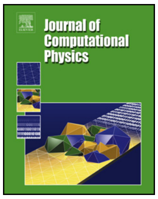
Journal of Computational Physics 433 (2021) 110017

Contents lists available at ScienceDirect

Journal of Computational Physics

www.elsevier.com/locate/jcp



# Development of an algebraic fractional step scheme for the primitive formulation of the compressible Navier-Stokes equations

Samuel Parada [a,b], Ramon Codina [a,b,*], Joan Baiges [b]

[a] *Centre Internacional de Mètodes Numèrics en Enginyeria (CIMNE), Edifici C1, Campus Nord UPC, Gran Capitán S/N, 08034 Barcelona, Spain*
[b] *Universitat Politècnica de Catalunya – Barcelona Tech, Jordi Girona 1-3, 08034 Barcelona, Spain*

ARTICLE INFO

*Article history:*
Received 1 April 2020
Received in revised form 12 November 2020
Accepted 17 November 2020
Available online 11 February 2021

*Keywords:*
Compressible flow
Variational multiscale method
Fractional step schemes
Finite element method
Weak imposition of boundary conditions

ABSTRACT

In this work we address the compressible Navier-Stokes equations written in the so-called primitive formulation. The proposed methodology is a finite-element solver based on a fractional step scheme in time, which allows to uncouple the calculation of the problem unknowns providing important savings in computational cost. In addition, we include a stabilization technique within the Variational Multi-Scale framework and, in particular, we consider orthogonal and dynamic definitions for the subscales. In order to overcome any wave reflections which may arise in aeroacoustic simulations at the low compressibility regime, we present a method for enforcing boundary conditions based on a combination of a zero order non-reflecting condition plus the weak imposition of Dirichlet boundary conditions over the external contours. Several representative benchmark flow simulations are performed, which demonstrate the suitability of the proposed algorithm for the subsonic regime.

## 1. Introduction

The so-called compressible Navier–Stokes equations are commonly used to model flow problems where compressibility effects become relevant, e.g. in the aerodynamic and aeroacoustic fields, with applications ranging from classical turbo-machinery design to modern speech therapy simulations. The general mathematical setting consists of the momentum, mass and energy conservation equations, which classically read as,

$$\rho \frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} - \nabla\cdot\boldsymbol{\sigma} = \rho\mathbf{f}, \tag{1a}$$

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho\nabla\cdot\mathbf{u} = 0, \tag{1b}$$

$$\rho\frac{\mathrm{d}e}{\mathrm{d}t} + \nabla\cdot\mathbf{q} - \boldsymbol{\sigma} : \boldsymbol{\varepsilon} = \rho r, \tag{1c}$$

$$\rho = \rho(p,\theta), \tag{1d}$$

$$e = e(p,\theta). \tag{1e}$$

[*] Corresponding author.
*E-mail addresses:* samuel.parada@upc.edu (S. Parada), ramon.codina@upc.edu (R. Codina), joan.baiges@upc.edu (J. Baiges).

https://doi.org/10.1016/j.jcp.2020.110017

Let us set some notation to be used throughout the text. In these equations, the operator $\mathrm{d}(\cdot)/\mathrm{d}t$ stands for the total time derivative, $\mathbf{u}$ is the velocity field, $p$ is the pressure, $\rho$ is the density, $\theta$ stands for the temperature, $\mathbf{f}$ is a possible acceleration vector, $e$ represents the internal energy, $\mathbf{q}$ stands for the heat flux vector, $\boldsymbol{\sigma}$ is the stress tensor, $\boldsymbol{\varepsilon}(\mathbf{u}) := (\nabla\mathbf{u} + \nabla^{\mathsf{T}}\mathbf{u})/2$ represents the strain rate tensor and $r$ is a possible heat source, which may include chemical reactions or electromagnetic effects. The last two equations Eqs. (1d)–(1e) are general representations of the thermal and caloric equations of state, respectively, which are needed to close the well-posedness of the problem, together with suitable initial and boundary data. The previous set of partial differential equations describes a wide range of scales including sound wave propagation, a phenomenon which is known to require accurate computational methods. In general, one could proceed either by choosing small mesh and time step sizes or by using high precision numerical schemes. Regardless of the selected approach, obtaining a representative solution of problem (1a)–(1c) is particularly demanding from the computational point of view and this still remains as one of the main limitations in compressible flow simulations in spite of the increasing amount of computing facilities available for the scientific and engineering communities.

The development of numerical approximations for the compressible Navier–Stokes equations which perform adequately at low compressibility regimes is one of the major concerns related to compressible flow simulations. However, the classical compressible flow solvers found in the literature display a deterioration in the solution when the free stream Mach number is reduced. According to [1], the rationale behind this misbehavior is a possible mismatch between numerical and continuous fluxes, a fact principally attributed to the broad difference in length and time scales of the solution. In this sense, a widespread alternative for performing wave-propagation simulations is to solve the classical incompressible Navier–Stokes equations supplemented with an aeroacoustic model, commonly referred in the literature as hybrid methods, see e.g. [2–5]. Nonetheless, this common approach generally involves a sequential calculation of aerodynamics and aeroacoustics, what directly leads to discard any feedback related to the sound waves into the flow. Apart from this, and although they are not suitable for wave-like problems, there exist the so-called unified methods, a completely different family of techniques which are intended to be suitable for either compressible and incompressible flows. Within this field, it is worth mentioning the Low Mach formulation in [6,7], but also the physical models developed in [8,9], which attempt to be valid at any compressibility regime.

Probably the main reason for developing a compressible formulation which could be properly applied to low Mach flows is the fact that very low Mach number zones can coexist with regions where the flow compressibility becomes significant. In addition to this, several applications traditionally solved with incompressible formulations could be successfully handled by appropriate compressible solution techniques. The general trend in the literature is to make use of the conservative variables (namely density, momentum, and total energy) for the compressible formulation, whereas primitive variables (pressure, velocity, and temperature) are preferred for the incompressible equations. However, including density as a variable in the compressible problem (like in the conservative formulation), might yield singularities for problems within the low Mach number limit. Hence, primitive and even entropic unknowns remain as the two main possibilities in order to solve compressible problems posed in low compressibility conditions. Particularly, the entropy variable formulation ensures a global entropy stability condition, but it is subject to the definition of the entropy function [10]. For a general review on the different set of variables for solving compressible and incompressible flow problems we refer to [11].

Another particular feature of nearly incompressible aeroacoustic flows is that external computational boundaries may produce artificial wave reflections related to the ingoing part of the sound waves and which can completely pollute the solution of the problem. Ingoing waves may not only interfere with the acoustic signal, but they can also produce numerical instabilities if the numerical method is not able to provide enough dissipation. Among the most remarkable numerical techniques which deal with the backscattering of waves in the acoustic field, we would like to highlight the damping of the compressible equations and the application of non-reflecting boundary conditions. Even though performing a damping of some terms of the compressible equations is a robust approach to face spurious reflections at the boundary (usually referred to as *buffer* or *sponge* zones), this technique brings an extra computational effort related to the new terms that need to be included over an augmented computational domain. Therefore, other approaches are often adopted, being one of the main alternatives the inclusion of non-reflecting boundary conditions, an approach that we favor in this work. The backscattering issue represents a traditional research topic in aeroacoustics and thus the literature on compressible boundary conditions is really extensive. Nevertheless, we refer to the early works in [12,13] and to the more recent reviews in [14,15] for a deeper understanding on this topic.

The principal objective of the present work is to discuss the development of pressure segregation methods for the transient compressible Navier–Stokes equations written in primitive variables and using a finite element approximation for the space discretization. As a reference in the comparisons, we will take the solution of the so-called *monolithic* problem, that is to say, the standard coupled calculation involving all the problem unknowns. Clearly, the fully discrete and linearized monolithic scheme leads to an algebraic system the structure of which can be exploited so as to solve independently for the velocity, the pressure and the temperature degrees of freedom. Referring to the time integration, on the following we will concentrate on first and second order implicit finite difference schemes. The backward Euler method will be used for the former, whereas for second order methods we will stick to backward differentiation (or Gear) schemes, yet our developments are not restricted and, in principle, any other discretization methods might be used to advance the solution in time. The technique we will discuss here corresponds to the classical *fractional step* algorithms. Fractional step methods, commonly referred in the literature as *splitting* or *projection* methods, started to gain adepts after the original works of Chorin [16] and Temam [17], who attempted the uncoupling of velocity and pressure at the continuous level for the incompressible equations. Since then, many works have been devoted to a proper understanding of the original schemes, their numerical analysis, their extension to higher order approximations and to the design of adequate boundary conditions (see e.g. [18] and references therein). In contrast to this, our approach in this work is to present the splitting of the equations at the pure algebraic level once the equations have already been discretized in space and in time (see e.g. [19–22] for algebraic approaches on the incompressible, viscoelastic and isentropic flow problems). This way to face the problem emerged after the identification in [23] of the classical pressure segregation method as an inexact factorization of the final algebraic system. In this work we favor such algebraic viewpoint since it is generally simpler and, although a particular treatment of boundary conditions will be discussed to avoid the reflection of sound waves, it makes possible to obviate a discussion on specific boundary conditions for the different steps of the fractional step scheme (see e.g. [19,24]). The literature regarding fractional step schemes for compressible flow applications is not as vast as for the incompressible problem. However, many innovative and diverse collections of uncoupling techniques [25–30] emerged, which involve distinct spatial discretizations (finite differences, finite volumes, finite elements), temporal schemes (explicit, implicit, semi-implicit) and also different sets of unknowns (conservative, primitive and even combinations of both).

It is well known that classical Galerkin formulations applied to problems of the form of (1a)–(1c) may suffer from different types of numerical instabilities. The main ones arise, precisely, from the non-elliptic nature of the equations and also due to an inf-sup-like condition which restrains the compatibility of the interpolation spaces of the different variables in play. Both inconveniences can be overcome by resorting to a stabilized finite element formulation. In particular, the one adopted in this work is based on the Variational MultiScale Method (VMS) [31–33]. For an in-depth review of VMS methods in CFD with several examples of application we refer the reader to [34] and, specifically in the context of compressible flows, to [35,36], being the latter the first attempt to introduce stabilization techniques strictly in the frame of VMS methods. The basic idea of this approach is to approximate the effect of the component of the continuous solution which cannot be resolved by the finite element mesh. An important feature of the formulation adopted herein is that the unresolved component, hereafter referred to as subgrid scale or subscale, is assumed to be orthogonal to the finite element space, as originally discussed in [37]. In addition to this, we will also consider the time tracking of the subscales, which has become an effective feature in order to eliminate numerical oscillations originated by initial transients while minimizing numerical dissipation [38,39]. Finally, since only subsonic flows are considered in this study, the extra control over the gradients of the solution given by discontinuity capturing operators is not required, and consequently, shock capturing techniques are not surveyed here.

The outline of the article is as follows: in Section 2 we introduce the compressible Navier–Stokes problem written in primitive variables, as well as its variational formulation. In Section 3 we discuss the imposition of boundary conditions in order to avoid the spurious wave reflections at the external boundaries of the computational domain, whereas in Section 4 we present the standard finite element approximation and the monolithic time discretization. Section 5 is devoted to the design of pressure-correction algorithms, taking into account the modifications due to the application of boundary conditions. In Section 6, we introduce the VMS stabilized finite element formulation and Section 7 collects the relevant adjustments that need to be considered on the stabilized version of the fractional step method previously mentioned. Numerical experiments are conducted in Section 8, and, finally, we close the paper with some concluding statements in Section 9.

## 2. Problem statement

### 2.1. Preliminaries

In this first subsection we introduce some simple relations which allow to transform the original set in Eqs. (1a)–(1c) into the so-called primitive version (see e.g. [40] for additional details). The constitutive equation for a Newtonian fluid relates the stress tensor $\boldsymbol{\sigma}$ with the strain rate tensor $\boldsymbol{\varepsilon}(\mathbf{u})$ through a linear equation, viz.,

$$\boldsymbol{\sigma} = -p\mathbf{I} + 2\mu\boldsymbol{\varepsilon}(\mathbf{u}) + (\mu_B + \lambda)(\nabla\cdot\mathbf{u})\mathbf{I}, \tag{2}$$

where $\mu$, $\mu_B$ and $\lambda$ are respectively, the so-called *molecular* (or *shear*), *bulk* (or *volumetric*) and *second* viscosity coefficients. A common practice in the analysis of the motion of compressible fluids is to make use of the well-known Stokes hypothesis, i.e. $\lambda + \frac{2}{3}\mu = \mu_B = 0$. Setting $\mu_B = 0$ is supported by the kinetic theory of gases and such assumption renders the mathematical treatment of compressible flows notably easier, but it has been the object of long-lasting discussions on compressible flows simulations. As a matter of fact, only in very particular conditions will the term $\mu_B(\nabla\cdot\mathbf{u})$ be of practical significance. According to experimental evidence, this may happen for example when the fluid is characterized by large values of $\mu_B$ (e.g., $CO_2$), or the motion is such that extremely large values of $\nabla\cdot\mathbf{u}$ occur, e.g. in hypersonic flows, which are out of the scope of the present article.

The third term in the energy conservation equation Eq. (1c) can be formally rewritten as,

$$\boldsymbol{\sigma} : \boldsymbol{\varepsilon}(\mathbf{u}) = -p(\nabla\cdot\mathbf{u}) + \Phi = -p(\nabla\cdot\mathbf{u}) + 2\mu\boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}) - \frac{2}{3}\mu(\nabla\cdot\mathbf{u})^2, \tag{3}$$

where $\Phi$ is the so-called *dissipation* function which represents the rate at which mechanical energy is being converted to thermal energy (Joule's effect). Usually, this term in the energy equation is negligible for the Newtonian fluid model, and it might only be taken into account when highly viscous flows are considered. In particular, this term is fundamental when the flow of viscoplastic materials is studied but it is also of paramount importance in the small scales of turbulence in compressible flows.

The heat flux vector is calculated using Fourier's law of heat conduction, which in general can be written as $\mathbf{q} = -\nabla g(\theta)$ where $g$ is a nonlinear function of the temperature. Although nonlinear diffusion problems are often found in practice, we will restrict ourselves to the classical linear relation in order to ease the discussion. Hence we write,

$$\mathbf{q} := -\kappa \nabla\theta, \tag{4}$$

being $\kappa$ the thermal conduction coefficient.

Henceforth, we will make use of the most frequently encountered form of the thermal equation of state, i.e. the ideal-gas law $p = \rho R_\mathrm{g}\theta$, where $R_\mathrm{g}$ is the constant of the gas under consideration, defined as $R_\mathrm{g} = R^0 M_\mathrm{w}^{-1}$ being $R^0 = 8.31$ J/(mol K) the universal gas constant and $M_\mathrm{w}$ the molecular weight. For a calorically perfect gas, the internal energy is a sole function of the temperature, implying that the caloric equation of state can be simplified to give $e = e(\theta) := c_\mathrm{v}\theta$. Here and in what follows, $c_\mathrm{v}$ denotes the specific heat at constant volume, and $c_\mathrm{p}$ stands for the specific heat at constant pressure. We also define the ratio of specific heats as $\gamma := c_\mathrm{p}/c_\mathrm{v}$.

Furthermore, two physical variables are introduced, which relate temperature, pressure and density derivatives. These are the so-called volume expansivity and isothermal compressibility coefficients, denoted respectively as $\alpha$ and $\beta$ hereafter. Their definitions are,

$$\alpha := -\frac{1}{\rho}\left(\frac{\partial\rho}{\partial\theta}\right)_p, \quad \beta := \frac{1}{\rho}\left(\frac{\partial\rho}{\partial p}\right)_\theta, \tag{5}$$

where $(\cdot)_\theta$ and $(\cdot)_p$ stand for a constant temperature and pressure constrain. Since the ideal gas law is considered, it is readily checked that the previous expressions simply reduce to $\beta = p^{-1}$ and $\alpha = \theta^{-1}$.

### 2.2. Initial and boundary value problem

Let $\Omega$ be an open, bounded and polyhedral domain of $\mathbb{R}^{n_\mathrm{sd}}$ ($n_\mathrm{sd} = 2$ or 3 is the number of space dimensions) and $[0, t_\mathrm{f}]$ the time interval of analysis. The unknowns of the problem are the *primitive* variables, i.e. the fluid velocity $\mathbf{u}(\mathbf{x}, t) : \Omega \times (0, t_\mathrm{f}) \to \mathbb{R}^{n_\mathrm{sd}}$, the thermodynamic pressure $p(\mathbf{x}, t) : \Omega \times (0, t_\mathrm{f}) \to \mathbb{R}$, and the temperature $\theta(\mathbf{x}, t) : \Omega \times (0, t_\mathrm{f}) \to \mathbb{R}$ which are the solution of the following system of partial differential equations:

$$\rho\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} - 2\nabla\cdot(\mu\boldsymbol{\varepsilon}) + \frac{2}{3}\nabla(\mu\nabla\cdot\mathbf{u}) + \nabla p = \rho\mathbf{f} \qquad \text{in } \Omega \times (0, t_\mathrm{f}), \tag{6a}$$

$$\beta\frac{\mathrm{d}p}{\mathrm{d}t} - \alpha\frac{\mathrm{d}\theta}{\mathrm{d}t} + \nabla\cdot\mathbf{u} = 0 \qquad \text{in } \Omega \times (0, t_\mathrm{f}), \tag{6b}$$

$$\rho c_\mathrm{v}\frac{\mathrm{d}\theta}{\mathrm{d}t} - \nabla\cdot(\kappa\nabla\theta) + p(\nabla\cdot\mathbf{u}) = Q \qquad \text{in } \Omega \times (0, t_\mathrm{f}), \tag{6c}$$

where we made use of the relations in Eqs. (2)–(5) from previous the subsection. In addition, $Q$ stands for the energy source terms, which accounts for mechanical dissipation into heat, chemical reactions or even electromagnetic effects. Therefore,

$$Q := \rho r + 2\mu\boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u}) - \frac{2}{3}\mu(\nabla\cdot\mathbf{u})^2. \tag{7}$$

Here and in what follows, $\mu$, $\kappa$ and $c_\mathrm{v}$ are assumed to be constant to ease the discussion. The reader should note that the coupling among all the variables in the primitive problem defined by Eqs. (6a)–(6c) is remarkably relevant, mostly through non-linear relations.

Boundary conditions need to be appended to the previous problem. Let us consider the disjoint splittings $\partial\Omega = \overline{\partial\Omega_{\mathrm{D},\mathbf{u}} \cup \partial\Omega_{\mathrm{N},\mathbf{u}}} = \overline{\partial\Omega_{\mathrm{D},\theta} \cup \partial\Omega_{\mathrm{N},\theta}}$. Subscript D refers to Dirichlet or essential boundary conditions, whereas N refers to Neumann or natural boundary conditions. The second subscript indicates the variable to which the condition is applied. Note that boundaries for velocity and temperature may overlap. Let $\mathbf{n}$ be the unit vector normal to $\partial\Omega$, $\mathbf{u}_\mathrm{g}$ the given velocity prescribed on $\partial\Omega_{\mathrm{D},\mathbf{u}}$, $\mathbf{t}$ the prescribed traction on $\partial\Omega_{\mathrm{N},\mathbf{u}}$, $\theta_\mathrm{g}$ the given temperature on $\partial\Omega_{\mathrm{D},\theta}$ and $\varphi$ the prescribed heat flux on $\partial\Omega_{\mathrm{N},\theta}$. The boundary conditions to be considered for all time $t \in (0, t_\mathrm{f}]$ are *initially* written as:

$$\mathbf{u} = \mathbf{u}_\mathrm{g} \qquad \text{on } \partial\Omega_{\mathrm{D},\mathbf{u}}, \tag{8a}$$

$$\mathbf{n}\cdot\boldsymbol{\sigma} = \mathbf{t} \qquad \text{on } \partial\Omega_{\mathrm{N},\mathbf{u}}, \tag{8b}$$

$$\theta = \theta_\mathrm{g} \qquad \text{on } \partial\Omega_{\mathrm{D},\theta}, \tag{8c}$$

$$-\kappa\mathbf{n}\cdot\nabla\theta = \varphi \qquad \text{on } \partial\Omega_{\mathrm{N},\theta}. \tag{8d}$$

Sometimes, the Neumann-type prescription for the temperature has to be generalized to a Robbin boundary condition to include the surface heat convection, although this is immaterial for what follows. To complete the definition of the problem we need to add initial conditions of the form $\mathbf{u}(\mathbf{x},0)=\mathbf{u}_0(\mathbf{x})$, $p(\mathbf{x},0)=p_0(\mathbf{x})$, $\theta(\mathbf{x},0)=\theta_0(\mathbf{x})$ all holding in the spatial domain ($\mathbf{x}\in\Omega$) at $t=0$.

