explicit simulation, whereas the red dots denote the results for the IMEX scheme.

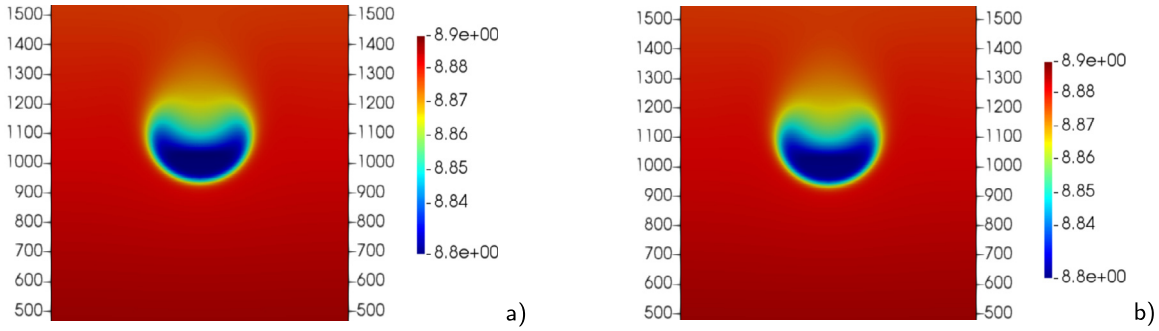

Fig. 10. Cold bubble test case, van der Waals EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, results at $t = T_f$, a) contour plot of β for the reference explicit simulation, b) contour plot of β for the simulation with IMEX scheme.

the temperature for (43) and for the corresponding third stage. We expect a behavior similar to that of the ideal gas, which is confirmed by the density profile reported in Fig. 9.

We then consider the case with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, which yield an average compressibility factor $z \approx 0.83$. In this case, we expect more significant effects due to conditions far from the ideal ones. We first compute a reference solution with the explicit time discretization. The time step for the IMEX simulation is kept equal to $\Delta t = 0.08$, yielding a maximum Courant number $C \approx 5.3$ and a maximum advective Courant number $C_u \approx 0.19$. Fig. 10 shows the contour plot for β at $t = T_f$ for both the reference explicit and the IMEX simulations. The expected behavior is retrieved and a good agreement with the reference results is established. Also in this case, a computational saving of around 90% with respect to the explicit simulation is obtained thanks to the IMEX scheme. Fig. 11 reports the profile of the density for $y = 1000$ at $t = T_f$. One can notice the very good agreement between the IMEX results and the reference ones. Furthermore, a clear discrepancy with respect to the ideal gas can be observed. The higher density values are due to the large value of $\tilde{a}$, which means that strong forces of attraction between the gas particles are present [42].

Concerning the adaptive simulations, since, as proven in Section 3, the quantity $\beta = \log(T) - 2\frac{\tilde{R}_g}{\tilde{c}_v}\operatorname{atanh}\left(2\rho\tilde{b} - 1\right)$ is constant in an isentropic process with $\frac{d\tilde{a}}{dT} = \frac{d\tilde{c}_v}{dT} = 0$, we define the local refinement indicator for each element as

$$\eta_K = \max_{i \in \mathcal{N}_K} |\nabla \beta|_i \,. \tag{66}$$

We allow to refine when η_K exceeds $4 \cdot 10^{-4}$ and to coarsen when the indicator is below $2 \cdot 10^{-4}$. The initial mesh is composed by 50×100 elements and we allow up to four local refinements. For this reason, in order to keep under control the advective Courant number, we need to reduce the time step $\Delta t = 0.02$, so as to obtain a maximum acoustic Courant number $C \approx 5.3$ and a maximum advective Courant number $C_u \approx 0.18$. Fig. 12 confirms that β is an appropriate quantity to track the bubble and the one-dimensional density profile in Fig. 13 shows that no significant loss in accuracy occurs. The final mesh consists of 8876 elements.

The same analysis is carried out using the Peng-Robinson EOS. Hence, we first consider $\tilde{R}_g = 2.87 \cdot 10^{-3}$, $\tilde{c}_v = 7.175 \cdot 10^{-3}$, $\tilde{a} = 5 \cdot 10^{-9}$ and $\tilde{b} = 5 \cdot 10^{-4}$, so that $z \approx 1$. The density profile reported in Fig. 14 highlights, as expected, a behavior entirely analogous to that of the ideal gas.

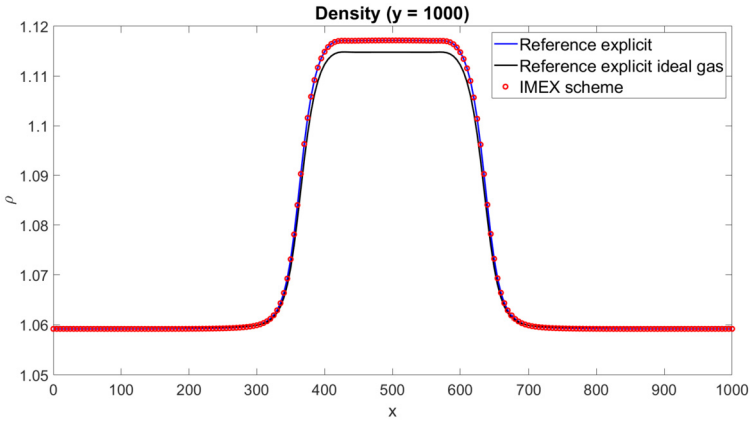

Fig. 11. Cold bubble test case, van der Waals EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the full explicit simulation, the continuous black line reports the results for the reference explicit simulation with an ideal gas, whereas the red dots denote the results for the IMEX scheme.

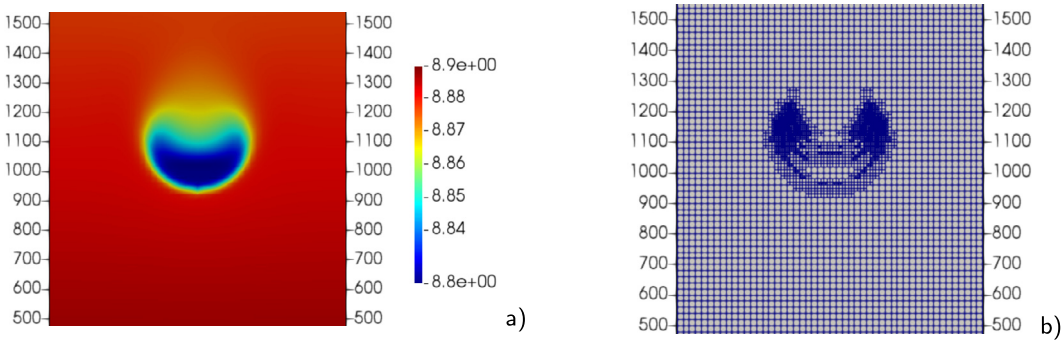

Fig. 12. Cold bubble test case, van der Waals EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, adaptive simulation, results at $t = T_f$, a) contour plot of β, b) computational grid.

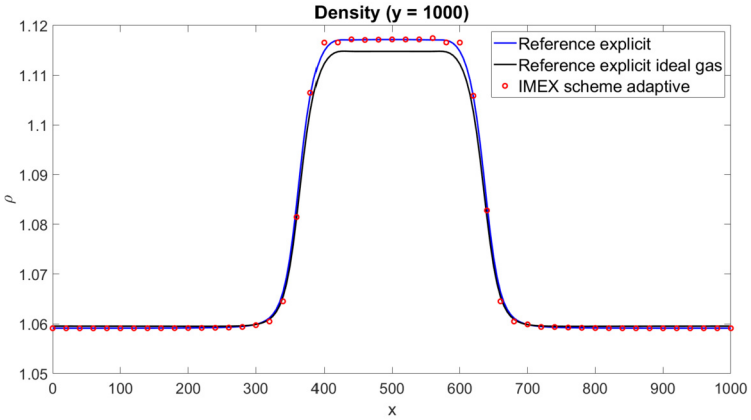

Fig. 13. Cold bubble test case, van der Waals EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, adaptive simulation, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the reference explicit simulation, the continuous black line reports the results for the reference explicit simulation with an ideal gas, whereas the red dots denote the results for the IMEX scheme in the non-ideal case.

