SEMI-IMPLICIT FORMULATIONS OF THE NAVIER-STOKES EQUATIONS: APPLICATION TO NONHYDROSTATIC ATMOSPHERIC MODELING

F.X. GIRALDO *, M. RESTELLI †, AND M. LÄUTER ‡

Abstract. We present semi-implicit (IMEX) formulations of the compressible Navier-Stokes equations (NSE) for applications in nonhydrostatic atmospheric modeling. The compressible NSE in nonhydrostatic atmospheric modeling include buoyancy terms that require special handling if one wishes to extract the Schur complement form of the linear implicit problem. We present results for five different forms of the compressible NSE and describe in detail how to formulate the semi-implicit time-integration method for these equations. Finally, we compare all five equations and compare the semi-implicit formulations of these equations both using the *Schur* and *No Schur* forms against an explicit Runge-Kutta method. Our simulations show that, if efficiency is the main criterion, it matters which form of the governing equations you choose. Furthermore, the semi-implicit formulations are faster than the explicit Runge-Kutta method for all the tests studied especially if the *Schur* form is used. While we have used the spectral element method for discretizing the spatial operators, the semi-implicit formulations that we derive are directly applicable to all other numerical methods. We show results for our five semi-implicit models for a variety of problems of interest in nonhydrostatic atmospheric modeling, including: inertia gravity waves, rising thermal bubbles (i.e., Rayleigh-Taylor instabilities), density current (i.e., Kelvin-Helmholtz instabilities), and mountain test cases; the latter test case requires the implementation of non-reflecting boundary conditions. Therefore, we show results for all five semi-implicit models using the appropriate boundary conditions required in nonhydrostatic atmospheric modeling: no-flux (reflecting) and non-reflecting boundary conditions. It is shown that the non-reflecting boundary conditions exert a strong impact on the accuracy and efficiency of the models.

Key words. compressible flow; element-based Galerkin methods; Euler; IMEX; Lagrange; Legendre; Navier-Stokes; nonhydrostatic; spectral elements; time-integration.

AMS subject classifications. 65M60, 65M70, 35L65, 86A10

1. Introduction. It can be argued that the single most important property of an operational nonhydrostatic mesoscale atmospheric is efficiency. Clearly, this efficiency should not come at the cost of accuracy but if a weather center has the choice between a very accurate model and one that is efficient, they will probably pick the efficient one; however, as numerical analysts, we would like to build models that are both accurate and efficient. One way to achieve this goal is to construct numerical models based on high-order methods: this class of methods offers exponential (spectral) convergence for smooth problems and achieves excellent scalability on modern multi-core systems if they are used in an element-based approach (i.e., if the approximating polynomials have compact/local support). This is the idea behind element-based Galerkin methods such as spectral element (SE) and discontinuous Galerkin (DG) methods (see [14] and [27] for nonhydrostatic models based on these methods) and in this work we use the SE method to approximate spatial derivatives. Almost all nonhydrostatic mesoscale models currently in existence are based on the finite difference (FD) method. The only nonhydrostatic atmospheric models not based on the FD method are the finite volume (FV) models found in [4], [2], and [1], and our SE and DG models found in [14] and [27]. One of the biggest advantages that finite element (FE), SE, and DG methods have over the FD method is that no terrain following coordinates of the type introduced in [10] need to be included in the governing equations. Of course, the orography (e.g., mountains) has to be accounted for in some manner but element-based Galerkin (EBG) methods, such as FE, SE, FV, and DG, incorporate the orography via the definition of the grid. EBG methods do not require either orthogonal grids (see [13, 17, 25, 15]) or grids with specific directions (such as the I and J indices in FD models); EBG models are inherently unstructured and, while requiring additional data structures for bookkeeping, completely liberate the method from the grid. This freedom from the grid has major repercussions in the implementation of these methods on distributed-memory computers in that no halo is required which translates to truly local algorithms that require very little communication across processors; instead, the communication stencil consists of the perimeter values of each processor (see [16] and

*Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93943, USA

†Ecuaciones Diferenciales y Análisis Numérico, Universidad de Sevilla, Facultad de Matemáticas, C/ Tarfia s/n, 41012 Sevilla, Spain

‡Alfred Wegener Institute for Polar and Marine Research in the Helmholtz Association, P.O. Box 60 01 49, 14401 Potsdam, Germany

[12]). Another advantage that FE, SE, and DG methods have over the FD and FV methods is that high-order solutions (greater than fourth order) can be constructed quite naturally within the framework - such high-order properties are desirable because they reduce the dispersion errors associated with the discrete spatial operators [11]. In fact, the SE formulation used in this paper allows for arbitrarily high-order spatial operators to be constructed by an input parameter; all the results presented in Sec. 4 use either 8th or 10th order polynomials per element.