Some dimensionless numbers which might be significant for the problem analyzed here are,

$$\mathrm{Re} := \frac{\rho U \ell}{\mu} \qquad \text{Reynolds number,} \tag{9a}$$

$$\mathrm{Ma} := \frac{U}{c} \qquad \text{Mach number,} \tag{9b}$$

$$\mathrm{Ra} := \frac{|\mathbf{g}|\rho^2 c_\mathrm{p}\Delta\theta}{\mu\kappa} \qquad \text{Rayleigh number,} \tag{9c}$$

$$\mathrm{Pr} := \frac{c_\mathrm{p}\mu}{\kappa} \qquad \text{Prandtl number,} \tag{9d}$$

$$\mathrm{Gr} := \frac{|\mathbf{g}|\rho^2\Delta\theta}{\mu^2} \qquad \text{Grashof number,} \tag{9e}$$

$$\mathrm{Fr} := \frac{U^2}{|\mathbf{g}|\Delta\theta\ell} \qquad \text{Froude number,} \tag{9f}$$

$$\mathrm{Nu} := \frac{\phi\ell}{\kappa} \qquad \text{Nusselt number,} \tag{9g}$$

where $\ell$ is the characteristic length of the problem, $U$ a characteristic velocity, $\Delta\theta$ a characteristic temperature difference, $\mathbf{g}$ a buoyancy force vector, $\phi$ is the heat transfer coefficient, and $c := \sqrt{\gamma R_\mathrm{g}\theta}$ is the speed of sound in the considered medium. The Mach number is particularly relevant in this work since it defines the compressibility regime. These dimensionless numbers are not independent from each other. Some important relations among them are: $\mathrm{Ra}=\mathrm{Gr}\ \mathrm{Pr}$, $\mathrm{Fr}=\mathrm{Re}^2\mathrm{Gr}^{-1}$.

**Remark 2.1.** For ideal gases at the low Mach number limit, the fluid usually presents very large values of pressure and temperature, specially if the international system is used. Therefore, both volume expansivity and isothermal compressibility coefficients tend to zero, and the system of Eqs. (6a)–(6c) recovers the classical incompressible form with almost constant density.

### 2.3. Variational formulation

In order to write the weak form of Eqs. (6a)–(6c) together with the boundary conditions in Eqs. (8a)–(8d) let $\mathbf{v}$, $q$, and $\eta$ be the test functions for $\mathbf{u}$, $p$ and $\theta$, respectively. We consider them time-independent, and additionally $\mathbf{v}$ and $\eta$ are assumed to vanish separately on $\partial\Omega_{\mathrm{D},\mathbf{u}}$ and $\partial\Omega_{\mathrm{D},\theta}$, respectively.

Let now $V_\mathbf{u}$, $\mathscr{V}_p$ and $V_\theta$ be, respectively, the proper functional spaces where each component of the velocity, the pressure and the temperature are well defined for each fixed time $t\in(0,t_\mathrm{f})$, with appropriate regularity. In addition, let us further introduce the following functional spaces of trial solutions and test functions,

$$\mathscr{V}_\mathbf{u} = \left\{\mathbf{u}\in[V_\mathbf{u}]^{n_\mathrm{sd}} \mid \mathbf{u}|_{\partial\Omega_{\mathrm{D},\mathbf{u}}}=\mathbf{u}_\mathrm{g}\right\}, \qquad \mathscr{W}_\mathbf{u} = \left\{\mathbf{v}\in[V_\mathbf{u}]^{n_\mathrm{sd}} \mid \mathbf{v}|_{\partial\Omega_{\mathrm{D},\mathbf{u}}}=\mathbf{0}\right\},$$
$$\mathscr{V}_\theta = \left\{\theta\in V_\theta \mid \theta|_{\partial\Omega_{\mathrm{D},\theta}}=\theta_\mathrm{g}\right\}, \qquad \mathscr{W}_\theta = \left\{\eta\in V_\theta \mid \eta|_{\partial\Omega_{\mathrm{D},\theta}}=0\right\}.$$

Once Eqs. (6a)–(6c) are multiplied by the corresponding test functions, integrated over the computational domain $\Omega$, second order terms integrated by parts, and the boundary conditions are also taken into account, the resulting variational form of the problem that we consider is given as follows: Find the triplet $[\mathbf{u},p,\theta]:(0,t_\mathrm{f})\to\mathscr{V}_\mathbf{u}\times\mathscr{V}_p\times\mathscr{V}_\theta$ such that,

$$\langle\rho\partial_t\mathbf{u},\mathbf{v}\rangle+\mathscr{A}_{\mathbf{uu}}(\rho;\mathbf{u},\mathbf{u},\mathbf{v})-b(p,\mathbf{v})=\ell_\mathbf{u}(\rho;\mathbf{v}), \tag{10a}$$

$$\langle\beta\partial_t p,q\rangle-\langle\alpha\partial_t\theta,q\rangle+\mathscr{A}_{pp}(\beta;\mathbf{u},p,q)-\mathscr{A}_{p\theta}(\alpha;\mathbf{u},\theta,q)+b(q,\mathbf{u})=0, \tag{10b}$$

$$\langle\rho c_\mathrm{v}\partial_t\theta,\eta\rangle+\mathscr{A}_{\theta\theta}(\rho;\mathbf{u},\theta,\eta)+\mathscr{A}_{\theta p}(\mathbf{u};p,\eta)=\ell_\theta(\rho,\mathbf{u};\eta), \tag{10c}$$

$$\langle\mathbf{u}(\mathbf{x},0),\mathbf{v}\rangle=\langle\mathbf{u}_0(\mathbf{x}),\mathbf{v}\rangle\,, \tag{10d}$$

$$\langle p(\mathbf{x},0),q\rangle=\langle p_0(\mathbf{x}),q\rangle\,, \tag{10e}$$

$$\langle\theta(\mathbf{x},0),\eta\rangle=\langle\theta_0(\mathbf{x}),\eta\rangle\,, \tag{10f}$$

which must hold for all test functions $\mathbf{v} \in \mathscr{W}_{\mathbf{u}}$, $q \in \mathscr{W}_p = \mathscr{V}_p$ and $\eta \in \mathscr{W}_\theta$ and for all $t \in (0, t_{\mathrm{f}})$. Here and in what follows, $\langle f, g \rangle := \int_\Omega fg \, \mathrm{d}\Omega$ where $f$ and $g$ are two generic functions such that the integral of their product is well defined. The different nonlinear and linear forms appearing in the previous weak form are defined as:

$$\mathscr{A}_{\mathbf{uu}}(\rho; \mathbf{a}, \mathbf{u}, \mathbf{v}) = \int_\Omega \rho\,[(\mathbf{a}\cdot\nabla)\mathbf{u}]\cdot\mathbf{v}\,\mathrm{d}\Omega + 2\mu\int_\Omega \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{v})\,\mathrm{d}\Omega - \frac{2}{3}\mu\int_\Omega (\nabla\cdot\mathbf{u})(\nabla\cdot\mathbf{v})\,\mathrm{d}\Omega, \tag{11a}$$

$$b(p, \mathbf{v}) = \int_\Omega p(\nabla\cdot\mathbf{v})\,\mathrm{d}\Omega, \tag{11b}$$

$$\mathscr{A}_{pp}(\beta; \mathbf{a}, p, q) = \int_\Omega \beta[(\mathbf{a}\cdot\nabla)p]q\,\mathrm{d}\Omega, \tag{11c}$$

$$\mathscr{A}_{p\theta}(\alpha; \mathbf{a}, \theta, q) = \int_\Omega \alpha[(\mathbf{a}\cdot\nabla)\theta]q\,\mathrm{d}\Omega, \tag{11d}$$

$$\mathscr{A}_{\theta\theta}(\rho; \mathbf{a}, \theta, \eta) = \int_\Omega \rho c_{\mathrm{v}}[(\mathbf{a}\cdot\nabla)\theta]\eta\,\mathrm{d}\Omega + \int_\Omega \kappa\nabla\theta\cdot\nabla\eta\,\mathrm{d}\Omega, \tag{11e}$$

$$\mathscr{A}_{\theta p}(\mathbf{u}; p, \eta) = \int_\Omega p(\nabla\cdot\mathbf{u})\eta\,\mathrm{d}\Omega, \tag{11f}$$

$$\ell_{\mathbf{u}}(\rho; \mathbf{v}) = \int_\Omega \rho\mathbf{f}\cdot\mathbf{v}\,\mathrm{d}\Omega + \int_{\partial\Omega_{\mathrm{N},\mathbf{u}}} \mathbf{t}\cdot\mathbf{v}\,\mathrm{d}\partial\Omega, \tag{11g}$$

$$\ell_\theta(\rho, \mathbf{u}; \eta) = \int_\Omega \rho r\eta\,\mathrm{d}\Omega + 2\mu\int_\Omega \boldsymbol{\varepsilon}(\mathbf{u}) : \boldsymbol{\varepsilon}(\mathbf{u})\eta\,\mathrm{d}\Omega - \frac{2}{3}\mu\int_\Omega (\nabla\cdot\mathbf{u})^2\eta\,\mathrm{d}\Omega + \int_{\partial\Omega_{\mathrm{N},\theta}} \varphi\eta\,\mathrm{d}\partial\Omega, \tag{11h}$$

where $\mathbf{a}(\mathbf{x}, t)$ represents a general convective velocity.

## 3. Non reflecting boundary conditions

The compressible Navier–Stokes equations represent a direct path to consistently deal with both aerodynamic and acoustic scales at once. As a consequence, acoustic waves and the flow boundary conditions must be treated consistently and a certain compatibility requirement should be introduced. In this section we review the boundary setting described in detail in [41] and we extend its application for the compressible problem in hand. To fix ideas, the main ingredients of the methodology are the weak prescription of essential boundary conditions together with the application of a zero-order non-reflecting boundary condition.

### 3.1. Unknown and boundary splitting

Let us start by considering a splitting of the velocity and pressure fields, respectively into mean and oscillatory components. For a given time instant $t \in [0, t_{\mathrm{f}}]$ and a point in the spatial boundary domain $\mathbf{x} \in \partial\Omega$, we have,

$$\mathbf{u}(\mathbf{x}, t) = \bar{\mathbf{u}}(\mathbf{x}, t) + \mathbf{u}'(\mathbf{x}, t), \qquad p(\mathbf{x}, t) = \bar{p}(\mathbf{x}, t) + p'(\mathbf{x}, t), \tag{12}$$

where the time-average variables are mathematically defined as,

$$\bar{\mathbf{u}}(\mathbf{x}, t) := \frac{1}{T_{\mathrm{w}}}\int_{t-T_{\mathrm{w}}}^{t} \mathbf{u}(\mathbf{x}, \chi)\,\mathrm{d}\chi, \qquad \bar{p}(\mathbf{x}, t) := \frac{1}{T_{\mathrm{w}}}\int_{t-T_{\mathrm{w}}}^{t} p(\mathbf{x}, \chi)\,\mathrm{d}\chi. \tag{13}$$

Hereafter, the oscillatory components correspond to the acoustic fluctuations and the mean components to the flow variables. In this definition, $T_{\mathrm{w}}$ represents an appropriate time window and thus it implicitly defines a filtering frequency for the acoustic waves, which must be chosen small enough to allow a damping of the acoustic perturbations without damaging the flow evolution.

Complementing the previous variable decomposition, we also divide the boundary where velocity conditions exist (either essential or natural) into internal and external contributions. Internal contributions are for instance those corresponding to solid walls in the interior of the domain. External contours correspond to inlet and outlet boundaries. The *external* boundary where velocity shall be prescribed $\partial\Omega^e_{\mathbf{u}}$ is divided into two different disjoint subsets, $\partial\Omega^e_{\mathrm{D},\mathbf{u}}$, $\partial\Omega^e_{\mathrm{N},\mathbf{u}}$. These subsets are such that $\overline{\partial\Omega^e_{\mathrm{D},\mathbf{u}} \cap \partial\Omega^e_{\mathrm{N},\mathbf{u}}} = \emptyset$, and $\overline{\partial\Omega^e_{\mathrm{D},\mathbf{u}} \cup \partial\Omega^e_{\mathrm{N},\mathbf{u}}} = \partial\Omega^e_{\mathbf{u}}$. We remark that this boundary splitting is performed at the external artificial contours of the domain. Should the domain contain any interior wall with prescribed velocity, we should expect sound waves to be reflected in such location and no particular boundary treatment is needed. On both $\partial\Omega^e_{\mathrm{D},\mathbf{u}}$ and $\partial\Omega^e_{\mathrm{N},\mathbf{u}}$, which are in the far field, it is assumed that the acoustic scales are dominant.

**Remark 3.1.** The key idea behind our boundary formulation is the introduction of the so-called Sommerfeld boundary condition. Such condition is derived from the wave equation written in mixed form, which represents a set of wave-like equations for $\mathbf{u}(\mathbf{x},t)$ and $p(\mathbf{x},t)$ [42]. Some might argue that there exits a wave-like equation for heat. However, following the ideas in [43], it is easily shown that although temperature oscillations may have the mathematical expression of a wave, they are not real travelling waves as there is no actual transport of energy, and hence they show neither wave fronts nor reflection/refraction phenomena. As a result, there is no need to introduce any splitting for the temperature into mean and oscillatory components and the only contribution to the weak form from the imposition of boundary conditions in the temperature equation would be Eq. (8d).

### 3.2. *Unified prescription of boundary conditions*

In this subsection, we summarize the different conditions to be applied on each boundary.

On the frontiers belonging to the truncation boundary $\partial\Omega^e_{\mathrm{D},\mathbf{u}}$, distinct conditions are enforced:

- The mean value of the velocity is prescribed to the flow given velocity,

$$\bar{\mathbf{u}} = \mathbf{u}_{\mathrm{g}} \quad \text{on} \quad \partial\Omega^e_{\mathrm{D},\mathbf{u}}, \tag{14}$$

  and this will be done weakly via Nitsche's method.
- A Sommerfeld-like non-reflecting boundary condition is prescribed for the acoustic component of the velocity field. In the normal direction to the boundary we set,

$$\mathbf{n} \cdot [\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}', p')] = -\sqrt{\frac{\rho}{\beta}}\mathbf{n} \cdot \mathbf{u}' \quad \text{on} \quad \partial\Omega^e_{\mathrm{D},\mathbf{u}}, \tag{15}$$

  and for the tangential direction we directly write,

$$\mathbf{m} \cdot [\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}', p')] = 0 \quad \text{on} \quad \partial\Omega^e_{\mathrm{D},\mathbf{u}}, \tag{16}$$

  for any vector $\boldsymbol{m}$ in the tangent direction to $\partial\Omega^e_{\mathrm{D},\mathbf{u}}$.

Finally, on the boundary $\partial\Omega^e_{\mathrm{N},\mathbf{u}}$, the following conditions are enforced:

- The mean value tractions are prescribed, i.e.,

$$\mathbf{n} \cdot \boldsymbol{\sigma}(\bar{\mathbf{u}}, \bar{p}) = \mathbf{t} \quad \text{on} \quad \partial\Omega^e_{\mathrm{N},\mathbf{u}}. \tag{17}$$

- The same approach as in $\partial\Omega^e_{\mathrm{D},\mathbf{u}}$ is used now for the fluctuating component. Therefore,

$$\mathbf{n} \cdot [\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}', p')] = -\sqrt{\frac{\rho}{\beta}}\mathbf{n} \cdot \mathbf{u}' \quad \text{on} \quad \partial\Omega^e_{\mathrm{N},\mathbf{u}}, \tag{18}$$

$$\mathbf{m} \cdot [\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}', p')] = 0 \qquad \text{on} \quad \partial\Omega^e_{\mathrm{N},\mathbf{u}}. \tag{19}$$

Taking now into account these definitions, the prescription of boundary conditions in the weak form of the problem can be done upon the modification of the boundary term on the right-hand-side of the momentum equation Eq. (10a), which after the introduction of the symmetric and the penalty terms for the imposition of $\bar{\mathbf{u}} = \mathbf{u}_{\mathrm{g}}$ using Nitsche's method reads,

$$\begin{aligned}\int_{\partial\Omega^e_{\mathbf{u}}} [\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{u}, p)] \cdot \mathbf{v}\, \mathrm{d}\partial\Omega = & \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} [\mathbf{n} \cdot \boldsymbol{\sigma}(\bar{\mathbf{u}}, \bar{p})] \cdot \mathbf{v}\, \mathrm{d}\partial\Omega - \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \sqrt{\frac{\rho}{\beta}}(\mathbf{u}' \cdot \mathbf{n})(\mathbf{v} \cdot \mathbf{n})\, \mathrm{d}\partial\Omega \\ & + \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} (\bar{\mathbf{u}} - \mathbf{u}_{\mathrm{g}}) \cdot [\mathbf{n} \cdot \boldsymbol{\sigma}(\mathbf{v}, q)]\, \mathrm{d}\partial\Omega - \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \zeta(\bar{\mathbf{u}} - \mathbf{u}_{\mathrm{g}}) \cdot \mathbf{v}\, \mathrm{d}\partial\Omega \\ & - \int_{\partial\Omega^e_{\mathrm{N},\mathbf{u}}} \sqrt{\frac{\rho}{\beta}}(\mathbf{u}' \cdot \mathbf{n})(\mathbf{v} \cdot \mathbf{n})\, \mathrm{d}\partial\Omega + \int_{\partial\Omega^e_{\mathrm{N},\mathbf{u}}} \mathbf{t} \cdot \mathbf{v}\, \mathrm{d}\partial\Omega, \end{aligned} \tag{20}$$

$\zeta$ being the numerical penalty parameter.