Next, we take $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, so that $z \approx 0.83$, and we perform both uniform mesh and adaptive simulations, using the same parameters employed for the van der Waals EOS. The results are compared with a reference solution computed with the explicit method. Fig. 15 shows similar contour plots for all the configurations as well as for the adaptive mesh at $t = T_f$, which consists of 8888 elements and it is able to track the bubble correctly. Fig. 16 reports the comparison for the one-dimensional profile of the density at $y = 1000$ and the same considerations of the van der Waals EOS are still valid. We want to test in this case the refinement indicator based on (35). More specifically, we set

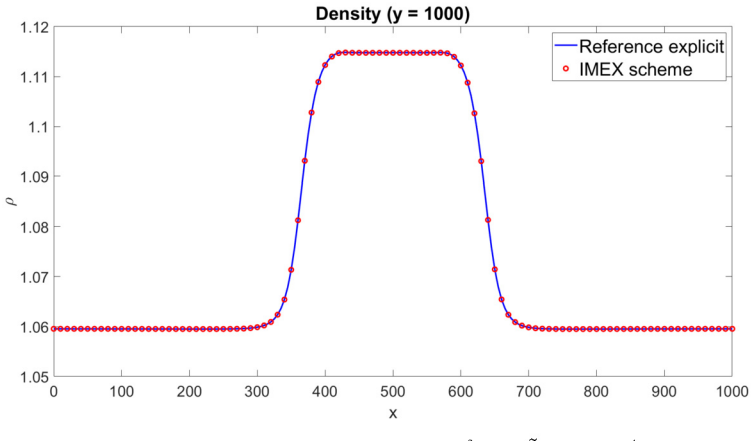

Fig. 14. Cold bubble test case, Peng-Robinson EOS with $\tilde{a} = 5 \cdot 10^{-9}$ and $\tilde{b} = 5 \cdot 10^{-4}$, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the reference explicit simulation, whereas the red dots denote the results for the IMEX scheme.

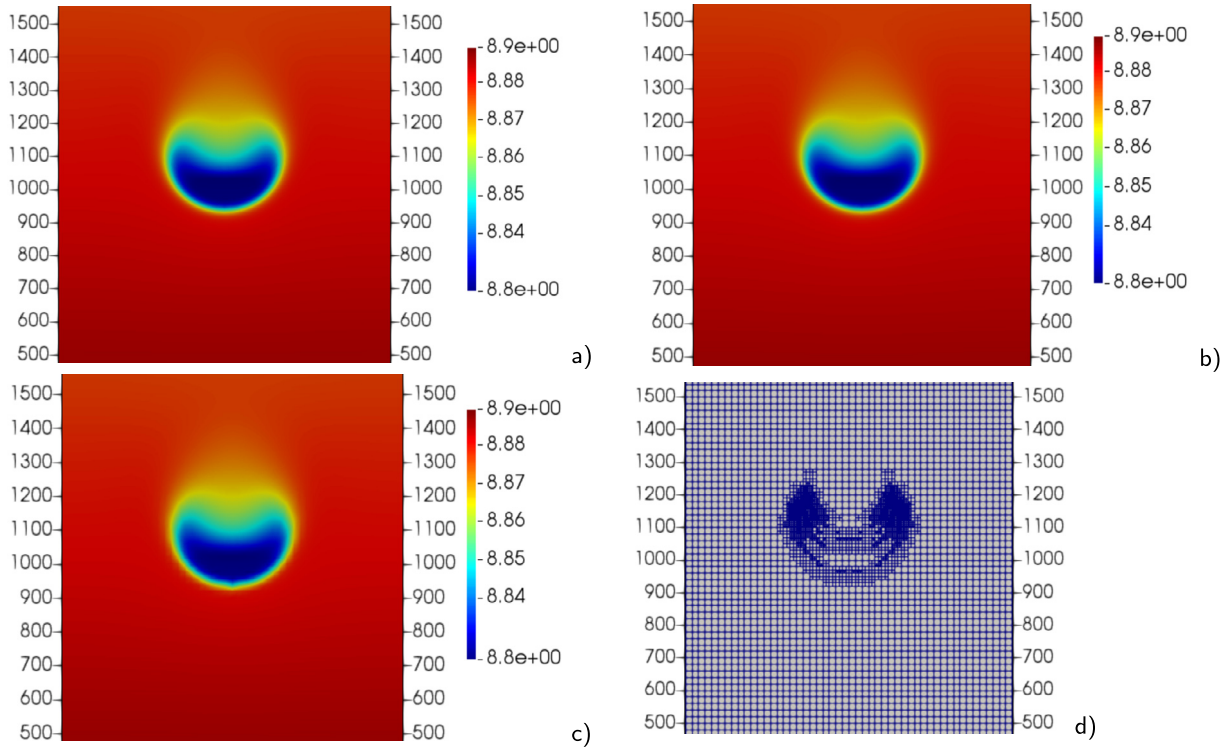

Fig. 15. Cold bubble test case, Peng-Robinson EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, results at $t = T_f$, a) contour plot of β for the reference explicit simulation, b) contour plot of β for the constant mesh simulation with IMEX scheme, c) contour plot of β for adaptive simulation with IMEX scheme, d) adaptive mesh.

$$\eta_K = \max_{i \in \mathcal{N}_K} \left| \nabla \left(\frac{p}{\rho^{\gamma_{p\rho}}} \right) \right|_i \tag{67}$$

and we allow to refine in case η_K is above $4 \cdot 10^{-4}$ and to coarsen below $2 \cdot 10^{-4}$ with the same remeshing procedure adopted so far for non-ideal gases. Fig. 17 shows the contour plot of (35) and the computational mesh at $t = T_f$. The mesh consists of 8168 elements and one can easily notice that more resolution is added only in correspondence of the bubble.

6.5. Warm bubble

In order to test the method also in presence of heat conduction, we now consider for an ideal gas the test case of a rising warm bubble proposed in [13]. The domain is the square box $\Omega = (-0.5, 1.5) \times (-0.5, 1.5)$ with periodic boundary

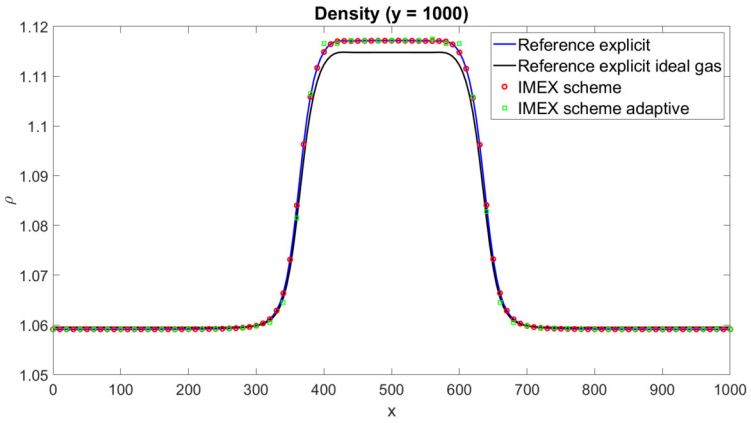

Fig. 16. Cold bubble test case, Peng-Robinson EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, adaptive simulation, results at $t = T_f$, density profile at $y = 1000$. The continuous blue line represents the results for the reference explicit simulation, the continuous black line reports the results for the reference explicit simulation with an ideal gas, the red dots denote the results for the IMEX scheme, whereas the green squares represent the results for the adaptive simulation with IMEX scheme.