Once the spatial discretization method has been selected, one is then faced with choosing a method for evolving time-dependent partial differential equations forward in time. The simplest choice is to use explicit time-integrators (e.g., Runge-Kutta methods) but these may not be the most efficient methods to use especially taking the following two points into consideration: 1) methods that are high-order in space require a much smaller time-step than low-order methods because the time-step is proportional to the polynomial order and 2) the fastest waves in the compressible Navier-Stokes equations are the acoustic waves that have little or no effect on the large-scale processes in the linear regime. The fact that the acoustic waves are so fast but have little significance in the accuracy of the simulations means that if one uses explicit methods, then one must adhere to a very small time-step restriction caused by a physical phenomenon that is essentially inconsequential. To overcome these issues almost all operational nonhydrostatic weather models use split-explicit methods [23] where the fast acoustic waves use a smaller time-step while the slower waves use a larger time-step, typically using a time-integration strategy based on explicit Runge-Kutta methods. Examples of models that use this approach include the operational models of the U.S. Navy [19], the National Center for Atmospheric Research [22], Penn State/NCAR [29], U.S. National Center for Environmental Prediction [20], German Weather Service [30], and the Japanese Meteorological Agency [28], to name only a few. Some centers have experimented with semi-implicit approaches but have found them lacking with respect to the currently used explicit approach [37].

To construct semi-implicit formulations (i.e., IMEX) that are competitive with the explicit approach currently used by all operational models requires the development of state-of-the-art iterative solvers and preconditioners. Our current work is a step towards building such models and, here, we show that the semi-implicit formulations are indeed more efficient than explicit Runge-Kutta methods, at least for our spatial discretization methods (high-order spectral element methods); however, our results should hold for all other spatial discretization methods, the construction of the Schur complement form for Godunov-type methods requires special treatment (see [27] for issues facing these methods); we will extend these results to Godunov-type methods shortly. The next step will be to show that semi-implicit formulations in all directions (as we have done here) are more efficient than semi-implicit formulations along the vertical; this we shall do in a future paper.

The remainder of the paper is organized as follows. Section 2 describes the five forms of the equations that we study. In Sec. 3 we describe the semi-implicit method used to march the equations in time. In this section, we discuss in detail the construction of the semi-implicit operators for all five equation sets and describe how to extract the Schur complement that is necessary in order to further increase the efficiency of the semi-implicit models. In Sec. 4 we present the results for all five semi-implicit models using four test cases. In addition, we compare the efficiency of an explicit method with the semi-implicit methods both with the Schur and No Schur forms. Finally, in Sec. 5 we summarize the key findings of this research and propose future directions.

2. Governing Equations. In this paper we study five different forms of the equations that govern the dynamics of nonhydrostatic atmospheric processes, namely the compressible Euler equations including the gravitational force and a diffusion-like term. Depending on the form of the diffusion term, the complete compressible Navier-Stokes equations can be recovered. Specifically, we study the following equation sets:

1. (set 1) the non-conservative form using Exner pressure, velocity, and potential temperature,
2. (set 2NC) the non-conservative form using density, velocity, and potential temperature,
3. (set 2C) the conservative form using density, momentum, and potential temperature density,
4. (set 3) the conservative form using density, momentum, and total energy, and
5. (set 4) the non-conservative form using density, velocity, and pressure.

For the purposes of this study we restrict ourselves to two dimensions (x-z) and omit the Coriolis terms. These two assumptions place no restrictions on the analysis of this paper but they do simplify the discussion considerably. Compared to standard problems considered in computational fluid dynamics, a distinctive feature of atmospheric flows is the important role played by the gravitational force, resulting in a vertically stratified fluid. In fact, the vertical profiles of pressure, density, and temperature are determined to first order by the hydrostatic balance, and nonhydrostatic effects typically represent perturbations from this equilibrium condition. This fact poses some challenges to prospective numerical methods and is usually dealt with by introducing a fixed hydrostatic state and using as prognostic variables the nonhydrostatic deviations from this state. We use this approach in the present work and describe it in more detail in the following summary of the considered equation sets. The fixed hydrostatic reference state will also prove useful in the construction of the semi-implicit time integrator. Let us now describe each of the five equations that we compare.