If we now group these boundary terms in the following forms,

$$\begin{aligned}
\mathscr{A}_{\mathbf{uu}}^{\partial\Omega}(\rho,\beta;\mathbf{u},\mathbf{v}) = {} & \int_{\partial\Omega^e_{\mathbf{u}}} \sqrt{\frac{\rho}{\beta}}(\mathbf{u}'\cdot\mathbf{n})(\mathbf{v}\cdot\mathbf{n})\,\mathrm{d}\partial\Omega + \int_{\partial\Omega_{\mathrm{D},\mathbf{u}}} \zeta\bar{\mathbf{u}}\cdot\mathbf{v}\,\mathrm{d}\partial\Omega - 2\mu\int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} [\mathbf{n}\cdot\boldsymbol{\varepsilon}(\bar{\mathbf{u}})]\cdot\mathbf{v}\,\mathrm{d}\partial\Omega \\
& + \frac{2}{3}\mu\int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \nabla\cdot\bar{\mathbf{u}}(\mathbf{n}\cdot\mathbf{v})\,\mathrm{d}\partial\Omega - 2\mu\int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} [\mathbf{n}\cdot\boldsymbol{\varepsilon}(\mathbf{v})]\cdot\bar{\mathbf{u}}\,\mathrm{d}\partial\Omega + \frac{2}{3}\mu\int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \nabla\cdot\mathbf{v}(\mathbf{n}\cdot\bar{\mathbf{u}})\,\mathrm{d}\partial\Omega, \qquad (21a)
\end{aligned}$$

$$b^{\partial\Omega}(p,\mathbf{v}) = \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \bar{p}(\mathbf{n}\cdot\mathbf{v})\,\mathrm{d}\partial\Omega, \qquad (21b)$$

$$\begin{aligned}
\ell_{\mathbf{u}}^{\partial\Omega}(\mathbf{v}) = {} & \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \zeta\mathbf{u}_{\mathrm{g}}\cdot\mathbf{v}\,\mathrm{d}\partial\Omega - 2\mu\int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} [\mathbf{n}\cdot\boldsymbol{\varepsilon}(\mathbf{v})]\cdot\mathbf{u}_{\mathrm{g}}\,\mathrm{d}\partial\Omega + \frac{2}{3}\mu\int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} \nabla\cdot\mathbf{v}(\mathbf{n}\cdot\mathbf{u}_{\mathrm{g}})\,\mathrm{d}\partial\Omega \\
& + \int_{\partial\Omega^e_{\mathrm{N},\mathbf{u}}} \mathbf{v}\cdot\mathbf{t}\,\mathrm{d}\partial\Omega, \qquad (21c)
\end{aligned}$$

$$\ell_{p}^{\partial\Omega}(q) = \int_{\partial\Omega^e_{\mathrm{D},\mathbf{u}}} q(\mathbf{n}\cdot\mathbf{u}_{\mathrm{g}})\,\mathrm{d}\partial\Omega, \qquad (21d)$$

the weak formulation of the Navier–Stokes compressible primitive problem would now consist in seeking the velocity, the pressure and the temperature satisfying,

$$\langle\rho\partial_t\mathbf{u},\mathbf{v}\rangle + \mathscr{A}_{\mathbf{uu}}(\rho;\mathbf{u},\mathbf{u},\mathbf{v}) - b(p,\mathbf{v}) + \mathscr{A}_{\mathbf{uu}}^{\partial\Omega}(\rho,\beta;\mathbf{u},\mathbf{v}) + b^{\partial\Omega}(p,\mathbf{v}) = \ell_{\mathbf{u}}(\rho;\mathbf{v}) + l_{\mathbf{u}}^{\partial\Omega}(\mathbf{v}), \qquad (22a)$$

$$\langle\beta\partial_t p,q\rangle - \langle\alpha\partial_t\theta,q\rangle + \mathscr{A}_{pp}(\beta;\mathbf{u},p,q) - \mathscr{A}_{p\theta}(\alpha;\mathbf{u},\theta,q) + b(q,\mathbf{u}) + b^{\partial\Omega}(q,\mathbf{u}) = \ell_{p}^{\partial\Omega}(q), \qquad (22b)$$

$$\langle\rho c_{\mathrm{v}}\partial_t\theta,\eta\rangle + \mathscr{A}_{\theta\theta}(\rho;\mathbf{u},\theta,\eta) + \mathscr{A}_{\theta p}(\mathbf{u},p,\eta) = \ell_{\theta}(\rho,\mathbf{u};\eta), \qquad (22c)$$

for all test functions and for all $t\in(0,t_{\mathrm{f}})$, and fulfilling weakly the initial conditions at $t=0$.

## 4. Numerical approximation

In this section, we study the discretization of the problem from Eqs. (22a)–(22c). For the time discretization we proposed backward differencing schemes (BDF). For the spatial approximation we will use in this section the standard Galerkin finite element method. For the sake of simplicity, we do not pose any restriction on the finite element spaces and let us assume for the moment that the raw Galerkin method provides meaningful results. However, this might not be the case when convection is dominant, or if there is some kind of incompatibility among the interpolation spaces of the different unknowns of the problem. The introduction of stabilization techniques is deferred until Section 6, where we present the stabilized finite element formulation.

### *4.1. Time discretization*

The problem can be discretized in time using a uniform partition of the time interval. Given a time step $\Delta t\in(0,t_{\mathrm{f}})$, we set $t^0=0$, $t^n=t^0+n\Delta t$ (for any $n\geq 1$), and $N=[t_{\mathrm{f}}/\Delta t]$. Denoting by $f^n$ an approximation to a time-dependent function $f(t)$ at time $t^n$, we consider backward difference (BDF) schemes with the operator,

$$D_k f^{n+1} = \frac{1}{\phi_k}\left(f^{n+1} - \sum_{i=0}^{k-1}\xi_k^i f^{n-i}\right),$$

being $\phi_k$ and $\xi_k^i$ numerical parameters depending on the order of the temporal approximation. In particular, for the first and second order schemes, i.e., $k=1,2$, it is found that:

$$\begin{aligned}
D_1 f^{n+1} &= \delta f^{n+1} = f^{n+1} - f^n, \\
D_2 f^{n+1} &= \frac{3}{2}\left(f^{n+1} - \frac{4}{3}f^n + \frac{1}{3}f^{n-1}\right).
\end{aligned}$$

For the design of fractional step schemes, it is particularly useful to make use of the backward extrapolation operators,

$$\hat{f}_0^{n+1} = 0, \tag{23a}$$

$$\hat{f}_1^{n+1} = f^n, \tag{23b}$$

$$\hat{f}_2^{n+1} = 2f^n - f^{n-1}, \tag{23c}$$

or, in general,

$$\hat{f}_k^{n+1} := f^{n+1} - \delta^{(k)} f^{n+1} = f^{n+1} + \mathscr{O}(\Delta t^k), \; k = 0, 1, 2 \dots \tag{24}$$

with $\delta^{(i+1)} f^{n+1} := \delta^{(i)} f^{n+1} - \delta^{(i)} f^n, \; i = 1, 2, 3, \dots$.

As the reader might have noticed, discrete expressions to compute the time-average components of $\mathbf{u}(\mathbf{x}, t)$ and $p(\mathbf{x}, t)$ still need to be provided. In this sense, we will consider the trapezoidal rule and the expression we use for the average values is,

$$\bar{\mathbf{u}}_h^{n+1} = \frac{\Delta t}{T_{\mathrm{w}}} \left( \frac{1}{2} \mathbf{u}_h^{n+1} + \sum_{j=n-N_{\mathrm{w}}+2}^{n} \mathbf{u}_h^j + \frac{1}{2} \mathbf{u}_h^{n-N_{\mathrm{w}}+1} \right),$$

and equivalently for the pressure. This relation maintains the integration implicit and second order accurate, but several time steps need to be run prior to its application so as to obtain representative data for a reliable mean computation. Finally, at the discrete level, the time window introduced in Eq. (13) is computed as $T_{\mathrm{w}} = N_{\mathrm{w}} \Delta t$ being $N_{\mathrm{w}}$ a certain amount of time steps.

### 4.2. Finite element discretization

Let $\mathscr{T}_h$ be a shape-regular and conforming partition of $\Omega$, such that $\overline{\Omega} = \cup_{e=1}^{n_{\mathrm{el}}} \overline{\Omega}^{(e)}$ being $n_{\mathrm{el}}$ the total number of elements in the partition. This triangulation is described by the characteristic mesh size, defined as $h := \max\{h^{(e)} \mid \Omega^{(e)} \in \mathscr{T}_h(\Omega)\}$ with $h^{(e)} = \mathrm{diam}(\Omega^{(e)})$. Let now $\mathscr{V}_{\mathbf{u},h} \subset \mathscr{V}_{\mathbf{u}}$, $\mathscr{V}_{p,h} \subset \mathscr{V}_p$ and $\mathscr{V}_{\theta,h} \subset \mathscr{V}_\theta$ be the velocity, pressure and temperature finite element spaces associated with the triangulation. In the following, finite element functions will be identified with a subscript $h$.

The discrete problem is obtained by approximating $\mathbf{u}$, $p$ and $\theta$. We assume that $\mathbf{u}_h^n$, $p_h^n$, and $\theta_h^n$ are constructed using the standard finite element interpolation from the nodal values, which we denote hereafter as $\mathbf{U}^n$, $\mathbf{P}^n$ and $\boldsymbol{\Theta}^n$, respectively. These are computed as the solution of a nonlinear algebraic system, which is the fully discrete counterpart of Eqs. (22a)–(22c), and it has the classical form $\mathbf{Ax} = \mathbf{f}$,

$$\begin{bmatrix} \mathbf{A}_{\mathbf{uu}} & \mathbf{A}_{\mathbf{u}p} & \mathbf{0} \\ \mathbf{A}_{p\mathbf{u}} & \mathbf{A}_{pp} & \mathbf{A}_{p\theta} \\ \mathbf{0} & \mathbf{A}_{\theta p} & \mathbf{A}_{\theta\theta} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{U}^{n+1} \\ \mathbf{P}^{n+1} \\ \boldsymbol{\Theta}^{n+1} \end{bmatrix} = \begin{bmatrix} \mathbf{F}_{\mathbf{u}}^{n+1} \\ \mathbf{F}_p^{n+1} \\ \mathbf{F}_\theta^{n+1} \end{bmatrix} \tag{25}$$

with,

$$\mathbf{A}_{\mathbf{uu}} := \mathbf{M}_{\mathbf{uu}}(\rho) \frac{D_k}{\Delta t} + \mathbf{K}_{\mathbf{uu}}(\rho, \mathbf{U}^{n+1}) + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega} + \mathbf{K}_{\mathbf{uu}}^{\partial\Omega}(\rho, \beta) \qquad \mathbf{A}_{\mathbf{u}p} := -\mathbf{K}_{\mathbf{u}p} + \mathbf{K}_{\mathbf{u}p}^{\partial\Omega} \tag{26a}$$

$$\mathbf{A}_{p\mathbf{u}} := \mathbf{K}_{p\mathbf{u}} + \mathbf{K}_{p\mathbf{u}}^{\partial\Omega} \qquad \mathbf{A}_{pp} := \mathbf{M}_{pp}(\beta) \frac{D_k}{\Delta t} + \mathbf{C}_{pp}(\beta, \mathbf{U}^{n+1}) \tag{26b}$$

$$\mathbf{A}_{p\theta} := -\mathbf{M}_{p\theta}(\alpha) \frac{D_k}{\Delta t} - \mathbf{C}_{p\theta}(\alpha, \mathbf{U}^{n+1}) \qquad \mathbf{A}_{\theta p} := \mathbf{K}_{\theta p} \tag{26c}$$

$$\mathbf{A}_{\theta\theta} := \mathbf{M}_{\theta\theta}(\rho) \frac{D_k}{\Delta t} + \mathbf{K}_{\theta\theta}(\rho, \mathbf{U}^{n+1}) \qquad \mathbf{F}_{\mathbf{u}}^{n+1} := \mathbf{R}_{\mathbf{u}}^{n+1}(\rho) + \mathbf{R}_{\mathbf{u}}^{\partial\Omega, n+1} \tag{26d}$$

$$\mathbf{F}_p^{n+1} := \mathbf{R}_p^{\partial\Omega, n+1} \qquad \mathbf{F}_\theta^{n+1} := \mathbf{R}_\theta^{n+1}(\rho, \mathbf{U}^{n+1}) \tag{26e}$$

The first subscript on the arrays refers to the momentum ($\mathbf{u}$), energy ($\theta$) and continuity ($p$) equation, and the second stands for the unknown to which the term refers to. The high non-linear character of the problem is made explicit in the system by including the dependence of the arrays on the variables in the parenthesis. The symbol $\mathbf{M}$ stands for the mass matrices and, in addition, the superscript $\partial\Omega$ stands for the terms arising from the special treatment of boundary conditions, as described in Section 3. All the arrays in the system are computed from the local assembly of the elemental contributions and the right-hand-side terms in Eqs. (26a)–(26e) have a one-to-one correspondence with the discrete version of the forms in Eqs. (22a)–(22c).

**Remark 4.1.** We want to remark that the unknowns of the problem are $\mathbf{u}_h^{n+1}$, $p_h^{n+1}$ and $\theta_h^{n+1}$ and that $\rho$, $\beta$ and $\alpha$ would be explicitly computed in the final algorithm by means of some linearization process. Hence their values are obtained from finite element quantities but not solved as unknowns. In other words, $\rho$, $\beta$ and $\alpha$ *do not* belong to the finite element spaces.

## 5. Fractional step methods

In this section we develop the algebraic fractional step methods for the compressible Navier–Stokes problem. The algorithms here proposed could be viewed as a natural extension of pressure-segregation schemes previously developed for viscoelastic [21] or incompressible flow problems [19]. The basic procedure entails a calculation of an intermediate velocity with a guess of the pressure, yet we will broaden here the original technique by computing also an intermediate temperature. After computing the pressure, we will finally correct the velocity and temperature intermediate values, so that to ensure that the global time accuracy of the method is maintained. We will not discuss here different segregation techniques such as velocity-correction methods [44] (based on the opposite procedure, a velocity guess is assumed to solve for the pressure) or predictor-multicorrector techniques [29].

### 5.1. The algebraic viewpoint: extrapolation

Let us start by writing the system in Eq. (25) in the following *equivalent* form,

$$\mathbf{M}_{\mathbf{uu}}(\rho)\frac{D_k}{\Delta t}\widetilde{\mathbf{U}}^{n+1} + \mathbf{K}_{\mathbf{uu}}(\rho,\mathbf{U}^{n+1})\mathbf{U}^{n+1} - \mathbf{K}_{\mathbf{u}p}\widehat{\mathbf{P}}^{n+1}_{k-1} + \mathbf{M}^{\partial\Omega}_{\mathbf{uu}}\mathbf{U}^{n+1} + \mathbf{K}^{\partial\Omega}_{\mathbf{uu}}(\rho,\beta)\mathbf{U}^{n+1} + \mathbf{K}^{\partial\Omega}_{\mathbf{u}p}\mathbf{P}^{n+1} = \mathbf{F}^{n+1}_{\mathbf{u}}, \tag{27a}$$

$$\mathbf{M}_{\mathbf{uu}}(\rho)\frac{1}{\phi_k\Delta t}(\mathbf{U}^{n+1} - \widetilde{\mathbf{U}}^{n+1}) - \mathbf{K}_{\mathbf{u}p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}^{n+1}_{k-1}) = \mathbf{0}, \tag{27b}$$

$$\mathbf{M}_{\theta\theta}(\rho)\frac{D_k}{\Delta t}\widetilde{\boldsymbol{\Theta}}^{n+1} + \mathbf{K}_{\theta\theta}(\rho,\mathbf{U}^{n+1})\boldsymbol{\Theta}^{n+1} + \mathbf{K}_{\theta p}(\mathbf{U}^{n+1})\widehat{\mathbf{P}}^{n+1}_{k-1} = \mathbf{F}^{n+1}_{\theta}, \tag{27c}$$

$$\mathbf{M}_{\theta\theta}(\rho)\frac{1}{\phi_k\Delta t}(\boldsymbol{\Theta}^{n+1} - \widetilde{\boldsymbol{\Theta}}^{n+1}) + \mathbf{K}_{\theta p}(\mathbf{U}^{n+1})(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}^{n+1}_{k-1}) = \mathbf{0}, \tag{27d}$$

$$\mathbf{M}_{pp}(\beta)\frac{D_k}{\Delta t}\mathbf{P}^{n+1} + \mathbf{C}_{pp}(\beta,\mathbf{U}^{n+1})\mathbf{P}^{n+1} - \mathbf{M}_{p\theta}(\alpha)\frac{D_k}{\Delta t}\boldsymbol{\Theta}^{n+1} - \mathbf{C}_{p\theta}(\alpha,\mathbf{U}^{n+1})\boldsymbol{\Theta}^{n+1} + \mathbf{K}_{p\mathbf{u}}\mathbf{U}^{n+1} + \mathbf{K}^{\partial\Omega}_{p\mathbf{u}}\mathbf{U}^{n+1} = \mathbf{F}^{n+1}_{p}, \tag{27e}$$

where $D_k\tilde{f}^{n+1}$ is computed as $D_k f^{n+1}$ but replacing $f^{n+1}$ by a yet undetermined function $\tilde{f}^{n+1}$, for $f = \mathbf{U}$ and $f = \boldsymbol{\Theta}$. The chosen ordering of the equations is consistent with the steps of the fractional step algorithm proposed below. The reader should note that adding up Eqs. (27a)–(27b) we obtain the former momentum equation, i.e. the first row of the system in Eq. (25). Adding up Eqs. (27c)–(27d) we recover the original energy equation, i.e. the last raw in Eq. (25). We shall refer to Eqs. (27a) and (27c) as the intermediate momentum and intermediate energy equations, respectively and to Eqs. (27b)–(27d) as the correction momentum and energy equations. Similarly, the auxiliary variables $\widetilde{\mathbf{U}}^{n+1}$ and $\widetilde{\boldsymbol{\Theta}}^{n+1}$ are the so-called intermediate velocity and intermediate temperature. Likewise, $\widehat{\mathbf{P}}^{n+1}_{k-1}$ is an extrapolation of the pressure of order $k-1$ at time step $n+1$. Note that we introduce a pressure extrapolation of one order less than the general time integration scheme, but we re-include this extrapolated term in the correction equations.

If we now multiply Eq. (27b) by $\mathbf{K}_{p\mathbf{u}}\mathbf{M}^{-1}_{\mathbf{uu}}$, we get,

$$\mathbf{K}_{p\mathbf{u}}\mathbf{U}^{n+1} = \mathbf{K}_{p\mathbf{u}}\widetilde{\mathbf{U}}^{n+1} + \phi_k\Delta t\mathbf{K}_{p\mathbf{u}}\mathbf{M}^{-1}_{\mathbf{uu}}\mathbf{K}_{\mathbf{u}p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}^{n+1}_{k-1}),$$

and multiplying Eq. (27d) by $\mathbf{C}_{p\theta}\mathbf{M}^{-1}_{\theta\theta}$,

$$\mathbf{C}_{p\theta}\boldsymbol{\Theta}^{n+1} = \mathbf{C}_{p\theta}\widetilde{\boldsymbol{\Theta}}^{n+1} - \phi_k\Delta t\mathbf{C}_{p\theta}\mathbf{M}^{-1}_{\theta\theta}\mathbf{K}_{\theta p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}^{n+1}_{k-1}),$$

and both expressions can be used in the equation for the pressure Eq. (27e) and this yields,

$$\begin{aligned}\mathbf{M}_{pp}(\beta)\frac{D_k}{\Delta t}\mathbf{P}^{n+1} &+ \mathbf{C}_{pp}(\beta,\mathbf{U}^{n+1})\mathbf{P}^{n+1} - \mathbf{M}_{p\theta}(\alpha)\frac{D_k}{\Delta t}\boldsymbol{\Theta}^{n+1} - \mathbf{C}_{p\theta}(\alpha,\mathbf{U}^{n+1})\widetilde{\boldsymbol{\Theta}}^{n+1}\\ &+ \mathbf{K}_{p\mathbf{u}}\widetilde{\mathbf{U}}^{n+1} + \phi_k\Delta t\mathbf{C}_{p\theta}\mathbf{M}^{-1}_{\theta\theta}\mathbf{K}_{\theta p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}^{n+1}_{k-1})\\ &+ \phi_k\Delta t\mathbf{K}_{p\mathbf{u}}\mathbf{M}^{-1}_{\mathbf{uu}}\mathbf{K}_{\mathbf{u}p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}^{n+1}_{k-1}) + \mathbf{K}^{\partial\Omega}_{p\mathbf{u}}\mathbf{U}^{n+1} = \mathbf{F}^{n+1}_{p}.\end{aligned}$$

At this point, several remarks are in order, since we have modified the original matrix version of the continuity equation by introducing some burden related to appearance of the inverse of the mass matrices $\mathbf{M}_{\theta\theta}$ and $\mathbf{M}_{\mathbf{uu}}$.