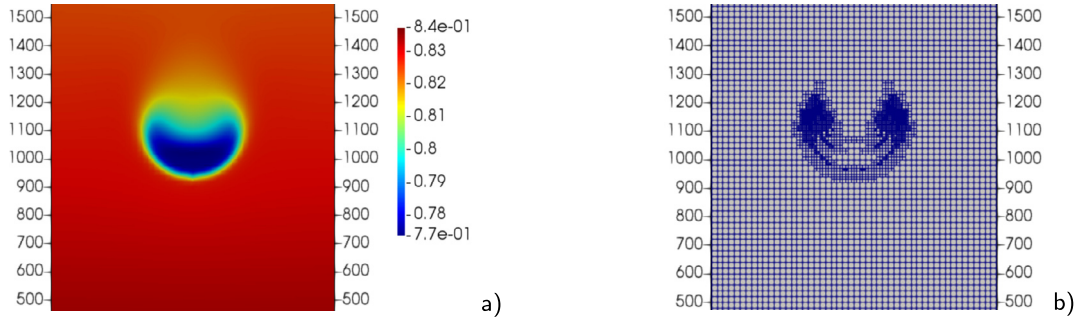

Fig. 17. Cold bubble test case, Peng-Robinson EOS with $\tilde{a} = 1.6 \cdot 10^{-1}$ and $\tilde{b} = 5 \cdot 10^{-4}$, adaptive simulation with criterion (67), results at $t = T_f$, a) contour plot of (35), b) adaptive mesh.

conditions on the lateral boundaries and wall boundary conditions on the top and on the bottom of the domain. The initial temperature corresponds to a truncated Gaussian profile

$$T(\mathbf{x}, 0) = \begin{cases} 386.48 & \text{if } \tilde{r} > r_0 \\ \dfrac{\tilde{p}_0}{\tilde{R}_{g,air}\left(1 - 0.1 e^{\frac{\tilde{r}^2}{\sigma^2}}\right)} & \text{if } \tilde{r} \leq r_0, \end{cases} \tag{68}$$

where $\tilde{r}^2 = (x - x_0)^2 + (y - y_0)^2$ is the distance from the center with coordinates $x_0 = 0.5$ and $y_0 = 0.35$, $r_0 = 0.25$ is the radius and $\sigma = 2$. In this Section, we consider unitary reference values for density, pressure and temperature and, therefore, we set $\tilde{p}_0 = 10^5$ and $\tilde{R}_{g,air} = 287$. Moreover, following [13], we consider:

$$Re = 804.9 \qquad Pr = 0.71 \qquad Fr \approx 0.004 \qquad M \approx 0.01.$$

The grid is composed by 120 elements along each direction and the time step is such that the maximum Courant number $C \approx 118$ and the maximum value of advective Courant number C_u is around 0.03. Fig. 18 shows the results at $t = 20$ s both in terms of contours and plots along the same specific sections along x-axis chosen in [13]. All the results are in good agreement with the reference ones and we are able to recover the development of the expected Kelvin-Helmholtz instability.

The same test is repeated using data for nitrous oxide (N_2O) from [40], which we report here for the convenience of the reader. At temperature of 386.48 K and pressure of 10^5 Pa, $\mu = 1.8884 \cdot 10^{-5}$ Pa·s and $\kappa = 2.4855 \cdot 10^{-2}$ W m^{-1} K^{-1}, so as to obtain

$$Re \approx 716.1 \qquad Pr \approx 0.73.$$

We consider the Peng-Robinson EOS, for which the expressions of $\tilde{a}(T)$ and $\tilde{b}$ are the following [20]:

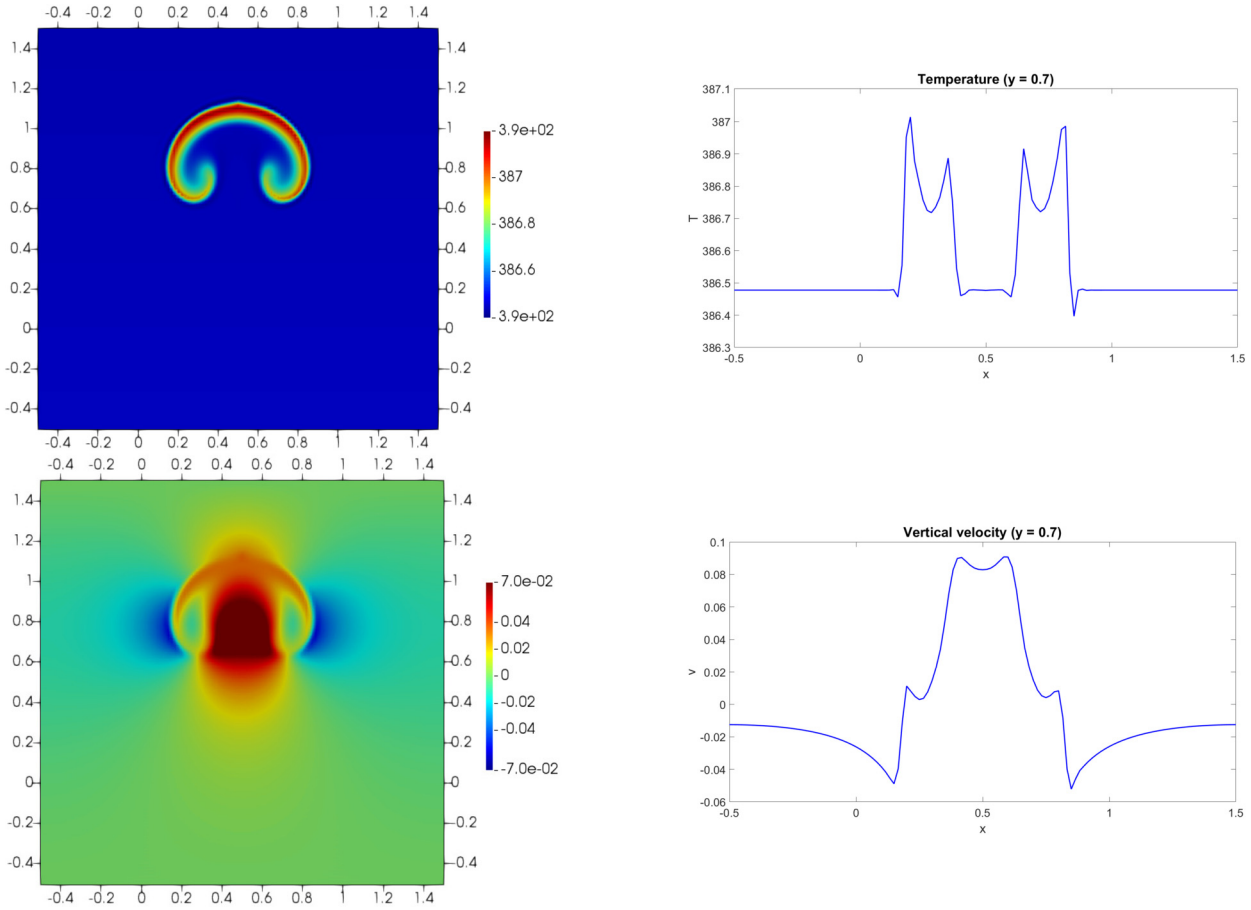

Fig. 18. Warm bubble test case, results at $t = 20$ s. From top to bottom: temperature and vertical velocity.