2.1. Equation Set 1 (SE1). Since none of the prognostic variables used in the SE1 equation set represents a conserved quantity, it is natural to state the problem in non-conservation form. We thus consider the system

$$\frac{\partial \pi}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \pi + (\gamma - 1)\pi \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + c_p \theta \boldsymbol{\nabla} \pi + g\boldsymbol{k} = \mu \nabla^2 \boldsymbol{u}$$

$$\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta = \mu \nabla^2 \theta \tag{2.1}$$

where the solution vector is $(\pi, \boldsymbol{u}^{\mathcal{T}}, \theta)^{\mathcal{T}}$, $\pi = \left(\frac{P}{P_A}\right)^{R/c_p}$ is the Exner pressure, $\boldsymbol{u} = (u, w)^{\mathcal{T}}$ is the velocity field, $\theta = \frac{T}{\pi}$ is the potential temperature, and $\mathcal{T}$ denotes the transpose operator. In these equations P is the pressure, P_A is a constant reference pressure at the surface ($P_A = 1 \times 10^5$ Pa) and T is the temperature. Other variables and symbols requiring definition are the gradient operator $\boldsymbol{\nabla} = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial z}\right)^{\mathcal{T}}$, the gravitational constant g, the gas constant $R = c_p - c_v$, the specific heats for constant pressure and volume, c_p and c_v, the specific heat ratio $\gamma = c_p/c_v$, and the directional vector along the vertical (z) direction $\boldsymbol{k} = (0, 1)^{\mathcal{T}}$.

Introducing the following splitting of the Exner pressure $\pi(\boldsymbol{x}, t) = \pi_0(z) + \pi'(\boldsymbol{x}, t)$ and potential temperature $\theta(\boldsymbol{x}, t) = \theta_0(z) + \theta'(\boldsymbol{x}, t)$ where the reference values are in hydrostatic balance, i.e., $c_p \theta_0 \frac{d\pi_0}{dz} = -g$, allows us to rewrite Eq. (2.1) as

$$\frac{\partial \pi'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \pi' + w\frac{d\pi_0}{dz} + (\gamma - 1)\left(\pi' + \pi_0\right)\boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + c_p \theta \boldsymbol{\nabla} \pi' - g\frac{\theta'}{\theta_0}\boldsymbol{k} = \mu \nabla^2 \boldsymbol{u}$$

$$\frac{\partial \theta'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta' + w\frac{d\theta_0}{dz} = \mu \nabla^2 \theta \tag{2.2}$$

that has been expanded and simplified in order to enforce hydrostasis for zero initial perturbation fields. It should be noted that the viscous terms on the right-hand-side of the momentum and energy equations are not the *true* Navier-Stokes viscous stresses but rather are *ad hoc* terms used to satisfy one of the test cases (i.e., the density current). We shall use a similar diffusion operator for all equation sets except for set 3 where it is natural to use the true viscous stresses.

2.2. Equation Set 2NC (SE2NC). These equations are written as follows:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = 0$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho}\boldsymbol{\nabla} P + g\boldsymbol{k} = \mu \nabla^2 \boldsymbol{u}$$

$$\frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta = \mu \nabla^2 \theta \tag{2.3}$$

where the prognostic variables are $(\rho, \boldsymbol{u}^{\mathcal{T}}, \theta)^{\mathcal{T}}$ and ρ is the density. The pressure P that appears in the momentum equation is obtained from the equation of state

$$P = P_A \left(\frac{\rho R \theta}{P_A} \right)^{\gamma}.$$

Introducing the following splitting of the density $\rho(\boldsymbol{x}, t) = \rho_0(z) + \rho'(\boldsymbol{x}, t)$ and potential tempera-ture $\theta(\boldsymbol{x}, t) = \theta_0(z) + \theta'(\boldsymbol{x}, t)$ where the reference values are in hydrostatic balance, i.e., $\frac{dP_0}{dz} = -\rho_0 g$, allows us to rewrite Eq. (2.3) as

$$\frac{\partial \rho'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho' + w \frac{d\rho_0}{dz} + (\rho' + \rho_0) \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho' + \rho_0} \boldsymbol{\nabla} P' + \frac{\rho'}{\rho' + \rho_0} g \boldsymbol{k} = \mu \nabla^2 \boldsymbol{u}$$

$$\frac{\partial \theta'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta' + w \frac{d\theta_0}{dz} = \mu \nabla^2 \theta$$