**Remark 5.1.** One should notice that the resulting matrix from $\mathbf{K}_{p\mathbf{u}}\mathbf{M}^{-1}_{\mathbf{uu}}\mathbf{K}_{\mathbf{u}p}$ can be viewed as an approximation to the discrete version of a Laplacian operator, [45]. In order to avoid dealing with this matrix, which is in general dense and might still be expensive to compute even when $\mathbf{M}_{\mathbf{uu}}$ is lumped, we use the approximation $\mathbf{K}_{p\mathbf{u}}\mathbf{M}^{-1}_{\mathbf{uu}}\mathbf{K}_{\mathbf{u}p} \approx \mathbf{L}$ where $\mathbf{L}$ is a Laplacian matrix obtained using the gradient of the standard shape functions $\Phi$ and computed at the element level as $[\mathbf{L}^{(e)}]^{ab} = \int_{\Omega^{(e)}} \frac{1}{\rho}\nabla\Phi^a \cdot \nabla\Phi^b \, \mathrm{d}\Omega$.

**Remark 5.2.** The case of $\mathbf{C}_{p\theta}\mathbf{M}_{\theta\theta}^{-1}\mathbf{K}_{\theta p}$ deserves an special comment. We need to take into account that, of course, we want to avoid computing inverse of matrices by all means. We approximate *the effect* of the term by $\mathbf{C}_{p\theta}\mathbf{M}_{\theta\theta}^{-1}\mathbf{K}_{\theta p} \approx \mathbf{Q}$, computed in each element as $[\mathbf{Q}^{(e)}]^{ab} = \int_{\Omega^{(e)}} \frac{\alpha}{\rho c_{\mathrm{v}}} (\mathbf{u}_h \cdot \nabla)\Phi^a (\nabla \cdot \mathbf{u}_h)\Phi^b \, \mathrm{d}\Omega$ and where the first factor in the integral is introduced to keep the proper scaling.

**Remark 5.3.** Note from the definition of the extrapolation operators, Eq. (24), that the difference $||\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}||$ is of order $\mathscr{O}(\Delta t^{k-1})$. Therefore, it is easy to see from Eqs. (27b) and (27d) that $\mathscr{O}(\|\mathbf{U}^{n+1} - \widetilde{\mathbf{U}}^{n+1}\|) = \mathscr{O}(\Delta t^k)$, $\mathscr{O}(\|\mathbf{\Theta}^{n+1} - \widetilde{\mathbf{\Theta}}^{n+1}\|) = \mathscr{O}(\Delta t^k)$ and thus, the global accuracy of the temporal integrator is formally maintained.

### 5.2. Pressure-correction scheme

Generally speaking, the fractional step approach that we favor to solve the fully compressible Navier–Stokes problem in primitive variables has five main steps: (i) compute an intermediate velocity from Eq. (27a), (ii) compute an intermediate temperature from Eq. (27c), (iii) compute an approximation to the end-of-step pressure by using Eq. (27e), (iv) update the end-of-step velocity with Eq. (27b), and (v) update the end-of-step temperature by using Eq. (27d). This procedure will make possible to segregate the calculation of the unknowns of the problem and provides a pressure-correction-like algorithm.

In Eqs. (27a)-(27e) there are still some terms which involve some coupling of the unknowns and hence some extra information is still needed in order to complete the algorithm. One of the key approximations we introduce to decouple the calculation of the problem unknowns is to substitute $\mathbf{U}^{n+1}$ by $\widetilde{\mathbf{U}}^{n+1}$ and $\mathbf{\Theta}^{n+1}$ by $\widetilde{\mathbf{\Theta}}^{n+1}$ in the rest of the terms in the intermediate momentum and energy equations, respectively Eqs. (27a) and (27c). This approximation is supported by Remark 5.3. In addition to this, the boundary pressure term in Eq. (27a) is made explicit by considering an extrapolation of order $k$, and hence it is not accounted for in any correction step. The remaining terms which could still couple the problem variables can be evaluated either with $\widetilde{\mathbf{U}}^{n+1}$ or $\widetilde{\mathbf{\Theta}}^{n+1}$ and the possible non-linearities are solved by taking the known values of the unknowns from the previous iteration, time step or from the intermediate equations. Taking all this information into account, the final algorithmic procedure is included in Algorithm 1.

**Algorithm 1:** Fractional-step method, order $k = 1, 2$.

```
• Read initial conditions $\mathbf{U}_0$, $\mathbf{P}_0$, $\mathbf{\Theta}_0$
• Set $n = 0$
• while $n < N$ do
    $n \leftarrow n + 1$ ;
        (1) Compute the intermediate velocity $\widetilde{\mathbf{U}}^{n+1}$ from:
            $\mathbf{M}_{\mathbf{uu}}(\rho)\frac{D_k}{\Delta t}\widetilde{\mathbf{U}}^{n+1} + \mathbf{K}_{\mathbf{uu}}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{U}}^{n+1} + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1} + \mathbf{K}_{\mathbf{uu}}^{\partial\Omega}(\rho, \beta)\widetilde{\mathbf{U}}^{n+1} = \mathbf{F}_{\mathbf{u}}^{n+1} + \mathbf{K}_{\mathbf{u}p}\widehat{\mathbf{P}}_{k-1}^{n+1} - \mathbf{K}_{\mathbf{u}p}^{\partial\Omega}\widehat{\mathbf{P}}_{k}^{n+1}$
        (2) Compute the intermediate temperature $\widetilde{\mathbf{\Theta}}^{n+1}$ after solving:
            $\mathbf{M}_{\theta\theta}(\rho)\frac{D_k}{\Delta t}\widetilde{\mathbf{\Theta}}^{n+1} + \mathbf{K}_{\theta\theta}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{\Theta}}^{n+1} = \mathbf{F}_{\theta}^{n+1} - \mathbf{K}_{\theta p}(\widetilde{\mathbf{U}}^{n+1})\widehat{\mathbf{P}}_{k-1}^{n+1}$
        (3) Compute the pressure $\mathbf{P}^{n+1}$ using the intermediate velocity and temperature:
            $\mathbf{M}_{pp}(\beta)\frac{D_k}{\Delta t}\mathbf{P}^{n+1} + \mathbf{C}_{pp}(\beta, \widetilde{\mathbf{U}}^{n+1})\mathbf{P}^{n+1} + \phi_k \Delta t \mathbf{Q}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1})$
                $+ \phi_k \Delta t \mathbf{L}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}) = \mathbf{F}_{p}^{n+1} - \mathbf{K}_{p\mathbf{u}}\widetilde{\mathbf{U}}^{n+1}$
                $+ \mathbf{C}_{p\theta}(\alpha, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{\Theta}}^{n+1} + \mathbf{M}_{p\theta}(\alpha)\frac{D_k}{\Delta t}\widetilde{\mathbf{\Theta}}^{n+1} - \mathbf{K}_{p\mathbf{u}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1}$
        (4) Velocity correction to obtain the end-of-step velocity $\mathbf{U}^{n+1}$:
            $\mathbf{M}_{\mathbf{uu}}(\rho)\frac{1}{\phi_k \Delta t}\mathbf{U}^{n+1} + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega}\mathbf{U}^{n+1} = \mathbf{M}_{\mathbf{uu}}(\rho)\frac{1}{\phi_k \Delta t}\widetilde{\mathbf{U}}^{n+1} + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1} + \mathbf{K}_{\mathbf{u}p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}).$
        (5) Temperature correction to obtain the end-of-step temperature $\mathbf{\Theta}^{n+1}$:
            $\mathbf{M}_{\theta\theta}(\rho)\frac{1}{\phi_k \Delta t}\mathbf{\Theta}^{n+1} = \mathbf{M}_{\theta\theta}(\rho)\frac{1}{\phi_k \Delta t}\widetilde{\mathbf{\Theta}}^{n+1} - \mathbf{K}_{\theta p}(\widetilde{\mathbf{U}}^{n+1})(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1})$
  end
```

The inclusion of the penalization matrix $\mathbf{M}_{\mathbf{uu}}^{\partial\Omega}$ in the velocity correction step aids to properly impose the boundary conditions of the problem avoiding boundary instabilities. It seems reasonable to take it into account, bearing in mind that the splitting of the momentum equation needs to be done considering boundary conditions, similarly to the case in which boundary conditions are enforced strongly. It is also important to note that this matrix displays a structure of mass matrix but for boundary contributions, what in turn would allow to solve directly the system for $\mathbf{U}^{n+1}$ if a lumping technique is used.

It is well known that the extrapolation of second order of the term $\mathbf{K}_{\mathbf{u}p}\mathbf{P}^{n+1}$ is unstable. In fact, this issue motivated the study of other methods to upgrade the temporal order of the fractional step scheme, such as the Yosida regularization technique (see e.g. [19]). On the contrary, we did not observe any erratic behavior of the term $\mathbf{K}^{\partial\Omega}_{\mathbf{u}p}\widehat{\mathbf{P}}^{n+1}_{k}$ when $k=2$ is chosen.

## 6. Variational multiscale stabilized formulation

This section is devoted to the discussion of Variational Multiscale method applied to the compressible Navier–Stokes problem. The basic idea of the VMS framework is to approximate the effect of the part of the solution of the continuous problem which cannot be reproduced by the finite element space. We shall start by presenting the procedure for a general nonlinear evolution problem, and later particularize it to the problem in hand. The presentation here is concise but formal and we avoid the introduction of excessive technicalities and details.

### 6.1. The subgrid scale framework

Let us consider a general evolution system of the form,

$$\boldsymbol{M}(\mathbf{y})\partial_t\mathbf{y}+\mathscr{L}(\mathbf{y};\mathbf{y})=\boldsymbol{F}, \tag{28}$$

in which the majority of problems in fluid mechanics may be cast, and where $\mathbf{y}$ contains the unknowns of the problem, (i.e. $[\mathbf{u},p,\theta]^{\mathsf{T}}$ in this case), $\boldsymbol{M}(\mathbf{y})$ is an operator containing coefficients related to the temporal derivatives (in general non-linear), and $\boldsymbol{F}$ is an external force vector. Assume homogeneous boundary conditions for the sake of simplicity. The spatial operator $\mathscr{L}(\mathbf{y};\mathbf{y})$, linear in the second argument, is clearly a convection-diffusion operator for the problem of interest, such that $\mathscr{L}(\mathbf{y};\mathbf{y})=\boldsymbol{A}_i(\mathbf{y})\partial_i\mathbf{y}-\partial_i(\boldsymbol{K}_{ij}(\mathbf{y})\partial_j\mathbf{y})$ where $\boldsymbol{A}_i$, $\boldsymbol{K}_{ij}$ are the classical $n_{\rm sd}+2\times n_{\rm sd}+2$ convective and diffusive matrices whose identification will be omitted here. Hereafter, repeated indexes imply summation up to the number of spatial dimensions $n_{\rm sd}$ and $\partial_i$ is the partial derivative with respect to the Cartesian coordinate $x_i$.

Let matrices $\mathbf{A}_i(\mathbf{y})$ be split as $\mathbf{A}^{\rm c}_i$ and $\mathbf{A}^{\rm f}_i$, where $\mathbf{A}^{\rm c}_i$ is the part of the convection matrices which is *not* integrated by parts and $\mathbf{A}^{\rm f}_i$ the part that *is* integrated by parts. In our case, matrices $\mathbf{A}^{\rm f}_i$ come from the first order derivatives of the pressure.

Let us denote by $\mathscr{Y}$ the functional space where the solution is to be sought. Then, the weak form of the problem consists in finding $\mathbf{y}:(0,t_{\rm f})\to\mathscr{Y}$, such that,

$$\mathscr{B}(\mathbf{y};\mathbf{y},\mathbf{z})-\ell(\mathbf{z})=0, \tag{29}$$

for any test functions $\mathbf{z}\in\mathscr{Y}$ with appropriate regularity and where,

$$\begin{aligned}\mathscr{B}(\mathbf{y}_0;\mathbf{y},\mathbf{z}) &:= \int_\Omega \mathbf{z}^{\mathsf{T}}\boldsymbol{M}(\mathbf{y}_0)\partial_t\mathbf{y}\,\mathrm{d}\Omega+\int_\Omega \mathbf{z}^{\mathsf{T}}\boldsymbol{A}^{\rm c}_i(\mathbf{y}_0)\partial_i\mathbf{y}\,\mathrm{d}\Omega-\int_\Omega \partial_i(\mathbf{z}^{\mathsf{T}}\boldsymbol{A}^{\rm f}_i)\mathbf{y}\,\mathrm{d}\Omega+\int_\Omega \partial_i\mathbf{z}^{\mathsf{T}}\boldsymbol{K}_{ij}(\mathbf{y}_0)\partial_j\mathbf{y}\,\mathrm{d}\Omega,\\ \ell(\mathbf{z}) &:= \int_\Omega \mathbf{z}^{\mathsf{T}}\boldsymbol{F}\,\mathrm{d}\Omega,\end{aligned}$$

where $\mathbf{y}_0$ is introduced to remark the strong non-linear character of the form $\mathscr{B}(\mathbf{y};\mathbf{y},\mathbf{z})$, taken into account in the first argument. The starting idea of VMS methods is to split the continuous space of the problem as $\mathscr{Y}=\mathscr{Y}_h\oplus\breve{\mathscr{Y}}$, where $\mathscr{Y}_h$ is the finite element space (and hence finite dimensional) and $\breve{\mathscr{Y}}$ is any complementary space which completes $\mathscr{Y}_h$ in $\mathscr{Y}$, usually termed subgrid space and which is in principle infinite dimensional. Such splitting of the space $\mathscr{Y}$ induces a scale separation of unknowns of the form $\mathbf{y}=\mathbf{y}_h+\breve{\mathbf{y}}$ (and same for the test functions). It is readily checked that the continuous problem can be written as the system of equations:

$$\mathscr{B}(\mathbf{y};\mathbf{y}_h,\mathbf{z}_h)+\mathscr{B}(\mathbf{y};\breve{\mathbf{y}},\mathbf{z}_h)=\ell(\mathbf{z}_h)\quad\forall\mathbf{z}_h\in\mathscr{Y}_h, \tag{31a}$$

$$\mathscr{B}(\mathbf{y};\mathbf{y}_h,\breve{\mathbf{z}})+\mathscr{B}(\mathbf{y};\breve{\mathbf{y}},\breve{\mathbf{z}})=\ell(\breve{\mathbf{z}})\quad\forall\breve{\mathbf{z}}\in\breve{\mathscr{Y}}. \tag{31b}$$

Introducing the following notation,

$$\int_{\Omega'}:=\sum_{e=1}^{n_{\rm el}}\int_{\Omega^{(e)}},\qquad \int_{\partial\Omega'}:=\sum_{e=1}^{n_{\rm el}}\int_{\partial\Omega^{(e)}},$$

and integrating by parts within each element in Eqs. (31a)– (31b), it is found that these two equations can be written as,

$$\mathscr{B}(\mathbf{y};\mathbf{y}_h,\mathbf{z}_h)+\int_{\Omega}\mathbf{z}^{\mathsf{T}}\boldsymbol{M}(\mathbf{y})\partial_t\breve{\mathbf{y}}\,\mathrm{d}\Omega+\int_{\Omega'}[\mathscr{L}^{*}(\mathbf{y};\mathbf{z}_h)]^{\mathsf{T}}\breve{\mathbf{y}}\,\mathrm{d}\Omega$$
$$+\int_{\partial\Omega'}\breve{\mathbf{y}}^{\mathsf{T}}n_i(\boldsymbol{K}_{ij}(\mathbf{y})\partial_j\mathbf{z}-\boldsymbol{A}_i^{\mathrm{c}}(\mathbf{y})\breve{\mathbf{z}})\,\mathrm{d}\partial\Omega=\ell(\mathbf{z}_h), \tag{32a}$$

$$\int_{\Omega'}\breve{\mathbf{z}}^{\mathsf{T}}[\boldsymbol{M}(\mathbf{y})\partial_t\breve{\mathbf{y}}+\mathscr{L}(\mathbf{y};\breve{\mathbf{y}})]\,\mathrm{d}\Omega+\int_{\partial\Omega'}\breve{\mathbf{z}}^{\mathsf{T}}n_i[\mathbf{K}_{ij}(\mathbf{y})\partial_j(\mathbf{y}_h+\breve{\mathbf{y}})-\mathbf{A}_i^{\mathrm{f}}(\mathbf{y}_h+\breve{\mathbf{y}})]\,\mathrm{d}\partial\Omega$$
$$=\int_{\Omega'}\breve{\mathbf{z}}^{\mathsf{T}}[\boldsymbol{F}-\boldsymbol{M}(\mathbf{y})\partial_t\mathbf{y}_h-\mathscr{L}(\mathbf{y};\mathbf{y}_h)]\,\mathrm{d}\Omega, \tag{32b}$$

where $n_i$ is now the $i$th component of the exterior normal to $\partial\Omega^{(e)}$ and $\mathscr{L}^{*}(\cdot,\cdot)$ is the formal adjoint operator of $\mathscr{L}(\cdot,\cdot)$, which is given by,

$$\mathscr{L}^{*}(\mathbf{y}_0;\mathbf{y}):=-[\partial_i\mathbf{A}_i^{\mathsf{T}}(\mathbf{y}_0)\mathbf{y}]-\partial_i[\mathbf{K}_{ij}^{\mathsf{T}}(\mathbf{y}_0)\partial_j\mathbf{y}]. \tag{33}$$

Eq. (32b) is equivalent to,

$$\boldsymbol{M}(\mathbf{y})\partial_t\breve{\mathbf{y}}+\mathscr{L}(\mathbf{y};\breve{\mathbf{y}})=\boldsymbol{F}-\boldsymbol{M}(\mathbf{y})\partial\mathbf{y}_h-\mathscr{L}(\mathbf{y};\mathbf{y}_h)+\mathbf{z}_h^{\perp}\qquad\text{in }\Omega^{(e)}, \tag{34a}$$
$$\breve{\mathbf{y}}=\breve{\mathbf{y}}_{\mathrm{ske}}\qquad\text{on }\partial\Omega^{(e)}, \tag{34b}$$

where $\mathbf{z}_h^{\perp}$ is obtained from the condition that $\breve{\mathbf{y}}$ must belong not to the whole space $\mathscr{Y}$, but to $\breve{\mathscr{Y}}$ and $\breve{\mathbf{y}}_{\mathrm{ske}}$ is a function defined over the skeleton of the computational domain and such that the normal component of the fluxes

$$\mathbf{q}_n:=n_i[\boldsymbol{K}_{ij}(\mathbf{y})\partial_j(\mathbf{y}_h+\breve{\mathbf{y}})-\boldsymbol{A}_i^{\mathrm{f}}(\mathbf{y}_h+\breve{\mathbf{y}})], \tag{35}$$

is continuous across interelement boundaries, and therefore the boundary term in Eq. (32b) vanishes.