$$\begin{cases} \tilde{a}(T) & = 0.45724 \frac{\tilde{R}_g^2 \tilde{T}_c^2}{\tilde{p}_c} \alpha(T)^2 \\ \alpha(T) & = 1 + \Gamma \left(1 - \sqrt{\frac{T}{\tilde{T}_c}}\right) \\ \Gamma & = 0.37464 + 1.54226\omega - 0.26992\omega^2 \\ \tilde{b} & = 0.0778 \frac{\tilde{R}_g \tilde{T}_c}{\tilde{p}_c}, \end{cases} \tag{69}$$

where $\tilde{T}_c$ denotes the non-dimensional critical temperature, $\tilde{p}_c$ the non-dimensional critical pressure and ω the acentric factor. For what concerns N_2O, we find from [40] $\tilde{T}_c = 309.52$, $\tilde{p}_c = 7.2450 \cdot 10^6$ and $\omega = 0.1613$. Finally, the function $\tilde{c}_v(T)$ is computed using the following polynomial from [40]:

$$\tilde{c}_v(T) = \frac{1}{T}\left[\left[A\frac{T}{1000} + \frac{1}{2}B\left(\frac{T}{1000}\right)^2 + \frac{1}{3}C\left(\frac{T}{1000}\right)^3 + \frac{1}{4}D\left(\frac{T}{1000}\right)^4 - E\frac{1000}{T}\right]\frac{10^6}{M_w} - \tilde{R}_{g,N_2O}T\right], \tag{70}$$

with $\tilde{R}_{g,N_2O} = 188.91$, $M_w = 44.0128$ and A, B, C, D, E denoting suitable coefficients whose values are reported in Table 11. It is worthwhile to recall once more that $\tilde{c}_v(T)$ is not a proper specific heat at constant volume, but it denotes the non-dimensional counterpart of $\frac{e^{\#}(T)}{T}$ from (10), as shown in (16). The test is initialized with the same temperature and the same pressure already used for the ideal gas. The same mesh and the time step of the previous case are used, yielding to $C \approx 92$ and $C_u \approx 0.03$. Fig. 19 shows the temperature, the horizontal and the vertical velocity at $t = 20$ s. One can easily notice that a good qualitative agreement compared with the results in Fig. 18 is obtained. For a more quantitative point of view, since an explicit solution cannot be computed easily in view of the very large acoustic Courant number and considering that the compressibility factor is $z \approx 0.997$, a simulation with the ideal gas law (28) is performed, using $\gamma = 1.2879$, which corresponds to $\frac{\tilde{c}_v(386.48)}{\tilde{R}_g} + 1$, so that the internal energy of the ideal gas at $T = 386.48$ is the same as in the case $e^{\#}(386.48)$. The temperature profile at $y = 0.8$ shown in Fig. 20 confirms the good quality of the solution, with only slight differences due to the different equations of state.

Table 11
Values for polynomial (70).

A	B	C	D	E
27.67988	51.14898	−30.64544	6.847911	−0.157906

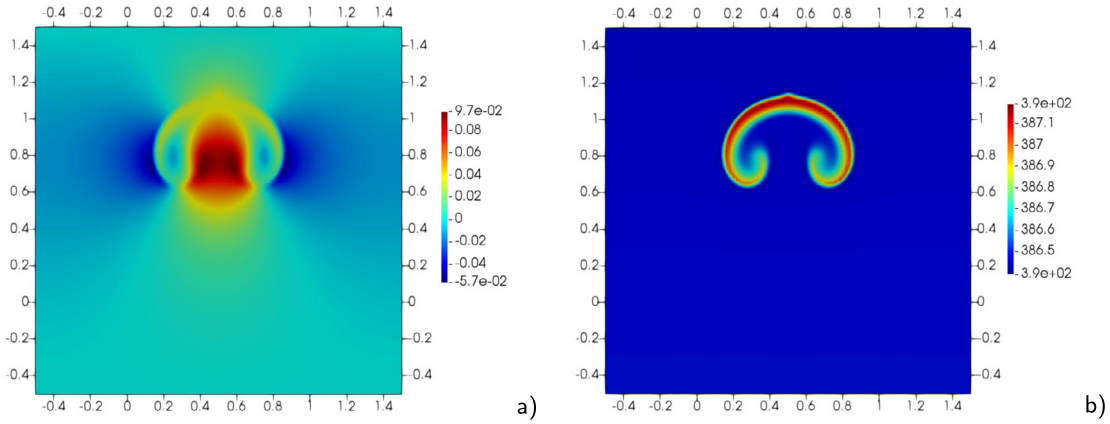

Fig. 19. Warm bubble test case for N_2O with Peng-Robinson EOS, results at $t = 20$ s, a) vertical velocity, b) temperature.

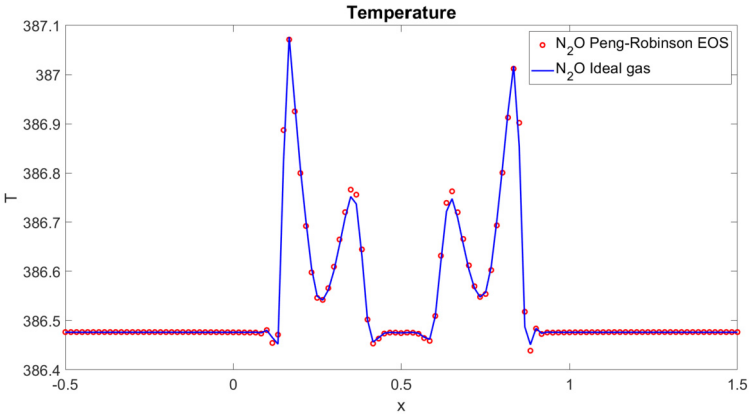

Fig. 20. Warm bubble test case for N_2O with Peng-Robinson EOS, temperature profile for $y = 0.8$ at $t = 20$ s.

In order to consider also non-ideal effects, we focus on the more challenging conditions closer to the vapor-liquid phase transition curve of N_2O. More in detail, we set as initial conditions $p = 4 \cdot 10^6$ and we consider the following temperature profile

$$T(\mathbf{x}, 0) = \begin{cases} 298 & \text{if } \tilde{r} > r_0 \\ \dfrac{\tilde{p}_0}{\tilde{R}_{g,air}\left(1-0.1e^{\frac{\tilde{r}^2}{\sigma^2}}\right)} - 88.48 & \text{if } \tilde{r} \leq r_0, \end{cases} \tag{71}$$

which corresponds to a translation with respect to (68), yielding $z \approx 0.72$. The maximum acoustic Courant is $C \approx 74.5$, whereas the maximum advective Courant number is $C_u \approx 0.06$. Fig. 21 shows the contour plots of the temperature at $t = 15$ s and $t = 20$ s. For these conditions of temperature and pressure, we obtain from [40] $\mu = 1.6680 \cdot 10^{-5}$ Pa·s, $\kappa = 2.1201 \cdot 10^{-2}$ $\mathrm{W\,m^{-1}\,K^{-1}}$ and $c_p = 1.5150 \cdot 10^3$ $\mathrm{J\,kg^{-1}\,K^{-1}}$, so that one has

$$Re \approx 810.7 \quad Pr \approx 1.19$$

One can easily notice the full development of the Kelvin-Helmholtz instability with the formation of secondary vortices and the fact that the bubble reaches a higher altitude with respect to the previous case. See also Fig. 22 for the vertical velocity.

Finally, we consider the SG-EOS with $\gamma = 1.0936$, $\tilde{c}_v = 1453.91$ and $\tilde{\pi}_\infty = \tilde{q}_\infty = 0$. The values for γ and $\tilde{c}_v$ are computed using the procedure described in [22]. The maximum acoustic Courant is $C \approx 67$, whereas the maximum advective Courant number is $C_u \approx 0.04$. Fig. 23 shows the contour plots of the temperature at $t = 15$ s and $t = 20$ s, whereas Fig. 24 shows

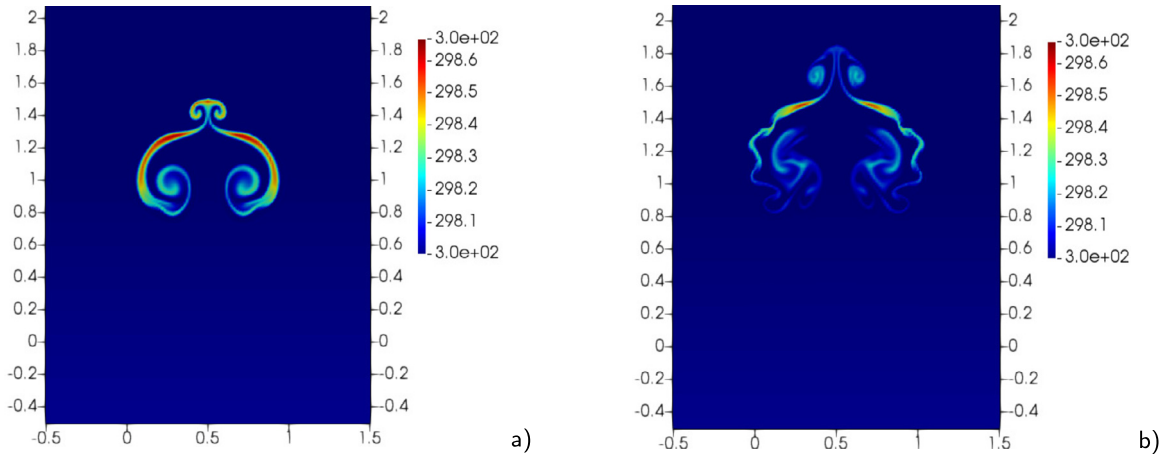

Fig. 21. Warm bubble test case for N_2O with Peng-Robinson EOS, a) temperature at $t = 15$ s, b) temperature at $t = 20$ s.