2.3. Equation Set 2C (SE2C). These equations are written as follows:

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} = 0$$

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P \mathcal{I}_2 \right) + \rho g \boldsymbol{k} = \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\boldsymbol{U}}{\rho} \right)$$

$$\frac{\partial \Theta}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\Theta \boldsymbol{U}}{\rho} \right) = \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\Theta}{\rho} \right) \qquad (2.4)$$

where the conserved, prognostic variables are $(\rho, \boldsymbol{U}^{\mathcal{T}}, \Theta)^{\mathcal{T}}$, $\boldsymbol{U} = (\rho u, \rho w)^{\mathcal{T}}$ is the momentum, $\Theta = \rho \theta$ is the potential temperature density, and $\mathcal{I}_2$ is a rank-2 identity matrix. The pressure P that appears in the momentum equation is obtained from the equation of state

$$P = P_A \left(\frac{R \Theta}{P_A} \right)^{\gamma}.$$

Introducing the following splitting of the density $\rho(\boldsymbol{x}, t) = \rho_0(z) + \rho'(\boldsymbol{x}, t)$ and potential temper-ature density $\Theta(\boldsymbol{x}, t) = \Theta_0(z) + \Theta'(\boldsymbol{x}, t)$ where the reference values are in hydrostatic balance, i.e., $\frac{dP_0}{dz} = -\rho_0 g$, allows us to rewrite Eq. (2.4) as

$$\frac{\partial \rho'}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} = 0$$

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P' \mathcal{I}_2 \right) + \rho' g \boldsymbol{k} = \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\boldsymbol{U}}{\rho} \right)$$

$$\frac{\partial \Theta'}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\Theta \boldsymbol{U}}{\rho} \right) = \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\Theta}{\rho} \right)$$

2.4. Equation Set 3 (SE3). Since these equations, when written in non-conservation form, are quite unwieldy, they are only discussed here in conservation form. We thus consider the system

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} = 0$$

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P \mathcal{I}_2 \right) + \rho g \boldsymbol{k} = \boldsymbol{\nabla} \cdot \boldsymbol{F}_{\boldsymbol{u}}^{\mathrm{visc}}$$

$$\frac{\partial E}{\partial t} + \boldsymbol{\nabla} \cdot \left[\frac{(E + P) \boldsymbol{U}}{\rho} \right] = \boldsymbol{\nabla} \cdot \boldsymbol{F}_e^{\mathrm{visc}} \qquad (2.5)$$

where the conserved, prognostic variables are $(\rho, \boldsymbol{U}^{\mathcal{T}}, E)^{\mathcal{T}}$, $E = \rho c_v T + \frac{1}{2} \frac{\boldsymbol{U} \cdot \boldsymbol{U}}{\rho} + \rho \phi$ is the total energy, and $\phi = gz$ is the geopotential. The pressure P is obtained from the equation of state that,

in terms of the solution variables, reads

$$P = (\gamma - 1) \left(E - \frac{\boldsymbol{U} \cdot \boldsymbol{U}}{2\rho} - \rho\phi \right).$$

The viscous fluxes $\boldsymbol{F}^{\text{visc}}$ are defined as follows:

$$\boldsymbol{F}_{\boldsymbol{u}}^{\text{visc}} = \mu \left[\boldsymbol{\nabla} \boldsymbol{u} + (\boldsymbol{\nabla} \boldsymbol{u})^{\mathcal{T}} + \lambda \left(\boldsymbol{\nabla} \cdot \boldsymbol{u} \right) \boldsymbol{\mathcal{I}}_2 \right]$$

and

$$\boldsymbol{F}_e^{\text{visc}} = \boldsymbol{u} \cdot \boldsymbol{F}_{\boldsymbol{u}}^{\text{visc}} + \frac{\mu c_p}{\text{Pr}} \boldsymbol{\nabla} T$$

where $\lambda = -\frac{2}{3}$ comes from the Stokes hypothesis, Pr is the Prandtl number, and μ is the dynamic viscosity.