Problem (31a)–(31b) is exactly equivalent to Eqs. (32a), (34a), (34b) in the distributional sense. The final approximate problem for the subscales is defined by the way in which problem (34a)–(34b) is solved as well as by the way in which functions $\breve{\mathbf{y}}_{\mathrm{ske}}$ and $\mathbf{z}_h^{\perp}$ are taken, what would give rise to different subgrid scales stabilization methods. However, the goal of this paper is not to propose a new stabilization technique, but rather to discuss the application of a well established formulation to the compressible Navier–Stokes equations written in primitive variables. This method can be obtained by solving for the subscales as a result of the following nonlinear evolution problem,

$$\boldsymbol{M}(\mathbf{y})\partial_t\breve{\mathbf{y}}+\boldsymbol{\tau}^{-1}(\mathbf{y})\breve{\mathbf{y}}=\breve{\mathscr{P}}[\mathscr{R}(\mathbf{y})]\qquad\text{in }\Omega^{(e)}, \tag{36a}$$
$$\breve{\mathbf{y}}=\breve{\mathbf{y}}_{\mathrm{ske}}\qquad\text{on }\partial\Omega^{(e)}, \tag{36b}$$

where $\boldsymbol{\tau}$ is the so-called matrix of stabilization parameters (and which we shall discuss later), $\breve{\mathscr{P}}$ denotes the projection operator onto $\breve{\mathscr{Y}}$, and $\mathscr{R}(\mathbf{y}_0):=[\boldsymbol{F}-\mathscr{L}(\mathbf{y}_0;\mathbf{y}_h)-\boldsymbol{M}(\mathbf{y}_0)\partial_t\mathbf{y}_h]$ is the finite element residual. There are two main possibilities for choosing $\breve{\mathscr{P}}$ and, therefore, for determining the space of subscales. The most common one is to take $\breve{\mathscr{P}}=I$, the identity, when applied to the finite element residual appearing in the right-hand-side of Eq. (36a). Another possibility, the one that we favor in this work, is to take $\breve{\mathscr{P}}=\mathscr{P}_h^{\perp}$ as the projection onto the space orthogonal to $\mathscr{Y}_h$, i.e. to the finite element space.

In order to fix ideas and close the approximation, we neglect the inter-element boundary terms, i.e. we assume $\breve{\mathbf{y}}_{\mathrm{ske}}=\mathbf{0}$, which can be understood as considering the subgrid scales as bubble functions, though this assumption could be relaxed, including the subscales on the element boundaries (see e.g. [46]). In addition, Eq. (36a) can be for instance integrated in time with a classical backward Euler scheme (following Section 4.1) and this yields,

$$\breve{\mathbf{y}}^{n+1}=\boldsymbol{\tau}_{\mathrm{dyn}}(\mathbf{y}^{n+1})\breve{\mathscr{P}}\left[\mathscr{R}(\mathbf{y}^{n+1})\right]+\boldsymbol{\tau}_{\mathrm{dyn}}(\mathbf{y}^{n+1})\boldsymbol{M}(\mathbf{y}^{n+1})\frac{\breve{\mathbf{y}}^{n}}{\Delta t}, \tag{37}$$

with $\boldsymbol{\tau}_{\mathrm{dyn}}=\left[\boldsymbol{M}(\mathbf{y}^{n+1})/\Delta t+\boldsymbol{\tau}^{-1}(\mathbf{y}^{n+1})\right]^{-1}$. The rationale behind using such a dissipative scheme for the time integration of the fine scales is precisely the assumption of bubble functions. However, it can be shown that this choice would not modify a second-order accuracy in time of the finite element solution (large scales) [39]. The final problem is: find $\mathbf{y}_h\in\mathscr{Y}_h$ such that,

$$\mathscr{B}(\mathbf{y}^{n+1};\mathbf{y}_h^{n+1},\mathbf{z}_h)+\int_{\Omega'}[\mathscr{L}^{*}(\mathbf{y}^{n+1};\mathbf{z}_h)]^{\mathsf{T}}\breve{\mathbf{y}}^{n+1}\,\mathrm{d}\Omega=\ell(\mathbf{z}_h)\quad\forall\,\mathbf{z}_h\in\mathscr{Y}_h, \tag{38}$$

which, upon substitution of the subscales by Eq. (37) with $\breve{\mathscr{P}}=\mathscr{P}_h^{\perp}$ yields the following discrete problem: find $\mathbf{y}_h\in\mathscr{Y}_h$ such that,

$$\begin{aligned}\mathscr{B}(\mathbf{y}^{n+1};\mathbf{y}_h^{n+1},\mathbf{z}_h) &+ \int_{\Omega'} [\mathscr{L}^*(\mathbf{y}^{n+1};\mathbf{z}_h)]^\top \boldsymbol{\tau}_{\text{dyn}}(\mathbf{y}^{n+1})\mathscr{P}_h^\perp\left[\mathscr{R}(\mathbf{y}^{n+1})\right] \mathrm{d}\Omega = \ell(\mathbf{z}_h) \\ &+ \int_{\Omega'} [\mathscr{L}^*(\mathbf{y}^{n+1};\mathbf{z}_h)]^\top \boldsymbol{\tau}_{\text{dyn}}(\mathbf{y}^{n+1})\boldsymbol{M}(\mathbf{y}^{n+1})\frac{\breve{\mathbf{y}}^n}{\Delta t}\,\mathrm{d}\Omega \quad \forall\,\mathbf{z}_h \in \mathscr{Y}_h. \end{aligned} \tag{39}$$

### 6.2. OSGS stabilized formulation of the compressible Navier–Stokes problem

In this subsection several details and simplifications about the stabilized finite element method arising from Eq. (39) are introduced. Up to now we have described the subgrid scale stabilization technique when applied to a general setting. The objective now is to specify its application to the compressible Navier–Stokes problem that we are considering in this paper.

#### 6.2.1. Some initial simplifications

As we have pointed out earlier, when the space of subscales is enforced to be $L^2$-orthogonal to the finite element space $\mathscr{Y}_h$, the method is termed Orthogonal SubGrid-Scale (OSGS or simply OSS). It corresponds to taking $\breve{\mathscr{P}} = \mathscr{P}_h^\perp = I - \mathscr{P}_h$ where $\mathscr{P}_h$ is the projection operator onto the appropriate finite element space without boundary conditions. The particularity of this method is that it makes the subscales active in regions which cannot be resolved by the finite element mesh. As a result, the OSGS method allows certain type of simplifications [37], which we state next.

- The orthogonal projection of the external forces might be neglected. External loads are assumed to belong to the finite element spaces or are approximated by an element of the corresponding space. Hence,
$$\mathscr{P}_h^\perp(\rho r) \approx 0, \quad \mathscr{P}_h^\perp(\rho\mathbf{f}) \approx 0.$$
- The orthogonal projection of terms involving temporal derivatives might be also neglected. Their orthogonal projection would be exactly zero if $\rho$, $\beta$ and $\alpha$ were constant. However we consider as true that those parameters are such the temporal terms already belong to the finite element space. This simplification is key in order to solve the problem by means of a fractional step method in time. This amounts to saying that,
$$\mathscr{P}_h^\perp(\rho\partial_t\mathbf{u}_h) \approx 0, \quad \mathscr{P}_h^\perp(\beta\partial_t p_h) \approx 0, \quad \mathscr{P}_h^\perp(\alpha\partial_t\theta_h) \approx 0, \quad \mathscr{P}_h^\perp(\rho c_{\mathrm{v}}\partial_t\theta_h) \approx 0.$$
- The evaluation of second derivatives in the stabilization terms is a costly and cumbersome process in finite element implementations which can be avoided using orthogonal subgrid scales. Second order derivatives are exactly zero for linear elements and for higher order interpolation, disregarding them leads to a weakly consistent formulation in the context of finite elements, that is to say, the stabilization terms do not cancel for the exact solution but vanish as the mesh size goes to zero.

Taking all this information into account, the finite element discrete residuals of the momentum, continuity and energy conservation equations can be written as,

$$\mathscr{R}(\mathbf{y}^{n+1}) = -\begin{bmatrix} \rho^{n+1}(\mathbf{a}^{n+1}\cdot\nabla)\mathbf{u}_h^{n+1} + \nabla p_h^{n+1} \\ \beta^{n+1}(\mathbf{a}^{n+1}\cdot\nabla)p_h^{n+1} - \alpha^{n+1}(\mathbf{a}^{n+1}\cdot\nabla)\theta_h^{n+1} + \nabla\cdot\mathbf{u}_h^{n+1} \\ \rho^{n+1}c_{\mathrm{v}}(\mathbf{a}^{n+1}\cdot\nabla)\theta_h^{n+1} + (\nabla\cdot\mathbf{u}_h^{n+1})p_h^{n+1} - 2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h^{n+1}):\boldsymbol{\varepsilon}(\mathbf{u}_h^{n+1}) + \frac{2}{3}\mu(\nabla\cdot\mathbf{u}_h^{n+1})^2 \end{bmatrix}, \tag{40}$$

with $\mathbf{a}^{n+1} = \mathbf{u}_h^{n+1}$ or $\mathbf{a}^{n+1} = \mathbf{u}_h^{n+1} + \breve{\mathbf{u}}^{n+1}$ if the non-linear character of the subscales is accounted for. In addition, the adjoint operator is identified with,

$$\mathscr{L}^*(\mathbf{y}^{n+1};\mathbf{z}_h) = -\begin{bmatrix} \nabla\cdot(\rho^{n+1}\mathbf{a}^{n+1}\otimes\mathbf{v}_h) + \nabla q_h \\ \nabla\cdot\mathbf{v}_h + \nabla\cdot(\beta^{n+1}\mathbf{a}^{n+1}q_h) \\ -\nabla\cdot(\alpha^{n+1}\mathbf{a}^{n+1}q_h) + c_{\mathrm{v}}\nabla\cdot(\rho^{n+1}\mathbf{a}^{n+1}\eta_h) \end{bmatrix}. \tag{41}$$

**Remark 6.1.** Note that we have not included in the adjoint operator of the temperature equation the terms corresponding to the compression work and the dissipation function. The hope is that their influence on the stability of the scheme is small. The formulation is weakly consistent, as explained above, and it is designed by taking the subscales in the orthogonal space to the finite-dimensional resolved space.

Furthermore, the subscales are time tracked in time by solving Eq. (37) which for the problem in hand reads,

$$\breve{\mathbf{u}}^{n+1} = -\tau_{\mathbf{u},\text{dyn}}\mathscr{P}_h^\perp\left[\rho(\mathbf{a}^{n+1}\cdot\nabla)\mathbf{u}_h^{n+1} + \nabla p_h^{n+1}\right] + \tau_{\mathbf{u}}\rho\frac{\breve{\boldsymbol{u}}^n}{\Delta t}, \tag{42a}$$

$$\breve{p}^{n+1} = -\tau_{p,\mathrm{dyn}}\mathscr{P}_h^{\perp}\left[\beta(\mathbf{a}^{n+1}\cdot\nabla)p_h^{n+1} - \alpha(\mathbf{a}^{n+1}\cdot\nabla)\theta_h^{n+1} + \nabla\cdot\mathbf{u}_h^{n+1}\right] + \tau_p\beta\frac{\breve{p}^n}{\Delta t}, \tag{42b}$$

$$\begin{aligned}\breve{\theta}^{n+1} &= -\tau_{\theta,\mathrm{dyn}}\mathscr{P}_h^{\perp}\left[\rho c_{\mathrm{v}}(\mathbf{a}^{n+1}\cdot\nabla)\theta_h^{n+1} + (\nabla\cdot\mathbf{u}_h^{n+1})p_h - 2\mu\boldsymbol{\varepsilon}(\mathbf{u}_h^{n+1}):\boldsymbol{\varepsilon}(\mathbf{u}_h^{n+1}) + \frac{2}{3}\mu(\nabla\cdot\mathbf{u}_h^{n+1})^2\right]\\ &\quad + \tau_\theta\rho c_{\mathrm{v}}\frac{\breve{\theta}^n}{\Delta t}.\end{aligned} \tag{42c}$$

*6.2.2. On the stabilization parameters*

Although the application of the VMS method to the compressible Navier–Stokes problem has already been discussed, the stabilization technique is not completed until one introduces a definition to compute the matrix of stabilization parameters $\boldsymbol{\tau}(\mathbf{y}^{n+1})$ (generally nonlinear). This is what we do next.

Up to our knowledge, there is no general rule to define it for systems of equations. It must be designed for each particular problem taking into account its stability deficiencies or even scaling requirements. For the problem in hand, we will take a simple diagonal expression. In the $n_{\mathrm{sd}} = 3$ scenario we define,

$$\boldsymbol{\tau}(\mathbf{y}^{n+1}) = \mathrm{diag}(\tau_{\mathbf{u}}\mathbf{I}, \tau_p, \tau_\theta). \tag{43}$$

The usual definition of the compressible stabilization parameters include a local sound velocity that arises from a linearization of the characteristic compressible flow problem (see [47] for a Fourier analysis). At the low Mach number limit the sound speed tends to infinity, and therefore such stabilization parameters are not suitable for the computations. In this work, we consider the following definition [48]:

$$\tau_u^{-1} = c_1\frac{\mu}{h^2} + c_2\frac{\rho a}{h}, \tag{44a}$$

$$\tau_p^{-1} = \frac{\tau_u}{h^2}, \tag{44b}$$

$$\tau_\theta^{-1} = c_1\frac{\kappa}{h^2} + c_2\frac{\rho c_{\mathrm{v}} a}{h}, \tag{44c}$$

where we recall that $h$ stands for the mesh size. It is understood that these expressions are evaluated element by element. The numerical constants $c_1$ and $c_2$ are independent of the physical parameters of the problem. In the numerical calculations we take them as $c_1 = 15\omega^4$, $c_2 = 2\omega$, $\omega$ being the order of the finite element interpolation. In the case of the largest characteristic velocity in the convective contribution, which theoretically should correspond to $(|\mathbf{u}| + c)$, we introduce a modified velocity $a$, that we take as the harmonic mean value of $|\mathbf{u}_h|$ and $c$, i.e. $a = 1/(1/|\mathbf{u}_h| + 1/c)$. This definition for the characteristic velocity accounts for the effect of the sound speed, and ensures the proper definition at the zero Mach limit $(c \to \infty)$.

## 7. Final numerical scheme OSS discrete problem: algebraic fractional step version

The derivation of the final matrix version of the algorithm is straightforward now and the matrix system that needs to be solved at each time step has a similar algebraic structure as Eq. (25) with the addition of the corresponding arrays coming from the adopted stabilization technique, which we denote, following the subscript notation in Section 5, as $\mathbf{S}_{\mathbf{uu}}(\rho, \mathbf{U})$, $\mathbf{S}_{\mathbf{u}p}(\rho, \beta, \mathbf{U})$, $\mathbf{S}_{\mathbf{u}\theta}(\rho, \alpha, \mathbf{U})$, $\mathbf{S}_{pp}(\rho, \beta, \alpha, \mathbf{U})$, $\mathbf{S}_{p\mathbf{u}}(\rho, \beta, \alpha, \mathbf{U})$, $\mathbf{S}_{p\theta}(\rho, \beta, \alpha, \mathbf{U})$, $\mathbf{S}_{\theta\theta}(\rho, \mathbf{U})$, $\mathbf{S}_{\theta p}(\rho, \mathbf{U})$ and $\mathbf{S}_{\theta\mathbf{u}}(\rho, \mathbf{U})$ for the left-hand-side matrices and $\mathbf{F}_{s,\mathbf{u}}(\rho, \mathbf{U})$, $\mathbf{F}_{s,p}(\rho, \beta, \alpha, \mathbf{U})$ and $\mathbf{F}_{s,\theta}(\rho, \mathbf{U})$ for their corresponding right-hand-side vectors containing the known projections and subscales. It is easy to identify their expressions from Eqs. (40)–(41) and (42a)–(42c) together with Eq. (39), respectively.

From this point, the derivation of the fractional step procedure can be facilely extended to account for the stabilization arrays and follows the same steps as done before. In addition, the unknown projections for a given equation can be computed with the values at the previous time step or iteration. The final algorithm is an extension of Algorithm 1 and it is summarized here as follows:

(1) Compute the intermediate velocity $\widetilde{\mathbf{U}}^{n+1}$ using the pressure extrapolations:

$$\begin{aligned}\mathbf{M}_{\mathbf{uu}}(\rho)\frac{D_k}{\Delta t}\widetilde{\mathbf{U}}^{n+1} &+ \mathbf{K}_{\mathbf{uu}}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{U}}^{n+1} + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1} + \mathbf{K}_{\mathbf{uu}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1} + \mathbf{S}_{\mathbf{uu}}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{U}}^{n+1}\\ &= \mathbf{F}_{\mathbf{u}}^{n+1}(\rho) + \mathbf{K}_{\mathbf{u}p}\widehat{\mathbf{P}}_{k-1}^{n+1} - \mathbf{K}_{\mathbf{u}p}^{\partial\Omega}\widehat{\mathbf{P}}_{k}^{n+1} - \mathbf{S}_{\mathbf{u}p}(\rho, \beta, \widetilde{\mathbf{U}}^{n+1})\mathbf{P}^n\\ &\quad - \mathbf{S}_{\mathbf{u}\theta}(\alpha, \widetilde{\mathbf{U}}^{n+1})\boldsymbol{\Theta}^n + \mathbf{F}_{s,\mathbf{u}}(\rho, \widetilde{\mathbf{U}}^{n+1}).\end{aligned}$$

(2) Compute the intermediate temperature $\widetilde{\boldsymbol{\Theta}}^{n+1}$ using the pressure extrapolations and the known intermediate velocity:

$$\mathbf{M}_{\theta\theta}(\rho)\frac{D_k}{\Delta t}\widetilde{\boldsymbol{\Theta}}^{n+1} + \mathbf{K}_{\theta\theta}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\boldsymbol{\Theta}}^{n+1} + \mathbf{S}_{\theta\theta}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\boldsymbol{\Theta}}^{n+1}$$

$$\begin{aligned}
&= \mathbf{F}_{\theta}^{n+1}(\rho) - \mathbf{K}_{\theta p}(\widetilde{\mathbf{U}}^{n+1})\widehat{\mathbf{P}}_{k-1}^{n+1} - \mathbf{S}_{\theta p}(\rho, \widetilde{\mathbf{U}}^{n+1})\mathbf{P}^{n} \\
&\quad - \mathbf{S}_{\theta \mathbf{u}}(\rho, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{U}}^{n+1} + \mathbf{F}_{s,\theta}(\rho, \widetilde{\mathbf{U}}^{n+1}).
\end{aligned}$$

(3) Compute the pressure $\mathbf{P}^{n+1}$ using the intermediate velocity and temperature:

$$\begin{aligned}
\mathbf{M}_{pp}(\beta)\frac{D_k}{\Delta t}\mathbf{P}^{n+1} &+ \mathbf{C}_{pp}(\beta, \widetilde{\mathbf{U}}^{n+1})\mathbf{P}^{n+1} + \phi_k \Delta t \mathbf{Q}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}) + \phi_k \Delta t \mathbf{L}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}) \\
&+ \mathbf{S}_{pp}(\rho, \beta, \alpha, \widetilde{\mathbf{U}}^{n+1})\mathbf{P}^{n+1} = \mathbf{F}_{p}^{n+1} - \mathbf{K}_{p\mathbf{u}}\widetilde{\mathbf{U}}^{n+1} + \mathbf{C}_{p\theta}(\alpha, \widetilde{\mathbf{U}}^{n+1})\widetilde{\boldsymbol{\Theta}}^{n+1} \\
&+ \mathbf{M}_{p\theta}(\alpha)\frac{D_k}{\Delta t}\widetilde{\boldsymbol{\Theta}}^{n+1} - \mathbf{K}_{p\mathbf{u}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1} - \mathbf{S}_{p\mathbf{u}}(\rho, \beta, \alpha, \widetilde{\mathbf{U}}^{n+1})\widetilde{\mathbf{U}}^{n+1} \\
&- \mathbf{S}_{p\theta}(\rho, \beta, \alpha, \widetilde{\mathbf{U}}^{n+1})\widetilde{\boldsymbol{\Theta}}^{n+1} + \mathbf{F}_{s,p}(\rho, \beta, \alpha, \widetilde{\mathbf{U}}^{n+1}).
\end{aligned}$$

(4) Velocity correction to obtain the end-of-step velocity $\mathbf{U}^{n+1}$:

$$\mathbf{M}_{\mathbf{uu}}(\rho)\frac{1}{\phi_k \Delta t}\mathbf{U}^{n+1} + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega}\mathbf{U}^{n+1} = \mathbf{M}_{\mathbf{uu}}(\rho)\frac{1}{\phi_k \Delta t}\widetilde{\mathbf{U}}^{n+1} + \mathbf{M}_{\mathbf{uu}}^{\partial\Omega}\widetilde{\mathbf{U}}^{n+1} + \mathbf{K}_{\mathbf{u}p}(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}).$$

(5) Temperature correction to obtain the end-of-step temperature $\boldsymbol{\Theta}^{n+1}$:

$$\mathbf{M}_{\theta\theta}(\rho)\frac{1}{\phi_k \Delta t}\boldsymbol{\Theta}^{n+1} = \mathbf{M}_{\theta\theta}(\rho)\frac{1}{\phi_k \Delta t}\widetilde{\boldsymbol{\Theta}}^{n+1} - \mathbf{K}_{\theta p}(\widetilde{\mathbf{U}}^{n+1})(\mathbf{P}^{n+1} - \widehat{\mathbf{P}}_{k-1}^{n+1}).$$

### *7.1. On the linearization of the discrete problem*

Apart from the classical nonlinearities appearing in the convective terms due to the advection velocity, there are others inherent to the nature of the compressible Navier–Stokes equations. In particular, all the temporal terms in the problem are nonlinear due to the presence of the density $\rho$, the thermal expansion coefficient $\alpha$ and the isothermal compressibility coefficient $\beta$. As in any fractional step scheme, the time step of the computation $\Delta t$ cannot be taken very large for the method to be effective, when it is compared for example to the critical time step of an explicit time integration scheme. This fact is remarked in compressible flow simulations, as the time step should be sufficiently small in order to reproduce the wide range of different scales that the compressible Navier–Stokes problem spans. Thus, for a given solution at a particular non-linear iteration, the next solution should be particularly close. Hence, only a few nonlinear iterations are required to converge and the fixed-point option is usually enough for the convective and temporal nonlinearities in flow problems. Otherwise one should expand by means of a Taylor series the expressions for density, volume expansivity and isothermal compressibility coefficients. Yet this may be appealing, it definitely introduces some burden in the formulation since the resulting terms involve rational expressions which are in general arduous to integrate numerically.