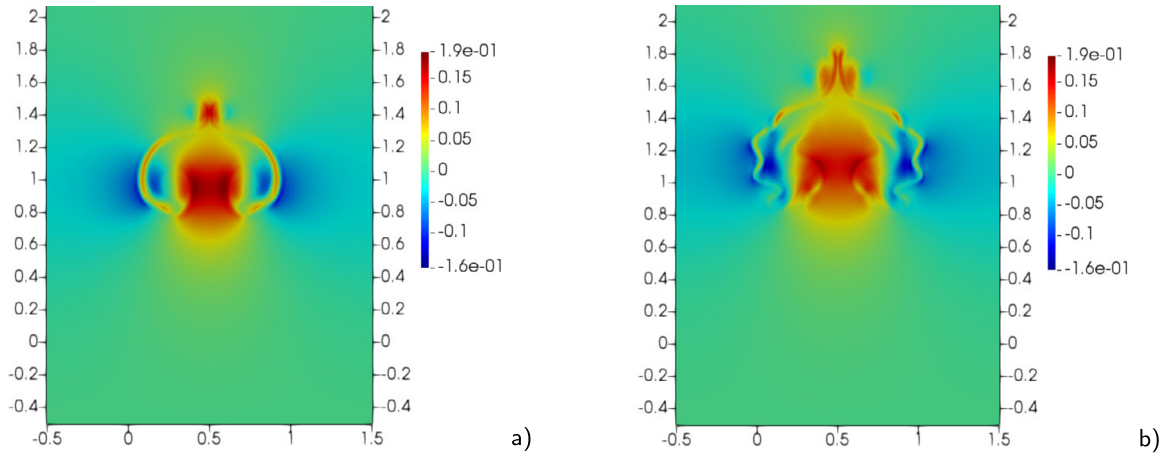

Fig. 22. Warm bubble test case for N_2O with Peng-Robinson EOS, a) vertical velocity at $t = 15$ s, b) vertical velocity at $t = 20$ s.

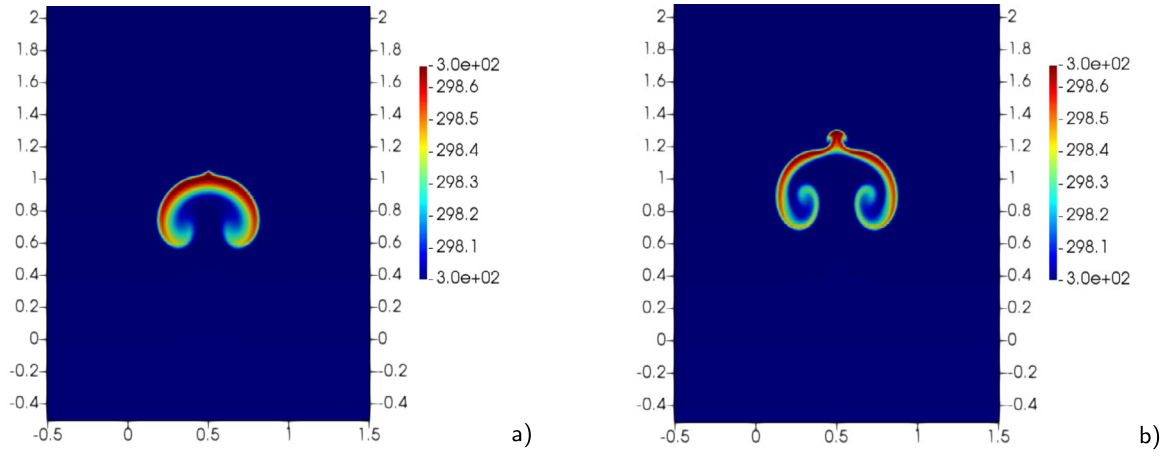

Fig. 23. Warm bubble test case for N_2O with SG-EOS, a) temperature at $t = 15$ s, b) temperature at $t = 20$ s.

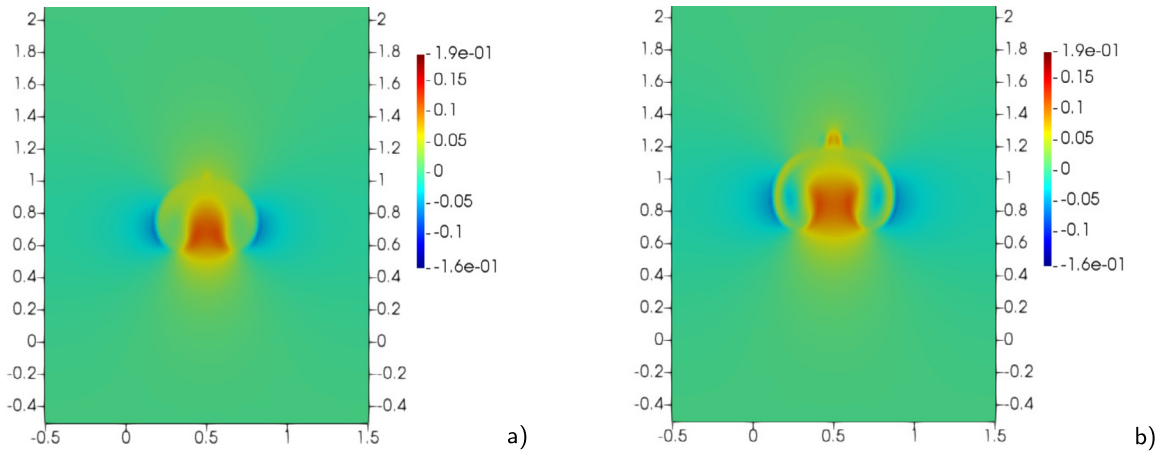

Fig. 24. Warm bubble test case for N_2O with SG-EOS, a) vertical velocity at $t = 15$ s, b) vertical velocity at $t = 20$ s.

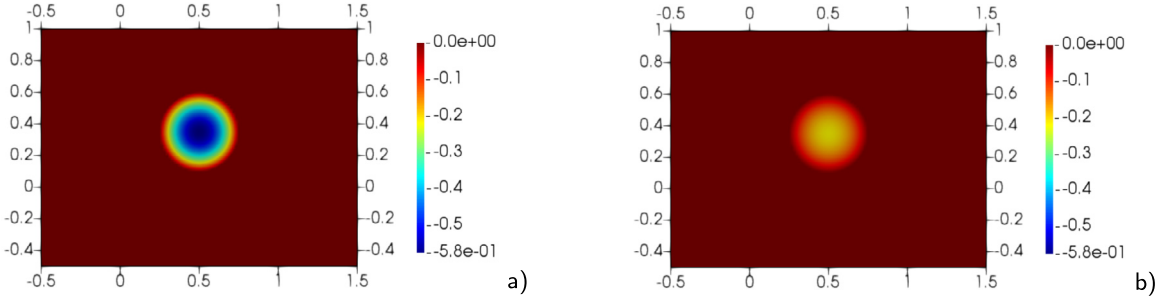

Fig. 25. Warm bubble test case for N_2O, density deviation from background state at $t = 0$, a) Peng-Robinson EOS, b) SG-EOS.

the contour plots of the vertical velocity. The different behavior between the two equations of state can be readily explained since, in the case of Peng-Robinson EOS, the difference between the density of the bubble and the background density is bigger with respect to SG-EOS, as evident from Fig. 25 and, therefore, a bigger upward buoyant force is exerted on the bubble. For this reason, in the simulation with Peng-Robinson EOS reaches a higher level compared to that in the SG-EOS simulation.