Let us also explain how we manage the orthogonal projections $\mathscr{P}_h^{\perp}$. When compared to the raw Galerkin method, the matrices emerging from the orthogonal projection of the unknowns show a wide stencil. In order to avoid dealing with them, at the $i$th iterion of the $n$th time step we may approximate $\mathscr{P}_h^{\perp}(f^{n,(i)}) \approx f^{n,(i)} - \mathscr{P}_h(f^{n,(i-1)})$ or $\mathscr{P}_h^{\perp}(f^{n,(i)}) \approx f^{n,(i)} - \mathscr{P}_h(f^{n-1})$ for any generic function $f$. In other words, we perform the projection by means of known values from either the previous iteration or time step. Numerical tests reveal that both options are effective, the latter being chosen in the numerical examples presented next.

### *7.2. A final note on implementation*

As stated before in this article, the primitive formulation for ideal gases in the low Mach number limit results in very large quantities in the continuity and energy equations, specially if the international system of units is to be used. When an implicit numerical scheme is used, the resulting linear system of equations which needs to be solved at each time step commonly contains very large values for these two equations, making it very inefficient to solve by using iterative methods. In order to overcome this difficulty, we perform a decomposition of the primitive (absolute) variables into a relative (or *gauge*) part and a reference (or *atmospheric*) part. Thus, we set,

$$p = p^* + p_{\mathrm{atm}} \quad ; \quad \mathbf{u} = \mathbf{u}^* \quad ; \quad \theta = \theta^* + \theta_{\mathrm{atm}},$$

assuming the atmospheric values to be constant. Taking this into account, we may write the initial system of equations using the relative part of the unknowns, but taking special care of the non-linear terms, where the complete contribution is needed to obtain reliable results. Furthermore, initial and boundary conditions must be set for the primitive problem written using relative unknowns.

## 8. Numerical examples

In this section we include some numerical examples in order to demonstrate, on the one hand, the performance of the proposed stabilized fractional scheme and, on the other hand, the suitability of the developed method to replace a standard monolithic strategy.

First, we perform a test with analytical solution in order to numerically check the convergence rate in time. Next, we include a classical benchmark in the field of aeroacoustics, which consists in simulating the sound generated by the flow past a cylinder, also known as the aeolian tones problem. The last example is based on simulating the differentially heated flow inside a cavity, which is a well-known benchmark for thermal problems. For all the upcoming numerical simulations, we consider an ideal gas with $\gamma = 1.4$ ($\mathrm{R_g} = 287$ J/(kg K)) and $c_\mathrm{p} = 1004.5$ J/(kg K).

All the implementations of the algorithms in this paper have been carried out into our in-house code `FEMUSS` (Finite Element Method Using Subgrid Scales). `FEMUSS` is an object-oriented `Fortran`-based finite element code which follows a modular approach for multiphysics interaction and performs parallel computations under `MPI` directives. In order to solve the final underlying systems of linear equations, if nothing else is stated, we make use of an iterative algorithm based on the stabilized version of the BiConjugate Gradient method `BiCGstab` [49], which is already implemented in the `PETSc` parallel solver library [50]. As previously discussed, the nonlinearities in the problem are solved via Picard's scheme. This leads to a monotonically decreasing relative error among consecutive iterations, ensuring the convergence of the method. A maximum of 10 iterations is set, and the relative numerical tolerance for the $L^2$-norm is $1 \times 10^{-5}$. All the plots are in SI units.

### 8.1. Convergence test

The first example is a convergence test where we analyze the time errors of the fractional step scheme. Here we make use of the method of manufactured solutions, which has been traditionally used to quantify the numerical error of partial differential equations solvers. In this example, an exact solution for pressure, velocity and temperature is specified in the computational domain, which are composed of smooth functions, yet they have no physical meaning. This example is similar to the one from [51]. We solve the Navier–Stokes problem over the unit square $\overline{\Omega} = [0, 1]^2$ and the force term is set so that the exact solution of the problem is,

$$p^*(x, y, t) = \cos(\pi x)\sin(\pi y)\sin(t), \tag{45a}$$

$$u(x, y, t) = \pi \sin(2\pi y)\sin^2(\pi x)\sin(t), \tag{45b}$$

$$v(x, y, t) = -\pi \sin(2\pi x)\sin^2(\pi y)\sin(t), \tag{45c}$$

$$\theta^*(x, y, t) = u(x, y, t). \tag{45d}$$

The contour plots of these fields are presented for completeness in Fig. 1. The finite element partition is structured and uniform and contains $Q_2/Q_2/Q_2$ finite elements of size $h = 1/200$. Both the boundary and initial conditions are evaluated from the previous equations, and particularized for each of the sides of the square at each time step and for $t = 0$, respectively. We use a range of time step sizes from 0.1 s to 0.003125 s and the problem is solved until $t_\mathrm{f} = 1$ s. In addition, we set $\mu = 0.001$ kg/(m s), $\kappa = 1$ W/(m K), $p_\mathrm{atm} = 1 \times 10^5$ Pa and $\theta_\mathrm{atm} = 300$ K and we make use of a sparse direct solver from the `MUMPS` library [52,53]. The error between the exact solution of the Navier–Stokes equations and the numerical one is measured in the $\ell^2$-norm of the sequence of spatial $L^2$-norms of the solutions, i.e.,

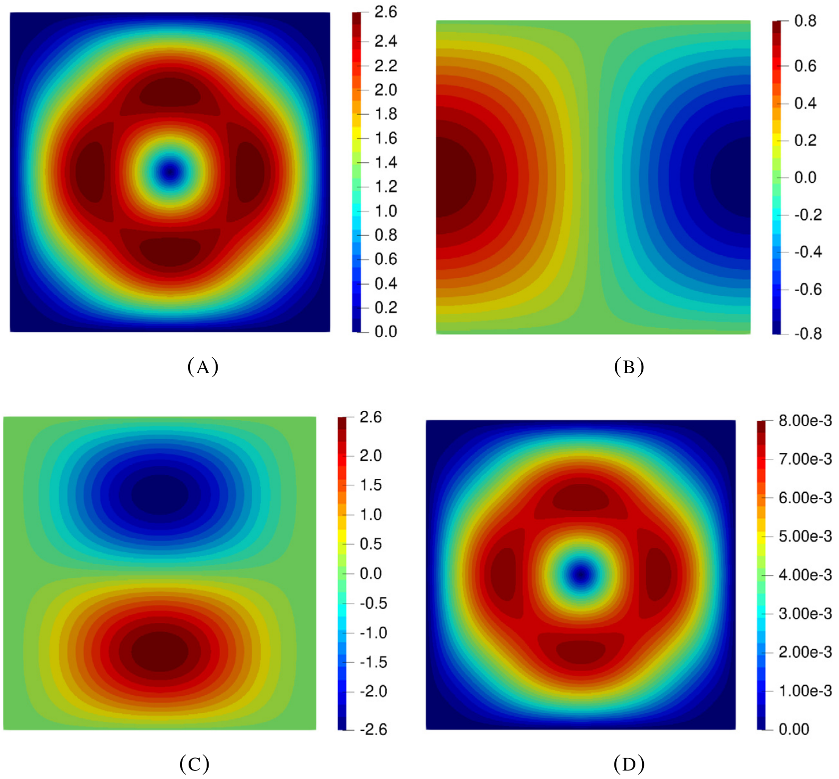
$$E_f := \left( \Delta t \sum_{n=1}^{N} \left\| \frac{f_h^n - f(t^n)}{f(t^n)} \right\|_{L^2}^2 \right)^{1/2},$$

for $f = \mathbf{u}$, $p$ or $\theta$, respectively.

In Fig. 2 we compare the convergence results of the standard monolithic solution with the ones obtained with the fractional step algorithm for both first and second order time integration schemes. The monolithic results can be used as a reference against which the results obtained with the fractional step can be compared because they have no splitting error. The expected convergence rate can be clearly seen for both temporal approximations and for all the time step sizes, yet the splitting errors of the fractional step approach become more noticeable for the BDF2 scheme. From the convergence plots it is also observed that the spatial error is not significant for the mesh size used.

### 8.2. Aeolian tones: flow past a cylinder

The next example consists in a two-dimensional circular cylinder embedded in a flow with free stream. We use this example to evaluate the performance of the VMS formulation in unsteady flows and to test the proposed non-reflecting boundary conditions. In this problem, a sequence of vortices is generated at the lee of the cylinder which are transported downstream. As a reaction to this fact, the cylinder undergoes lift fluctuations that lead to the emission of sound, with



(A)

(B)

(C)

(D)

**Fig. 1.** Convergence test: contour plots of the exact solutions from Eq. (45) at $t = 1$ s, (a) velocity magnitude $\|\mathbf{u}(\mathbf{x}, t)\|$, (b) relative pressure $p^*(\mathbf{x}, t)$ and (c) relative temperature $\theta^*(\mathbf{x}, t)$. Also the associated Mach number Ma distribution (d) is included.

a wave frequency established at the fixed value of the wake fluctuation. The emitted noise is commonly referred as an aeolian tone, and a typical example of this phenomenon is the wire whistle that can be heard when wind impacts power transmission lines.

The computational domain for this simulation is $\overline{\Omega} = [0, 0] \times [50, 50] \setminus \mathscr{C}$, being $\mathscr{C}$ the cylinder region of diameter $D = 0.1$ m which is placed at point $(20, 25)$. The domain is big enough to describe the far field conditions far away from the cylinder. A free stream enters from the left-most boundary with the following data: $U_\infty = 70$ m/s (horizontal), $p_\infty = p_{\text{atm}} = 101325$ Pa and $\theta_\infty = \theta_{\text{atm}} = 300$ K, which allows to match the tested flow conditions of $\text{Re}_\infty = 150$ and $\text{Ma}_\infty = 0.2$, for which meaningful 2D calculations can be performed. The Prandtl number is $\text{Pr} = 0.75$. No slip-adiabatic boundary conditions are prescribed on the cylinder surface and the free stream values are used as initial condition. We solve the compressible Navier–Stokes equations on an unstructured mesh of linear triangular elements, progressively refined on the cylinder surface, accounting for a total of $\sim$ 225,000 elements. The time step size used in the computation is constant and equal to $\Delta t = 0.0001$ s. The second order BDF2 scheme has been used for the large scales time evolution, while a first-order scheme has been used for the tracking of the subscales. In order to overcome all the burden related to any possible wave reflections at the external walls, we use the boundary formulation described in Section 3 and for this purpose we set a filtering frequency of 200 Hz.

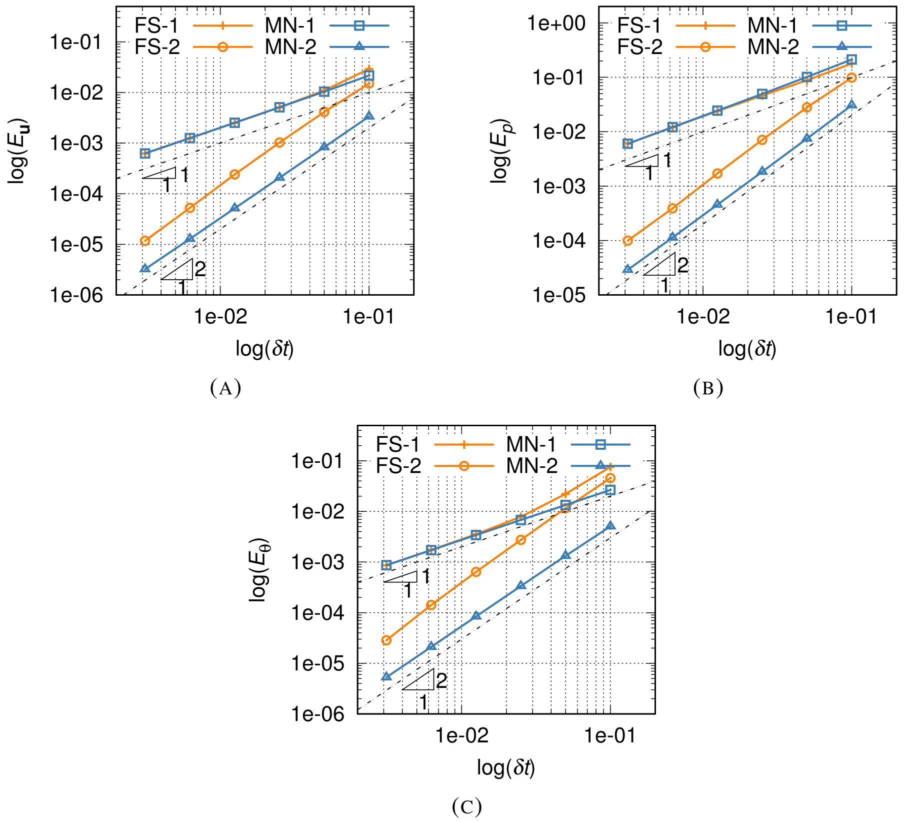
In Fig. 3 we present a depiction of the near-field developed flow, including the instantaneous contours of velocity (magnitude), pressure and temperature obtained with the fractional step algorithm. As we can see, they display the classical oscillating wake after the cylinder surface.

Since no notable differences were observed between the results of both the monolithic and fractional step counterpart, we provide a quantitative comparison by computing some aerodynamic integral values of the flow. In particular, we calculate the lift and drag non-dimensional coefficients,

$$C_{\text{d}} = \frac{F_x}{\frac{1}{2}\rho_\infty U_\infty^2 D}, \; C_{\text{l}} = \frac{F_y}{\frac{1}{2}\rho_\infty U_\infty^2 D}, \tag{46}$$

computed from the exerted force of the fluid over the cylinder,

$$\mathbf{F} = -\int_{\partial\Omega_{\text{Cylinder}}} \left(-p\mathbf{I} + 2\mu\boldsymbol{\varepsilon}(\mathbf{u}) - \frac{2}{3}\mu(\nabla\cdot\mathbf{u})\mathbf{I}\right) \cdot \mathbf{n}\, \mathrm{d}\partial\Omega. \tag{47}$$

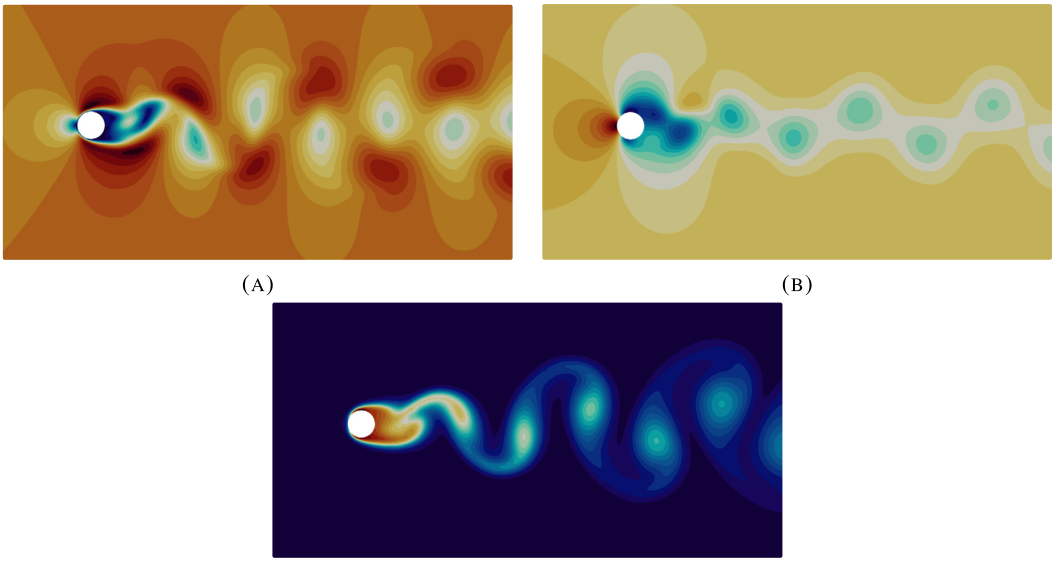


(A)

(B)

(C)

**Fig. 2.** Convergence test: time convergence of the relative errors of the three unknown variables measured in the $\ell^2(0, t_{\mathrm{f}}, L^2(\Omega))$-norm. Here and in what follows, FS stands for fractional step results and MN for monolithic results. The number after the dash symbol stands for first (1) or second (2) order scheme in time.



(A)

(B)

(C)

**Fig. 3.** 2D flow past a cylinder: near-field results, (a) velocity magnitude $\|\mathbf{u}\|$, (b) relative pressure $p^*$ and (c) relative temperature $\theta^*$ contours obtained with the fractional step scheme. Contours levels range, respectively, from 0 to 96 m/s, $-4.3 \times 10^3$ to $3.1 \times 10^3$ Pa and 0 to 6 K.

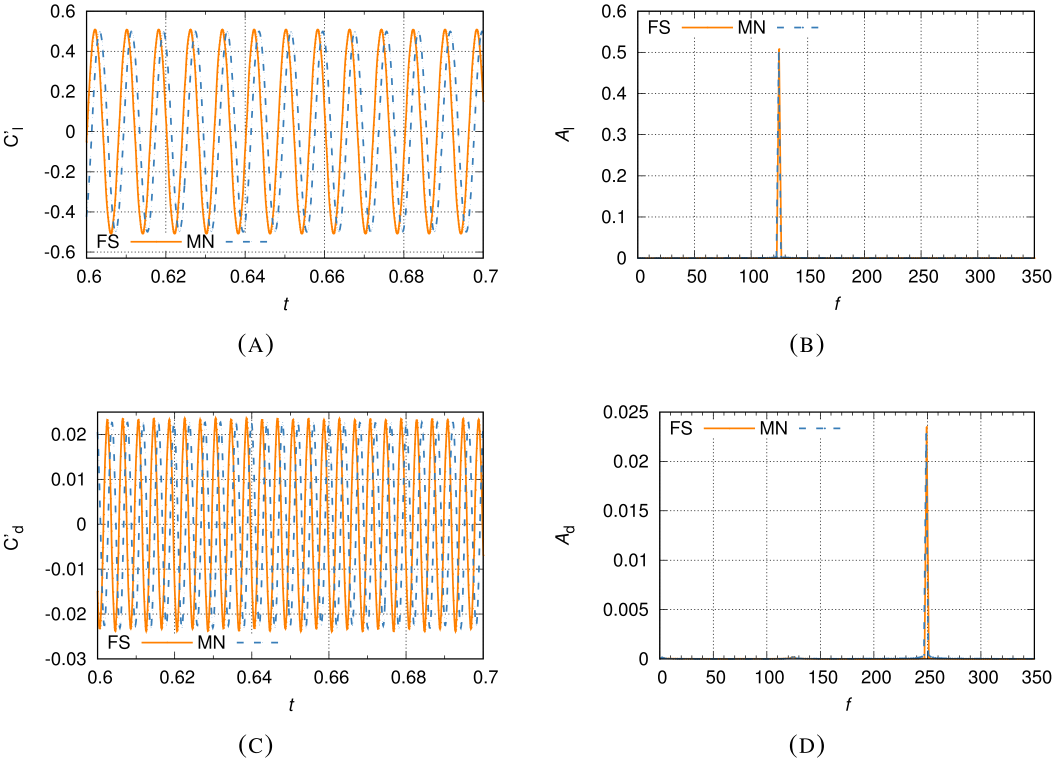


**Fig. 4.** 2D flow past a cylinder: part of the time history of the fluctuating non-dimensional lift (a) and drag (c) coefficients, together with FFTs obtained for time varying fluctuations in lift (b) and drag (d).

**Table 1**
2D flow past a cylinder: aerodynamic statistics for $\text{Re}_\infty = 150$ and $\text{Ma}_\infty = 0.2$. IBM: Immersed boundary method, FD-DNC: Finite difference Direct Numerical Simulation, LBM: Lattice Boltzmann method, FE: Finite Element.

| Ref. | Method | $\overline{C_\text{d}}$ | $A_\text{d}$ | $A_\text{l}$ | St |
|---|---|---|---|---|---|
| Margnat [54] | Hybrid IBM | 1.35 | 0.026 | 0.52 | 0.189 |
| Inoue & Hatakeyama [57] | FD-DNC | 1.32 | 0.026 | 0.52 | 0.183 |
| Laffite & Pérot [58] | LBM | 1.39 | 0.028 | 0.56 | 0.183 |
| Ganta *etal* [59] | FD-DNC | 1.34 | 0.026 | 0.53 | 0.183 |
| Present | FE-Monolithic | 1.328 | 0.023 | 0.499 | 0.178 |
| Present | FE-Segregated | 1.334 | 0.022 | 0.508 | 0.178 |

In addition, we compute the Strouhal number as

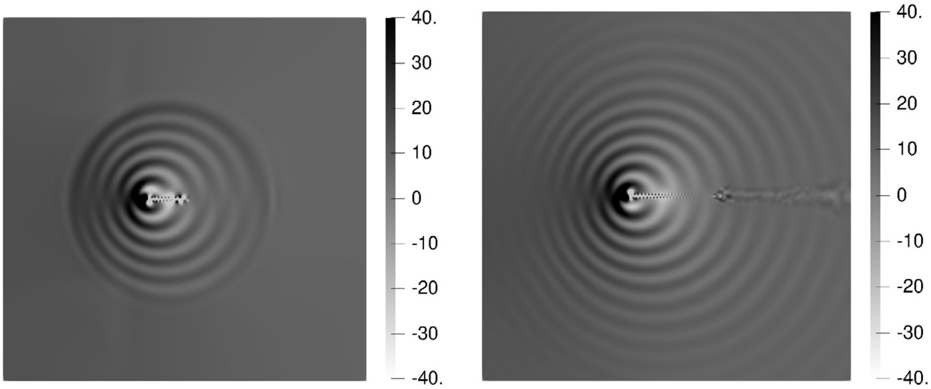
$$\text{St} = \frac{f D}{U_\infty}, \tag{48}$$

being $f$ the vortex shedding frequency.

Several studies on flow past bluff bodies discuss that time varying fluctuating loads over the surface of the studied body (fluctuations in lift and drag coefficients) are the main mechanism for generation and propagation of sound waves [54–56]. Here we evaluate the fluctuations in the lift and drag coefficients (which we denote hereafter as $C'_\text{l}$ and $C'_\text{d}$, respectively) as the difference between their instantaneous and mean values, respectively. A part of the time history of the lift and drag fluctuations is presented in Figs. 4a and 4c for both monolithic and fractional step schemes. The Fast Fourier Transform (FFT) of these time varying signals has been obtained, and the plots of the corresponding Fourier amplitudes ($A_\text{l}$ and $A_\text{d}$) as a function of the Fourier frequency are presented in Figs. 4b and 4d.

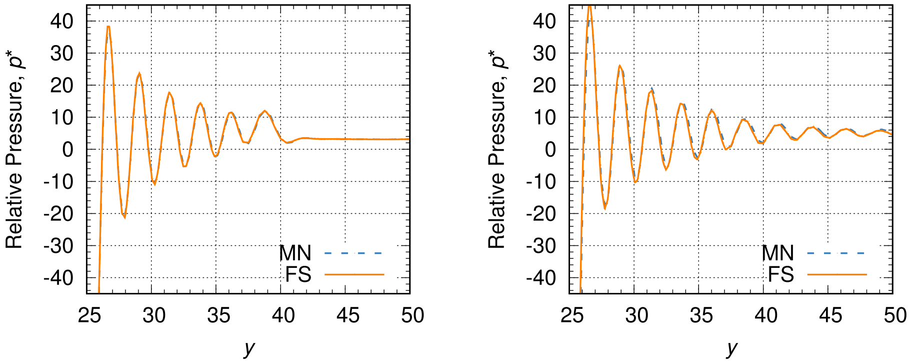
Both monolithic and segregated methods basically provide the same frequencies of oscillation for lift and drag. The lift coefficient oscillates at a frequency of 124.673 Hz ($\text{St} = 0.178$), and the drag coefficient at 249.341 Hz (twice of the vortex shedding one). Additional statistic results are collected in Table 1, together with some reference values reported in the literature and which cover a different variety of numerical methods [54,57–59]. An overall agreement can be observed.

Regarding the propagation of acoustic waves, Fig. 5 contains the far field pressure contours at two certain time steps obtained with the fractional step algorithm, one before the acoustic waves reach the external boundaries and another one on the long term once the waves hit the artificial contours. The vortices located at the wake of the cylinder perturb the



(A) (B)

**Fig. 5.** 2D flow past a cylinder: relative pressure $p^*$ contours showing the propagation of sound waves towards the far field at two different times. Monolithic contour plots are omitted since they are remarkably similar [48].



(A) (B)

**Fig. 6.** 2D flow past a cylinder: instantaneous pressure along the positive $y$ direction from the center of the cylinder.

pressure field, which is propagated in the form of sound waves. These pressure pulses evolve radially from the cylinder area with time, yet they do not propagate normally to the flow direction since this case is based on an uniform flow. The inclusion of non-reflecting boundary conditions (or any other dumping technique) is essential for the adequate reproduction of the pressure waves, due to the fact that reflections may occur at the external boundary and hence local numerical oscillations may end up polluting the solution. In this sense, we successfully accomplish the (direct) numerical simulation of the acoustic pressure propagation thanks to the above-developed technique for the prescription of boundary conditions. In addition to this, a plot concerning the pressure wave along the positive $y$ direction is included in Fig. 6. We observe that our fractional step scheme is able to agreeably replicate the dissipation of the radiated sound wave of the monolithic scheme (which serves as reference here, [48]), yet minor discrepancies might be noticed. These minor differences should come from the splitting error introduced by the segregated approach as well as from the approximate boundary condition.

### 8.3. Differentially heated flow inside a cavity

In this numerical simulation, we model the flow in a differentially heated cavity both in 2D and 3D configurations (see e.g. [7,60,61]). This is a natural convection flow problem in which the fluid is driven both by a large temperature gradient between two walls and by a gravity force, which plays a significant role in the development of the buoyancy flow patterns. The purpose of this example is twofold. First, we use this example to evaluate the applicability of the developed compressible formulation in the low Mach number limit and to investigate its behavior in unsteady chaotic flows. Second, we will perform some simple tests in order to provide some insights on the actual computational savings that the fractional step implementation offers with respect to the standard monolithic solver.

#### 8.3.1. 2D flow in a differentially heated cavity with aspect ratio 8: $Pr = 0.71$, $Ra = 1 \times 10^6$

Let us start by solving a 2D configuration. The computational domain is given by a rectangular cavity $[0, L] \times [0, H]$ with $H/L = 8$ and $L = 1$ m. The prescription of boundary conditions is as follows: the temperatures on the left and right

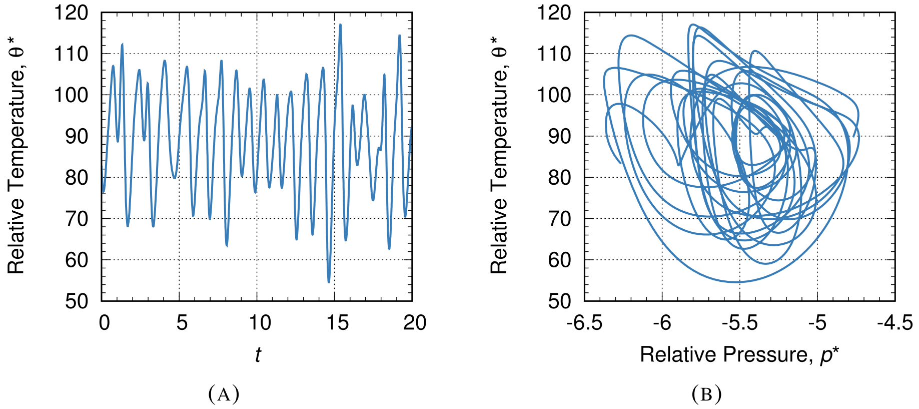


**Fig. 7.** 2D differentially heated cavity: (a) temperature evolution over time and (b) phase diagram pressure-temperature with values from the chosen control point (0.05,4).

boundaries, respectively hot and cold walls, are fixed to $\theta_{\text{g,hot}}$ = 600 K and $\theta_{\text{g,cold}}$ = 300 K, while adiabatic boundary conditions are prescribed on the upper and lower walls. No slip conditions are set for the velocity over all the boundaries. No pressure boundary conditions are prescribed for this example, yet we include an iterative penalization term of the form $\left\langle \psi(p_h^{n+1,(i+1)} - p_h^{n+1,(i)}), q_h \right\rangle$ to the mass conservation equation at the $i$th iteration. This aims to overcome the mechanical pressure resolution for transient and confined flows, where small local oscillations of pressure might generate overall distortions on the global pressure levels. The factor $\psi$ needs to be defined in such a way that it does not detriment neither the algebraic solver nor the convergence of the iterative procedure. For this simulation, we take $\psi = 1 \times 10^{-6}/\mu$, definition which renders this new term dimensionally consistent. The initial conditions to start the computations are prescribed all over the whole domain as: $\mathbf{u}_0 = \mathbf{0}$ m/s, $p_0 = p_{\text{atm}} = 152525$ Pa and $\theta_0 = \theta_{\text{atm}} = 450$ K. The gravity acceleration is specified with a modulus $|\mathbf{g}| = 9.81$ m/s$^2$, acting on the negative $y$ direction. For this problem we consider $\text{Pr} = 0.71$ and $\text{Ra} = 1 \times 10^6$, conditions which are reported in the literature as chaotic. The Rayleigh number is computed here as,

$$\text{Ra} = \frac{|\mathbf{g}|\rho^2 \Delta\theta c_{\text{p}}}{\kappa\mu} \quad \text{with} \quad \Delta\theta := 2\frac{\theta_{\text{g,hot}} - \theta_{\text{g,cold}}}{\theta_{\text{g,hot}} + \theta_{\text{g,cold}}}, \tag{49}$$

where $\Delta\theta$ stands for the dimensionless temperature ratio. The remaining parameters for the simulation are $\mu = 2.5 \times 10^{-3}$ kg/(m s) and $\kappa = 3.55$ W/(m K). The results down below are obtained from an structured mesh of 20,451 nodal points containing 20,000 bilinear quadrilateral elements. The time step size chosen for the computations is $\Delta t = 0.005$ s.

In this problem, a cyclic flow is established around the cavity due to the combination of both hot and cold shear layers. In order to analyze the dynamical response, in Fig. 7 we have plotted the temperature evolution at a control point in the horizontal middle plane of the cavity, as well as the pressure-temperature cycle. The temperature evolution is completely smooth and the pressure-temperature cycle shows the capability of the fractional step scheme to reproduce, without oscillations, the randomness associated with chaotic flows. Should the flow be non-chaotic the pressure-temperature plot would be a completely closed curve. In Fig. 8 snapshots of temperature and velocity contours at a certain time step and for the BDF2 version of the fractional step algorithm are shown. The flow patterns replicated in these images agree with instant contours of fields previously presented in the literature for the same Rayleigh number and cavity ratio (see e.g. results in [61]).

In order to further assess the transient behavior of the flow field and quantitatively compare the performance of the algorithms, we present the calculations of the non-dimensional Nusselt number, which we compute here in the cold and hot walls as,

$$\text{Nu}(\mathbf{x}, t) := \frac{L}{\theta_{\text{g,hot}} - \theta_{\text{g,cold}}} \int_{\partial\Omega} \mathbf{n} \cdot \nabla\theta(\mathbf{x}, t)\, \text{d}\partial\Omega, \qquad \mathbf{x} \in \partial\Omega, t > 0. \tag{50}$$

We can observe the unsteady character of the Nusselt number in Fig. 9 for both hot and cold walls, where we also compare the values obtained with both monolithic and fractional step schemes. The computed average Nusselt numbers are, respectively, $\overline{\text{Nu}}_{\text{FS,hot}} = 41.8162$, $\overline{\text{Nu}}_{\text{MN,hot}} = 41.8908$ and $\overline{\text{Nu}}_{\text{FS,cold}} = -46.2698$, $\overline{\text{Nu}}_{\text{MN,cold}} = -46.2418$. We observe that the predicted values of the segregated computation show an agreement with the ones from the monolithic counterpart. However, there exist some differences which should come, precisely, from the temporal error associated to the fractional step scheme, as already stated.

Now that we have checked that the solutions from both monolithic and fractional step schemes are significantly comparable (up to the splitting error, which was shown to maintain the temporal accuracy in time), we now include a comparison

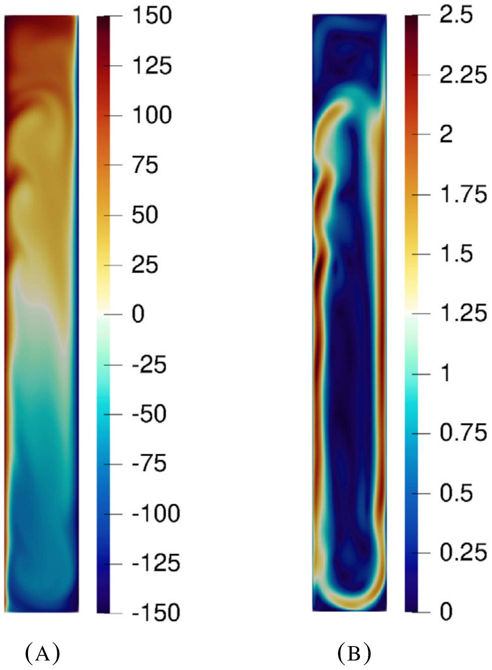


**Fig. 8.** 2D differentially heated cavity: (a) relative temperature $\theta^*$ and (b) velocity magnitude $\|\mathbf{u}\|$ contours of the fractional step scheme. Monolithic contour plots are omitted since they are remarkably similar.

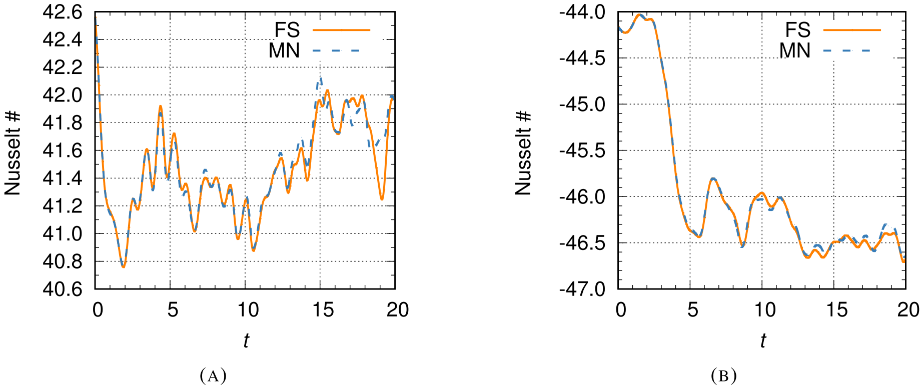


**Fig. 9.** 2D differentially heated cavity: Nusselt number evolution on (a) hot wall, (b) cold wall computed with both fractional step and monolithic second order schemes.

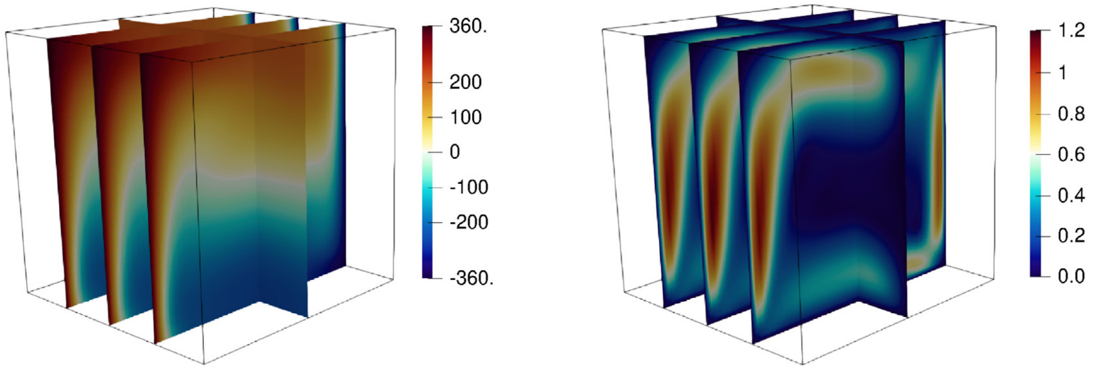
on the actual performance of the monolithic time integrator against the fractional step version, with an emphasis on the CPU time savings. We include the quotient between the CPU time of the fractional step scheme over the CPU time of the corresponding monolithic counterpart. A simple sequential implementation over a series of time steps has been used in all cases, so that to avoid any bias in the data resulting from a parallel computation. To this end, the results are presented in Table 2 and there we distinguish between the total CPU time needed by the algorithm and the time needed by the PETSC solver to obtain the solution of the final algebraic system (the remaining time is spent on assembling the different element contributions, updating variables, allocation and deallocation processes, etc.). As one can observe, the obtained savings in CPU time are remarkable, with a reduction of roughly 78 % of the total time and around 95 % in the solver time for the performed tests. This result is directly related with the information in Table 3, where the number of iterations needed by the solver to obtain the solution of the system of equations and the number of nonlinear iterations used to obtain converged results are collected. The number of iterations needed by the monolithic method is greater both for the linear solver and for the nonlinear algorithm. The subsystems for the different steps of the segregated algorithm are smaller, and in general better conditioned what translates into a substantial diminution on the number of iterations and hence in the time of the computations, as already mentioned. It is observed that the pressure represents the bottleneck in the computations, as it shows the greatest number of solver iterations, and this is in fact the reason why the monolithic case delays the conver-

**Table 2**
2D differentially heated cavity: ratio of CPU time of the fractional step schemes over the monolithic formulations in the 2D chaotic case (second order schemes).

| Case | Total time ratio | Solver time ratio |
|---|---|---|
| $\mathrm{Pr}=0.71$, $\mathrm{Ra}=1\times10^6$ | 0.2263 | 0.054 |

**Table 3**
2D differentially heated cavity: number of iterations of the monolithic and fractional step algorithms (second order schemes). Here, $\overline{\mathrm{nni}}$ is the average number of nonlinear iterations to achieve convergence and $\overline{\mathrm{nsi}}$ stands for the average number of iterations needed by the iterative algebraic solver.

| | Monolithic | Fractional step | | |
|---|---|---|---|---|
| Case | $\overline{\mathrm{nni}}/\overline{\mathrm{nsi}}$ | $\overline{\mathrm{nni}_{\mathbf{u}}}/\overline{\mathrm{nsi}_{\mathbf{u}}}$ | $\overline{\mathrm{nni}_{\theta}}/\overline{\mathrm{nsi}_{\theta}}$ | $\overline{\mathrm{nni}_{p}}/\overline{\mathrm{nsi}_{p}}$ |
| $\mathrm{Pr}=0.71$, $\mathrm{Ra}=1\times10^6$ | 8/105 | 3/3 | 3/2 | 2/146 |



(A) (B)

**Fig. 10.** 3D differentially heated cavity: contour plots of (a) relative temperature $\theta^*$ and (b) velocity magnitude $\|\mathbf{u}\|$ at the final time step of the simulation using the BDF2 fractional step scheme.

gence of the total system. Likewise, the greater number of nonlinear iterations needed by the monolithic method confirms the better treatment of the nonlinearities when the problem is solved in a decoupled manner.