7. Conclusions and future perspectives

We have proposed an efficient, *h*-adaptive IMEX-DG solver for the compressible Navier-Stokes equations with non-ideal EOS and we have shown how to apply an effective implicit adaptive procedure also to the case of general cubic equations of state. The solver combines ideas from the discretization approaches in [15,17,26,36] and proposes an improvement in the choice of the free parameter employed by the explicit part of the IMEX scheme described in [26]. The resulting method is implemented in the framework of the numerical library *deal.II* and exploits its *h*-adaptive capabilities on the basis of physically based adaptation criteria that have been proposed specifically for the non-ideal gas case. A number of numerical experiments validate the proposed method and show its potential for low Mach number problems. In future work, we plan to extend the scheme to multiphase flows and to demonstrate its potential for application to atmospheric flows.

CRediT authorship contribution statement

Giuseppe Orlando: Data curation, Software, Validation, Writing – review & editing. **Paolo Francesco Barbante:** Conceptualization, Methodology, Writing – review & editing. **Luca Bonaventura:** Conceptualization, Methodology, Writing – original draft, Writing – review & editing.

Declaration of competing interest

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

Data availability

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

Acknowledgements

We thank M. Tavelli for providing the original data of the cavity flow simulation discussed in Section 6. We also gratefully acknowledge several useful discussions with A. Della Rocca on numerical methods related to those presented here and the constructive comments of two anonymous reviewers, which have greatly helped in improving the quality of the paper. The simulations have been partly run at CINECA thanks to the computational resources made available through the SIDICoNS - HP10CLPLXI ISCRA-C project. This work was supported by the ESCAPE-2 project, European Union's Horizon 2020 Research and Innovation Programme (Grant Agreement No. 800897).

Appendix A. Stability and monotonicity of the explicit time discretization

In this Appendix, we study the stability and monotonicity of the explicit part of the IMEX scheme applied in the paper. We recall that the Butcher tableaux for the explicit part of the method is given by

$$\begin{array}{c|ccc} 0 & 0 & & \\ \chi & \chi & 0 & \\ 1 & 1-a_{32} & a_{32} & 0 \\ \hline & \frac{1}{2}-\frac{\chi}{4} & \frac{1}{2}-\frac{\chi}{4} & \frac{\chi}{2} \end{array}$$

In [26], the choice $a_{32} = \frac{7-2\chi}{6}$ was made to maximize the stability region of the resulting scheme, but this coefficient is indeed a free parameter and can also be chosen in different ways, as long as stability is not compromised. In order to identify possible alternative choices, we perform an analysis using the concepts introduced in [35], [30], [21] (see also the review in [29]). A similar analysis for the implicit part of the IMEX scheme was carried out in [8], to which we refer for a summary of the related theoretical results. We then define

$$A = \begin{bmatrix} 0 & 0 & 0 \\ \chi & 0 & 0 \\ 1-a_{32} & a_{32} & 0 \end{bmatrix} \qquad b^T = \begin{bmatrix} \frac{1}{2}-\frac{\chi}{4} & \frac{1}{2}-\frac{\chi}{4} & \frac{\chi}{2} \end{bmatrix}$$

with $\chi = 2-\sqrt{2}$. We define for $\xi \in \mathbb{R}$ the quantities

$$\begin{aligned} A(\xi) &= A\,(I-\xi A)^{-1} \qquad b^T(\xi) = b^T\,(I-\xi A)^{-1} \\ e(\xi) &= (I-\xi A)^{-1}\,e \qquad \varphi(\xi) = 1+\xi b^T\,(I-\xi A)^{-1}\,e \end{aligned} \tag{72}$$

where I is the 3×3 identity matrix and e is a vector whose all components are equal to 1. Therefore, for the specific scheme, we obtain

$$A(\xi) = \begin{bmatrix} 0 & 0 & 0 \\ \chi & 0 & 0 \\ 1+a_{32}\,(\chi\xi-1) & a_{32} & 0 \end{bmatrix} \quad b^T(\xi) = \begin{bmatrix} \frac{1}{4}\left[2+\chi\,(-1+\xi\,(4-\chi+2a_{32}\,(\chi\xi-1)))\right] \\ \frac{1}{4}\left[2+\chi\,(2a_{32}\xi-1)\right] \\ \frac{\chi}{2} \end{bmatrix}$$

$$e(\xi) = \begin{bmatrix} 1 \\ 1+\chi\xi \\ 1+\xi+a_{32}\chi\xi^2 \end{bmatrix} \quad \varphi(\xi) = 1+\xi+\frac{\xi^2}{2}+\left(3-2\sqrt{2}\right)a_{32}\xi^3.$$

A method with tableaux $\left(A, b^T\right)$ is absolutely monotone at $\xi \in \mathbb{R}$ if $A(\xi)\geq 0$, $b^T(\xi)\geq 0$, $e(\xi)\geq 0$ and $\varphi(\xi)\geq 0$ elementwise; moreover the radius of absolute monotonicity is defined for all ξ in $-r\leq\xi\leq 0$ as

$$R(a,b) = \sup\left[r | r\geq 0, A(\xi)\geq 0, b^T(\xi)\geq 0, e(\xi)\geq 0, \varphi(\xi)\geq 0\right].$$

Fig. 26 shows the behavior of the radius of absolute monotonicity as a_{32} varies, along with the behavior of the stability region along the imaginary axis. As already mentioned before, $a_{32} = \frac{7-2\chi}{6}$ was chosen originally to maximize the stability region, but in this case $R = \frac{2\sqrt{2}-3}{2+\sqrt{2}} \approx 0.05$, so that the region of absolute monotonicity is quite small. It can be shown that the region of absolute stability is given by

$$S = \left\{z\in\mathbb{C} : \left|1+z+a_{32}\chi z^2\right| < 1\right\}.$$

The alternative value $a_{32} = 0.5$ maximizes the region of absolute monotonicity without compromising too much the stability. The impact of this alternative choice on numerical results is discussed in Section 6.

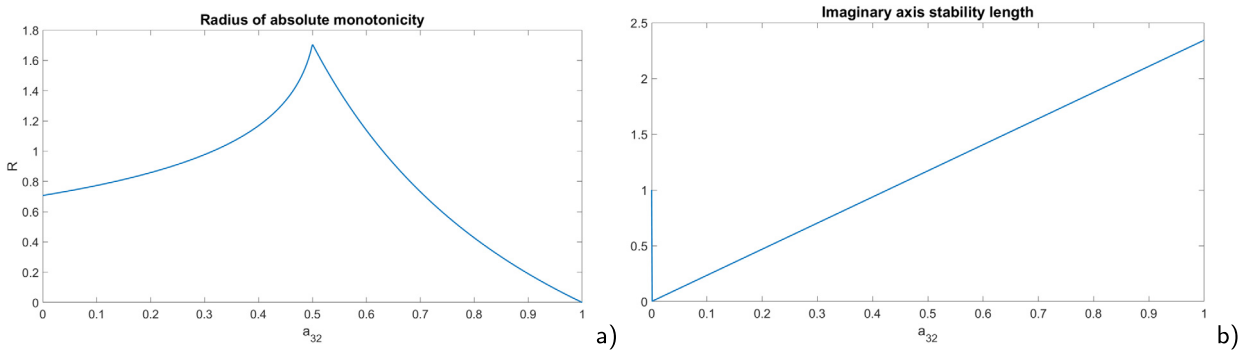

Fig. 26. Analysis of the explicit part of IMEX scheme: a) Radius of absolute monotonicity as function of a_{32}, b) Size of stability region along the imaginary axis as a_{32} varies.

Appendix B. Eigenvalues of 1D Euler equations

In this Appendix we compute the eigenvalues for the Euler equations in non-dimensional form for a general equation of state. For the sake of simplicity, we focus on 1D case and so the equations can be written as follows

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \frac{\partial}{\partial x}(\rho u) &= 0 \\ \frac{\partial \rho u}{\partial t} + \frac{\partial}{\partial x}\left(\rho u^2\right) + \frac{1}{M^2}\frac{\partial p}{\partial x} &= 0 \\ \frac{\partial \rho E}{\partial t} + \frac{\partial}{\partial x}\left[(\rho E + p)\, u\right] &= 0. \end{aligned} \tag{73}$$

This is equivalent to

$$\begin{aligned} \frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + \rho\frac{\partial u}{\partial x} &= 0 \\ \frac{\partial \rho}{\partial t}u + \frac{\partial u}{\partial t}\rho + u^2\frac{\partial \rho}{\partial x} + 2\rho u\frac{\partial u}{\partial x} + \frac{1}{M^2}\frac{\partial p}{\partial x} &= 0 \\ \frac{\partial \rho}{\partial t}E + \frac{\partial E}{\partial t}\rho + (\rho E + p)\frac{\partial u}{\partial x} + u\left(\frac{\partial \rho}{\partial x}E + \frac{\partial E}{\partial x}\rho + \frac{\partial p}{\partial x}\right) &= 0. \end{aligned} \tag{74}$$

Thanks to the continuity equation and to the relation $E = e + \frac{1}{2}M^2u^2$, we obtain

$$\begin{aligned} \frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + \rho\frac{\partial u}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho M^2}\frac{\partial p}{\partial x} &= 0 \\ \frac{\partial e}{\partial t} + \frac{p}{\rho}\frac{\partial u}{\partial x} + u\frac{\partial e}{\partial x} &= 0. \end{aligned} \tag{75}$$

In general $e = e(p, \rho)$, so that $\frac{\partial e}{\partial t} = \frac{\partial e}{\partial \rho}\frac{\partial \rho}{\partial t} + \frac{\partial e}{\partial p}\frac{\partial p}{\partial t}$ and $\frac{\partial e}{\partial x} = \frac{\partial e}{\partial \rho}\frac{\partial \rho}{\partial x} + \frac{\partial e}{\partial p}\frac{\partial p}{\partial x}$. Hence, the system reduces to

$$\begin{aligned} \frac{\partial \rho}{\partial t} + u\frac{\partial \rho}{\partial x} + \rho\frac{\partial u}{\partial x} &= 0 \\ \frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + \frac{1}{\rho M^2}\frac{\partial p}{\partial x} &= 0 \\ \frac{\partial p}{\partial t} + \frac{\left(\frac{p}{\rho} - \rho\frac{\partial e}{\partial \rho}\right)}{\frac{\partial e}{\partial p}}\frac{\partial u}{\partial x} + u\frac{\partial p}{\partial x} &= 0, \end{aligned} \tag{76}$$

which can be thought in the following vector form as $\frac{\partial \mathbf{Q}}{\partial t} + \mathbf{A}\frac{\partial \mathbf{Q}}{\partial x} = \mathbf{0}$ with

$$\mathbf{Q} = \begin{bmatrix} \rho \\ u \\ p \end{bmatrix} \tag{77}$$

and

$$\mathbf{A} = \begin{bmatrix} u & \rho & 0 \\ 0 & u & \frac{1}{\rho M^2} \\ 0 & \frac{\left(\frac{p}{\rho} - \rho \frac{\partial e}{\partial \rho}\right)}{\frac{\partial e}{\partial p}} & u \end{bmatrix}. \tag{78}$$

The eigenvalues of (78) are $u - \frac{1}{M}\frac{1}{\rho}\sqrt{\frac{p - \frac{\partial e}{\partial \rho}\rho^2}{\frac{\partial e}{\partial p}}}$, u and $u + \frac{1}{M}\frac{1}{\rho}\sqrt{\frac{p - \frac{\partial e}{\partial \rho}\rho^2}{\frac{\partial e}{\partial p}}}$. The first law of thermodynamics, already recalled in Section 3, provides us the following relation

$$Tds = de - \frac{p}{\rho^2}d\rho = \left(\frac{\partial e}{\partial \rho} - \frac{p}{\rho^2}\right)d\rho + \frac{\partial e}{\partial p}dp, \tag{79}$$

or, equivalently,

$$dp = \frac{\frac{p}{\rho^2} - \frac{\partial e}{\partial \rho}}{\frac{\partial e}{\partial p}}d\rho + \frac{T}{\frac{\partial e}{\partial p}}ds. \tag{80}$$

Hence, following [56], we have

$$c^2 = \left.\frac{\partial p}{\partial \rho}\right|_s = \frac{\frac{p}{\rho^2} - \frac{\partial e}{\partial \rho}}{\frac{\partial e}{\partial p}} \tag{81}$$

and, therefore, the eigenvalues of (78) are

$$u + \frac{c}{M} \quad u \quad u + \frac{c}{M}$$

also for a generic equation of state, and not only in the case of an ideal gas, as already discussed in [41]. This justifies the definition (57) also in case of non-ideal gases.

References

[1] D. Arndt, W. Bangerth, M. Feder, M. Fehling, R. Gassmöller, T. Heister, L. Heltai, M. Kronbichler, M. Maier, P. Munch, J.P. Pelteret, S. Sticko, B. Turcksin, D. Wells, The deal.II Library, Version 9.4, J. Numer. Math. (2022).

[2] D. Arnold, An interior penalty finite element method with discontinuous elements, SIAM J. Numer. Anal. 19 (1982) 742–760.

[3] W. Bangerth, R. Hartmann, G. Kanschat, Deal II: a general-purpose object-oriented finite element library, ACM Trans. Math. Softw. 33 (2007) 24–51.

[4] R. Bank, W. Coughran, W. Fichtner, E.H. Grosse, D.J. Rose, R. Smith, Transient simulation of silicon devices and circuits, IEEE Trans. Electron Devices 32 (1985) 1992–2007.

[5] F. Bassi, L. Botti, A. Colombo, A. Ghidoni, F. Massa, Linearly implicit Rosenbrock-type Runge–Kutta schemes applied to the discontinuous Galerkin solution of compressible and incompressible unsteady flows, Comput. Fluids 118 (2015).

[6] F. Bassi, A. Crivellini, D. Di Pietro, S. Rebay, An implicit high-order discontinuous Galerkin method for steady and unsteady incompressible flows, Comput. Fluids 36 (2007) 1529–1546.

[7] L. Bonaventura, A semi-implicit, semi-Lagrangian scheme using the height coordinate for a nonhydrostatic and fully elastic model of atmospheric flows, J. Comput. Phys. 158 (2000) 186–213.

[8] L. Bonaventura, A. Della Rocca, Unconditionally strong stability preserving extensions of the TR-BDF2 method, J. Sci. Comput. 70 (2017) 859–895.

[9] L. Bonaventura, R. Redler, R. Budich, Earth System Modelling 2: Algorithms, Code Infrastructure and Optimisation, Springer Verlag, New York, 2012.

[10] S. Boscarino, J. Qiu, G. Russo, T. Xiong, High order semi-implicit WENO schemes for all-Mach full Euler system of gas dynamics, SIAM J. Sci. Comput. 44 (2022) B368–B394, https://doi.org/10.1137/21M1424433.

[11] W. Boscheri, L. Pareschi, High order pressure-based semi-implicit IMEX schemes for the 3D Navier-Stokes equations at all Mach numbers, J. Comput. Phys. 434 (2021) 110206.

[12] K. Burrage, T. Tian, Stiffly accurate Runge–Kutta methods for stiff stochastic differential equations, Comput. Phys. Commun. 142 (2001) 186–190, https://doi.org/10.1016/S0010-4655(01)00324-1.

[13] S. Busto, M. Tavelli, W. Boscheri, M. Dumbser, Efficient high order accurate staggered semi-implicit discontinuous Galerkin methods for natural convection problems, Comput. Fluids 198 (2020) 104399.

[14] J. Butcher, Numerical Methods for Ordinary Differential Equations, 2 ed., Wiley, 2008.

[15] V. Casulli, D. Greenspan, Pressure method for the numerical solution of transient, compressible fluid flows, Int. J. Numer. Methods Fluids 4 (1984) 1001–1012.