### 8.3.2. *3D flow in a differentially heated cubic cavity: $Pr = 0.71$, $Ra = 3.5\times10^5$*

We will now consider a 3D version of the previous cavity problem. The computational domain is defined now as the unit cube $\overline{\Omega}=[0,L]^3$ with $L=1$ m. The prescription of boundary conditions is similar as in the previous example: the temperatures on the boundaries perpendicular to the $x$-coordinate (horizontal) are fixed to $\theta_{\mathrm{g,hot}}$ = 960 K (most left boundary) and $\theta_{\mathrm{g,cold}}$ = 240 K (most right boundary), while adiabatic boundary conditions are prescribed on the remaining walls. No slip conditions are set for the velocity over all the walls. Furthermore, the initial conditions we consider now are: $\mathbf{u}_0=\mathbf{0}$ m/s, $p_0=p_{\mathrm{atm}}=101325$ Pa and $\theta_0=\theta_{\mathrm{atm}}=600$ K. The gravity acceleration is specified to be acting in the negative $y$-coordinate direction (vertical) with a modulus $|\mathbf{g}|=9.81$ m/s$^2$. For this problem we consider $\mathrm{Pr}=0.71$ and $\mathrm{Ra}=3.5\times10^5$ (non-chaotic case), with both non-dimensional numbers defined as in the previous 2D problem and with $\Delta\theta=1.2$. The remaining parameters are set so that to match the desired flow conditions. We perform the simulation using an uniform structured mesh composed of 64,000 hexahedron elements and we use a constant time step of $\Delta t=0.02$ s.

In Fig. 10 some graphical results for this 3D cavity using the second order integrator (both the BDF1 fractional step scheme and the monolithic case provide practically identical contour plots and they are not included for the sake of conciseness).

We will again compare the CPU times for the different algorithms selected for this problem as well the number of iterations needed by the solver and the non-linear iterative scheme. All these results are included in Table 4 and Table 5, respectively. For the monolithic formulation, the total number of iterations of the linear system solver is really affected by the pressure. This seems to be a general trend for systems with primitive variables, as also discussed in [21] for the non-Newtonian incompressible case. This fact is one of the main downsides when compared to a fractional step algorithm, in which each step requires a different number of iterations to solve the corresponding linear system. Finally, although in these examples we have used the same solver for all subsystems arising in the segregation method, specific solving techniques could be exploited in order to improve the performance of fractional step schemes even further.

**Table 4**
3D differentially heated cavity: ratio of CPU times of the fractional step schemes over the monolithic formulation for the 3D cavity (BDF1 results were equivalent).

| Case | Total time ratio | Solver time ratio |
|---|---|---|
| $\mathrm{Pr} = 0.71$, $\mathrm{Ra} = 3.5 \times 10^5$ | 0.3410 | 0.1016 |

**Table 5**
3D differentially heated cavity: number of iterations of the monolithic and fractional step algorithm in the 3D cavity for the BDF2 scheme (BDF1 results were equivalent).

| | Monolithic | Fractional step | | |
|---|---|---|---|---|
| Case | $\overline{\mathrm{nni}}/\overline{\mathrm{nsi}}$ | $\overline{\mathrm{nni}}_{\mathbf{u}}/\overline{\mathrm{nsi}}_{\mathbf{u}}$ | $\overline{\mathrm{nni}}_{\theta}/\overline{\mathrm{nsi}}_{\theta}$ | $\overline{\mathrm{nni}}_{p}/\overline{\mathrm{nsi}}_{p}$ |
| $\mathrm{Pr} = 0.71$, $\mathrm{Ra} = 3.45 \times 10^5$ | 8/32 | 3/6 | 3/2 | 2/63 |

## 9. Conclusions

In this article, we have introduced a pressure-segregation technique to solve the compressible Navier–Stokes equations using primitive variables. The development of this new methodology, which is up to second order in time, has been designed at the pure algebraic level, considering as starting point the fully discretized monolithic problem both in space and in time and by using the extrapolation concept.

From the numerical viewpoint, the herein proposed fractional step compressible model is based on a stabilized VMS method and an implicit scheme to advance in time. In addition, other ingredients were appended, such as the orthogonal and dynamic definition of subscales, the weak imposition of Dirichlet boundary conditions, the application of non-reflecting boundary conditions (a major issue in low Mach compressible solvers) and the decomposition of the pressure and temperature unknowns into relative plus atmospheric part in order to solve nearly incompressible cases.

We have shown that the fractional step method introduces an splitting error which maintains the general temporal accuracy of the time integration scheme. Furthermore, the differentially heated cavity problem has been used to test the performance of the algorithm in dynamic and chaotic cases and the possibility of directly computing acoustic pressure waves has been verified with the simulation of the aeolian tones problem, for which the dumping of acoustic perturbations at the external computational boundaries was also evaluated. Likewise, the fractional step formulation has been shown to be efficient since an important reduction in the CPU time has been obtained with respect to the monolithic counterpart for the tests performed, which have a relatively small number of degrees of freedom; much more important savings should be expected in large scale problems.

## CRediT authorship contribution statement

**Samuel Parada:** Conceptualization, Methodology, Software, Validation, Writing – review & editing. **Ramon Codina:** Conceptualization, Methodology, Writing – review & editing. **Joan Baiges:** Conceptualization, Methodology, Software, 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.

## Acknowledgements

Samuel Parada gratefully acknowledges the support received from the Agència de Gestió d'Ajuts Universitaris i de Recerca through the predoctoral FI grant 2019-FI-B-00607. Joan Baiges gratefully acknowledges the support of the Spanish Government through the Ramón y Cajal grant RYC-2015-17367. Ramon Codina gratefully acknowledges the support received from the ICREA Acadèmia Research Program of the Catalan Government. This work was partially funded through the TOP-FSI: RTI2018-098276-B-I00 project of the Spanish Government. CIMNE is a recipient of a "Severo Ochoa Programme for Centers of Excellence in R&D" grant (CEX2018-000797-S) by the Spanish Ministry of Economy and Competitiveness.

## References

[1] J. Wong, D. Darmofal, J. Peraire, The solution of the compressible Euler equations at low Mach numbers using a stabilized finite element algorithm, Comput. Methods Appl. Mech. Eng. 190 (43) (2001) 5719–5737.
[2] J. Hardin, D. Pope, An acoustic/viscous splitting technique for computational aeroacoustics, Theor. Comput. Fluid Dyn. 6 (5–6) (1994) 323–340.
[3] W.Z. Shen, J.N. Sørensen, Aeroacoustic modelling of low-speed flows, Theor. Comput. Fluid Dyn. 13 (4) (1999) 271–289.
[4] W.Z. Shen, W. Zhu, J.N. Sørensen, Aeroacoustic computations for turbulent airfoil flows, Amer. Inst. Aeronaut. Astronaut. 47 (6) (2009) 1518–1527.

[5] O. Guasch, R. Codina, Computational aeroacoustics of viscous low speed flows using subgrid scale finite element methods, J. Comput. Acoust. 17 (3) (2009) 309–330.

[6] J. Principe, R. Codina, Mathematical models for thermally coupled low speed flows, Adv. Theor. Appl. Mech. 2 (2009) 93–112.

[7] M. Avila, J. Principe, R. Codina, A finite element dynamical nonlinear subscale approximation for the low Mach number flow equations, J. Comput. Phys. 230 (2011) 7988–8009.

[8] O. Zienkiewicz, R. Codina, A general algorithm for the compressible and incompressible flows. Part I: the split, characteristic-based scheme, Int. J. Numer. Methods Fluids 20 (8–9) (1995) 869–885.

[9] L. Pesch, J.J. van der Vegt, A discontinuous Galerkin finite element discretization of the Euler equations for compressible and incompressible fluids, J. Comput. Phys. 11 (2008) 5426–5446.

[10] T.J. Hughes, L. Franca, M. Mallet, A new finite element formulation for computations fluid dynamics. I. Symmetric forms of the compressible Euler and Navier-Stokes equations and the second law of thermodynamics, Comput. Methods Appl. Mech. Eng. 54 (2) (1986) 223–234.

[11] G. Hauke, T.J. Hughes, A comparative study of different sets of variables for solving compressible and incompressible flows, Comput. Methods Appl. Mech. Eng. 153 (1–2) (2007) 1–44.

[12] D.H. Rudy, J.C. Strikwerda, A nonreflecting outflow boundary conditions for subsonic Navier-Stokes calculation, J. Comput. Phys. 36 (1) (1980) 55–70.

[13] T. Colonius, S.K. Lele, P. Moin, Boundary conditions for direct computation of aerodynamic sound generation, Amer. Inst. Aeronaut. Astronaut. 31 (9) (1993) 1574–1582.

[14] P. Fosso, H. Deniau, N. Lamarque, T. Poinsot, Comparison of outflow boundary conditions for subsonic aeroacoustic simulations, Int. J. Numer. Methods Fluids 68 (10) (2012) 1207–1233.

[15] T. Colonious, Modeling artificial boundary conditions for compressible flow, Annu. Rev. Fluid Mech. 36 (2004) 315–345.

[16] A.J. Chorin, A numerical method for solving incompressible viscous problems, J. Comput. Phys. 2 (1967) 12–26.

[17] R. Teman, Sur l'approximation de la solution des equations de Navier–Stokes par la méthode des pas fractionaires (I), Arch. Ration. Mech. Anal. 32 (1969) 135–153.

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

[19] S. Badia, R. Codina, Algebraic pressure segregation methods for the incompressible Navier–Stokes equations, Arch. Comput. Methods Eng. 15 (3) (2007) 1–52.

[20] S. Badia, R. Codina, Pressure segregation methods based on a discrete pressure Poisson equation. An algebraic approach, Int. J. Numer. Methods Fluids 56 (2008) 351–382.

[21] E. Castillo, R. Codina, First, second and third order fractional step methods for the three-field viscoelastic flow problem, J. Comput. Phys. 296 (2015) 113–137.

[22] S. Parada, J. Baiges, R. Codina, A fractional step method for computational aeroacoustics using weak imposition of Dirichlet boundary conditions, Comput. Fluids 197 (2020).

[23] J.B. Perot, An analysis of the fractional step method, J. Comput. Phys. 108 (1993) 51–58.

[24] R. Codina, Pressure stability in fractional step finite element methods for incompressible flows, J. Comput. Phys. 170 (2001) 112–140.

[25] J.M. Hérard, O. Hurisse, A fractional step method to compute a class of compressible gas-liquid flows, Comput. Fluids 55 (2012) 57–69.

[26] T. Gallouët, L. Gastaldo, J.C.L.R. Herbin, An unconditionally stable pressure correction scheme for the compressible barotropic Navier-Stokes equations, ESAIM: Math. Model. Numer. Anal. 42 (2) (2008) 303–331.

[27] K. Liu, R.H. Pletcher, A fractional step method for solving the compressible Navier–Stokes equations, J. Comput. Phys. 226 (2) (2007) 1930–1951.

[28] R. Codina, M. Vázquez, O. Zienkiewicz, A general algorithm for the compressible and incompressible flows. Part III: the semi-implicit form, Int. J. Numer. Methods Fluids 27 (1996) 13–32.

[29] G. Hauke, A. Landaberea, I. Garmendia, J. Canales, A segregated method for compressible flow computation. Part II: general divariant compressible flows, Int. J. Numer. Methods Fluids 49 (2005) 183–209.

[30] D. Grapsas, R. Herbin, W. Kheriji, J.C. Latché, An unconditionally stable staggered pressure correction scheme for the compressible Navier-Stokes equations, SMAI J. Comput. Math. 2 (2016) 51–97.

[31] T. Hughes, G. Feijóo, L. Mazzei, J. Quincy, The variational multiscale method – a paradigm for computational mechanics, Comput. Methods Appl. Mech. Eng. 166 (1998) 3–24.

[32] T.J. Hughes, Multiscale phenomena: Greenś functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods, Comput. Methods Appl. Mech. Eng. 127 (1) (1995) 387–401.

[33] T. Hughes, G. Scovazzi, L. Franca, Multiscale and stabilized methods, in: Encyclopedia of Computational Mechanics, Wiley, 2017.

[34] R. Codina, S. Badia, J. Baiges, J. Príncipe, Variational multiscale methods in computational fluid dynamics, in: Encyclopedia of Computational Mechanics, second edition, 2018, pp. 1–28.

[35] T.J.R. Hughes, G. Scovazzi, T.E. Tezduyar, Stabilized methods for compressible flows, J. Sci. Comput. 43 (2010) 343–368.

[36] F. Rispoli, R. Saavedra, A stabilized finite element method based on sgs models for compressible flows, Comput. Methods Appl. Mech. Eng. 196 (2006) 652–664.

[37] R. Codina, Stabilized finite element approximation of transient incompressible flows using orthogonal subscales, Comput. Methods Appl. Mech. Eng. 191 (2002) 4295–4321.

[38] R. Codina, J. Principe, O. Guasch, S. Badia, Time dependent subscales in the stabilized finite element approximation of incompressible flow problems, Comput. Methods Appl. Mech. Eng. 196 (2007) 2413–2430.

[39] R. Codina, J. Principe, Dynamic subscales in the finite element approximation of thermally coupled incompressible flows, Int. J. Numer. Methods Fluids 54 (2007) 707–730.

[40] G.K. Batchelor, An Introduction to Fluid Dynamics. Cambridge Mathematical Library, Cambridge University Press, 2000.

[41] A. Pont, R. Codina, J. Baiges, O. Guasch, Unified solver for fluid dynamics and aeroacoustics in isentropic gas flows, J. Comput. Phys. 363 (2018) 11–29.

[42] H. Espinoza, R. Codina, S. Badia, A Sommerfeld non-reflecting boundary condition for the wave equation in mixed form, Comput. Methods Appl. Mech. Eng. 276 (2014) 122–148.

[43] A. Salazar, Energy propagation of thermal waves, Eur. J. Phys. 27 (6) (2006) 1349–1355.

[44] H. Owen, R. Codina, A third-order velocity correction scheme obtained at the discrete level, Int. J. Numer. Methods Fluids 69 (2012) 57–72.

[45] R. Codina, A. Folch, A stabilized finite element predictor–corrector scheme for the incompressible Navier–Stokes equations using a nodal–based implementation, Int. J. Numer. Methods Fluids 44 (2004) 483–503.

[46] R. Codina, J. Príncipe, J. Baiges, Subscales on the element boundaries in the variational two-scale finite element method, Comput. Methods Appl. Mech. Eng. 198 (2009) 838–852.

[47] C. Bayona, J. Baiges, R. Codina, Variational multi-scale finite element solution of the compressible Navier–Stokes equations, Int. J. Numer. Methods Heat Fluid Flow 26 (2015) 1240–1271.

[48] C. Bayona, J. Baiges, R. Codina, Solution of low Mach number aeroacoustic flows using a variational multi-scale finite element formulation of the compressible Navier–Stokes equations written in primitive variables, Comput. Methods Appl. Mech. Eng. 344 (2019) 1073–1103.

[49] H.A.V. der Vorst, Bi-CGSTAB: a fast and smoothly converging variant of Bi-CG for the solution of nonsymmetric linear systems, SIAM J. Sci. Stat. Comput. 13 (2) (1992) 631–644.

[50] S. Balay, S. Abhyankar, M.F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W.D. Gropp, D. Kaushik, M.G. Knepley, D.A. May, L.C. McInnes, R.T. Mills, T. Munson, K. Rupp, P. Sanan, B.F. Smith, S. Zampini, H. Zhang, H. Zhang, PETSc web page, http://www.mcs.anl.gov/petsc, 2015.

[51] J.L. Guermond, J. Shen, Velocity-correction projection method for incompressible flows, SIAM J. Numer. Anal. 41 (2003) 112–134.

[52] P.R. Amestoy, I.S. Duff, J. Koster, J.-Y. L'Excellent, A fully asynchronous multifrontal solver using distributed dynamic scheduling, SIAM J. Matrix Anal. Appl. 23 (1) (2001) 15–41.

[53] P.R. Amestoy, A. Guermouche, J.-Y. L'Excellent, S. Pralet, Hybrid scheduling for the parallel solution of linear systems, Parallel Comput. 32 (2) (2006) 136–156.

[54] F. Margnat, Hybrid prediction of the aerodynamic noise radiated by a rectangular cylinder at incidence, Comput. Fluids 109 (2015).

[55] Y. Liow, B.T. Tan, M.C. Thompson, K. Hourigan, Sound generated in laminar flow past a two-dimensional rectangular cylinder, J. Sound Vib. 295 (1–2) (2006) 407–427.

[56] B. Mahato, N. Ganta, Y.G. Bhumka, Direct simulation of sound generation by a two-dimensional flow past a wedge, Phys. Fluids 30 (o) (2018) 096101.

[57] O. Inoue, N. Hatakeyama, Sound generation by a two-dimensional circular cylinder in a uniform flow, J. Fluid Mech. 471 (2002) 285–314.

[58] A. Laffite, F. Pérot, Investigation of the noise generated by cylinder flows using a direct lattice-Boltzmann approach, in: 15th AIAA/CEAS Aeroacoustics Conference (30th AIAA Aeroacoustics Conference), 2009.

[59] N. Ganta, B. Mahato, Y.G. Bhumka, Analysis of sound generation by flow past a circular cylinder performing rotary oscillations using direct simulation approach, Phys. Fluids 31 (026104) (2019).

[60] M.A. Christon, P.M. Gresho, S.B. Sutton, Computational predictability of time-dependent natural convection flows in enclosures (including a benchmark solution), Int. J. Numer. Methods Heat Fluid Flow 40 (8) (2002) 953–980.

[61] R. Codina, J. Principe, M. Avila, Finite element approximation of turbulent thermally coupled incompressible flows with numerical sub-grid scale modelling, Int. J. Numer. Methods Heat Fluid Flow 20 (2010) 492–516.