[16] M. Cullen, A test of a semi-implicit integration technique for a fully compressible non-hydrostatic model, Q. J. R. Meteorol. Soc. 116 (1990) 1253–1258.

[17] M. Dumbser, V. Casulli, A conservative, weakly nonlinear semi-implicit finite volume scheme for the compressible Navier-Stokes equations with general equation of state, Appl. Math. Comput. 272 (2016) 479–497.

[18] M. Dumbser, O. Zanotti, R. Loubère, S. Diot, A posteriori subcell limiting of the discontinuous Galerkin finite element method for hyperbolic conservation laws, J. Comput. Phys. 278 (2014) 47–75.

[19] N. Fehn, M. Kronbichler, C. Lehrenfeld, G. Lube, P. Schroeder, High-order DG solvers for under-resolved turbulent incompressible flows: a comparison of l^2 and $h(div)$ methods, Int. J. Numer. Methods Fluids 91 (2019) 533–556.

[20] M. Fernandez, Propellant tank pressurization modeling for a hybrid rocket, Master's thesis, Rochester Institute of Technology, 2009.

[21] L. Ferracina, M. Spijker, Stepsize restrictions for the total-variation-diminishing property in general Runge-Kutta methods, SIAM J. Numer. Anal. 42 (2004) 1073–1093.

[22] M. Gandolfi, Baer-Nunziato type models for the simulation of hybrid rockets self-pressurizing tanks, Master's thesis, Politecnico di Milano, 2019.
[23] U. Ghia, K. Ghia, C. Shin, High-Re solutions for incompressible flow using the Navier-Stokes equations and a multigrid method, J. Comput. Phys. 48 (1982) 387–411, https://doi.org/10.1016/0021-9991(82)90058-4.
[24] F. Giraldo, Semi-implicit time-integrators for a scalable spectral element atmospheric model, Q. J. R. Meteorol. Soc. 131 (2005) 2431–2454.
[25] F. Giraldo, An Introduction to Element-Based Galerkin Methods on Tensor-Product Bases, Springer Nature, 2020.
[26] F. Giraldo, J. Kelly, E. Constantinescu, Implicit-explicit formulations of a three-dimensional nonhydrostatic unified model of the atmosphere (NUMA), SIAM J. Sci. Comput. 35 (2013) 1162–1194.
[27] F. Giraldo, M. Restelli, M. Läuter, Semi-implicit formulations of the Navier–Stokes equations: application to nonhydrostatic atmospheric modeling, SIAM J. Sci. Comput. 32 (2010) 3394–3425.
[28] S. Gottlieb, C. Shu, Total variation diminishing Runge-Kutta schemes, Math. Comput. 67 (1998) 73–85.
[29] S. Gottlieb, C. Shu, E. Tadmor, Strong stability-preserving high-order time discretization methods, SIAM Rev. 43 (2001) 89–112.
[30] I. Higueras, On strong stability preserving time discretization methods, J. Sci. Comput. 21 (2004) 193–223.
[31] M. Hosea, L. Shampine, Analysis and implementation of TR-BDF2, Appl. Numer. Math. 20 (1996) 21–37.
[32] G. Karniadakis, S. Sherwin, Spectral *hp*-Element Methods for Computational Fluid Dynamics, Oxford University Press, 2005.
[33] C. Kennedy, M. Carpenter, Additive Runge-Kutta schemes for convection-diffusion-reaction equations, Appl. Numer. Math. 44 (2003) 139–181.
[34] R. Klein, N. Botta, T. Schneider, C.D. Munz, S. Roller, A. Meister, L. Hoffmann, T. Sonar, Asymptotic adaptive methods for multi-scale problems in fluid mechanics, J. Eng. Math. 39 (2001) 261–343.
[35] J. Kraaijevanger, Contractivity of Runge-Kutta methods, BIT 31 (1991) 482–528.
[36] C. Kühnlein, W. Deconinck, R. Klein, S. Malardel, Z. Piotrowski, P. Smolarkiewicz, J. Szmelter, N. Wedi, FVM 1.0: a nonhydrostatic finite-volume dynamical core formulation for IFS, Geosci. Model Dev. 12 (2019) 651–676.
[37] O. Le Métayer, R. Saurel, The Noble-Abel stiffened-gas equation of state, Phys. Fluids 28 (2016) 046102, https://doi.org/10.1063/1.4945981.
[38] E. Lemmon, R. Span, Short fundamental equations of state for 20 industrial fluids, J. Chem. Eng. Data 51 (2006) 785–850.
[39] R. LeVeque, Finite Volume Methods for Hyperbolic Problems, Cambridge University Press, 2002.
[40] S. Lias, J. Bartmess, J. Liebman, J. Holmes, R. Levin, G. Mallard, NIST Chemistry Webbook Standard Reference Database Number 69, National Institute of Standards Technology, 2010.
[41] C. Munz, S. Roller, R. Klein, K. Geratz, The extension of incompressible flow solvers to the weakly compressible regime, Comput. Fluids 32 (2003) 173–196.
[42] P. Nederstigt, Real gas thermodynamics: and the isentropic behavior of substances, Master's thesis, Delft University of Technology, 2017.
[43] G. Orlando, A filtering monotonization approach for DG discretizations of hyperbolic problems, Technical Report, 2022, arXiv:2204.08693.
[44] G. Orlando, A. Della Rocca, P. Barbante, L. Bonaventura, N. Parolini, An efficient and accurate implicit DG solver for the incompressible Navier-Stokes equations, Int. J. Numer. Methods Fluids (2022) 1–33, https://doi.org/10.1002/fld.5098.
[45] L. Pareschi, G. Russo, Implicit–explicit Runge–Kutta schemes and applications to hyperbolic systems with relaxation, J. Sci. Comput. 25 (2005) 129–155.
[46] M. Restelli, Semi-Lagrangian and semi-implicit discontinuous Galerkin methods for atmospheric modeling applications, Ph.D. thesis, Politecnico di Milano, 2007.
[47] M. Restelli, F. Giraldo, A conservative discontinuous Galerkin semi-implicit formulation for the Navier-Stokes equations in nonhydrostatic mesoscale modeling, SIAM J. Sci. Comput. 31 (2009) 2231–2257.
[48] A. Robert, A semi-Lagrangian and semi-implicit numerical integration scheme for the primitive meteorological equations, J. Meteorol. Soc. Jpn. 60 (1982) 319–325.
[49] P. Smolarkiewicz, C. Kühnlein, N. Wedi, Semi-implicit integrations of perturbation equations for all-scale atmospheric dynamics, J. Comput. Phys. 376 (2019) 145–159.
[50] G. Sod, A survey of several finite difference methods for systems of nonlinear hyperbolic conservation laws, J. Comput. Phys. 27 (1978) 1–31, https://doi.org/10.1016/0021-9991(78)90023-2.
[51] R. Span, Multiparameter Equations of State, Springer, 2000.
[52] J. Steppeler, R. Hess, G. Doms, U. Schättler, L. Bonaventura, Review of numerical methods for nonhydrostatic weather prediction models, Meteorol. Atmos. Phys. 82 (2003) 287–301.
[53] G. Strang, On the construction and comparison of difference schemes, SIAM J. Numer. Anal. 5 (1968) 506–517.
[54] M. Tavelli, M. Dumbser, A pressure-based semi-implicit space–time discontinuous Galerkin method on staggered unstructured meshes for the solution of the compressible Navier–Stokes equations at all Mach numbers, J. Comput. Phys. 341 (2017) 341–376.
[55] G. Tumolo, L. Bonaventura, A semi-implicit, semi-Lagrangian discontinuous Galerkin framework for adaptive numerical weather prediction, Q. J. R. Meteorol. Soc. 141 (2015) 2582–2601.
[56] J. Vidal, Thermodynamics: applications to chemical engineering and petroleum industry, Editions Technip., 2001.
[57] J. Zeifang, J. Schütz, K. Kaiser, A. Beck, M. Lukáčová-Medvid'ová, S. Noelle, A novel full-Euler low Mach number IMEX splitting, Commun. Comput. Phys. 27 (2019) 292–320, https://doi.org/10.4208/cicp.OA-2018-0270.