Introducing the following splitting of the density $\rho(\boldsymbol{x}, t) = \rho_0(z) + \rho'(\boldsymbol{x}, t)$ and total energy $E(\boldsymbol{x}, t) = E_0(z) + E'(\boldsymbol{x}, t)$, where ρ_0 and E_0 are in hydrostatic balance, allows us to rewrite Eq. (2.5) as

$$\frac{\partial \rho'}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} = 0$$

$$\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P' \boldsymbol{\mathcal{I}}_2 \right) + \rho' g \boldsymbol{k} = \boldsymbol{\nabla} \cdot \boldsymbol{F}_{\boldsymbol{u}}^{\text{visc}}$$

$$\frac{\partial E'}{\partial t} + \boldsymbol{\nabla} \cdot \left[\frac{(E + P) \boldsymbol{U}}{\rho} \right] = \boldsymbol{\nabla} \cdot \boldsymbol{F}_e^{\text{visc}}$$

where the system satisfies hydrostasis for zero initial perturbation fields.

2.5. Equation Set 4 (SE4). As for SE1, it is natural to consider this equation set in the non-conservation form

$$\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) = 0$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho} \boldsymbol{\nabla} P + g \boldsymbol{k} = \mu \nabla^2 \boldsymbol{u}$$

$$\frac{\partial P}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} P + \gamma P \boldsymbol{\nabla} \cdot \boldsymbol{u} = \mu \gamma \frac{P}{\theta} \nabla^2 \theta \qquad (2.6)$$

where the prognostic variables are $(\rho, \boldsymbol{u}^{\mathcal{T}}, P)^{\mathcal{T}}$.

Introducing the following splitting of the density $\rho(\boldsymbol{x}, t) = \rho_0(z) + \rho'(\boldsymbol{x}, t)$ and pressure $P(\boldsymbol{x}, t) = P_0(z) + P'(\boldsymbol{x}, t)$ where the reference values are in hydrostatic balance, i.e., $\frac{dP_0}{dz} = -\rho_0 g$, allows us to rewrite Eq. (2.6) as

$$\frac{\partial \rho'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho' + w \frac{d\rho_0}{dz} + (\rho' + \rho_0) \boldsymbol{\nabla} \cdot \boldsymbol{u} = 0$$

$$\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho' + \rho_0} \boldsymbol{\nabla} P' + \frac{\rho'}{\rho' + \rho_0} g \boldsymbol{k} = \mu \nabla^2 \boldsymbol{u}$$

$$\frac{\partial P'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} P' + w \frac{dP_0}{dz} = \mu \gamma \frac{P}{\theta} \nabla^2 \theta.$$

Before describing the semi-implicit time-integration for all five equation sets, let us say a few words about the spatial discretization method. Although we have chosen to use the spectral element method, the semi-implicit method for all five equation sets does not change for other discretization methods as long as the resulting mass matrix is diagonal as is the case for finite difference and spectral element methods. For the construction of semi-implicit methods for Godunov-type methods, such as finite volumes and discontinuous Galerkin methods, see Restelli and Giraldo (2009) [27] where the method is described only for equation set 3. In a forthcoming paper, we will perform a similar analysis of the semi-implicit method on various forms of the equation sets with the DG discretization; this analysis will then be applicable to all other Godunov-type methods. For further details on the spectral element discretization for the equations described herein see [14]. Let us now describe the semi-implicit formulation of the five equation sets.

3. Semi-Implicit Time-Integration. The governing equations can be written in the compact vector form

$$\frac{\partial \boldsymbol{q}}{\partial t} = S(\boldsymbol{q}) \tag{3.1}$$

where, e.g., for set 3 $\boldsymbol{q} = (\rho', \boldsymbol{U}^{\mathcal{T}}, E')^{\mathcal{T}}$ and the right-hand side $S(\boldsymbol{q})$ represents the remaining terms in the equations apart from the time derivatives. In order to obtain the semi-implicit time discretization of (3.1), we introduce a linear operator $L(\boldsymbol{q})$ which approximates $S(\boldsymbol{q})$ and contains the terms responsible for the acoustic and gravity waves (the precise form of which will be defined in Sect. 3.2), rewrite (3.1) as

$$\frac{\partial \boldsymbol{q}}{\partial t} = \{S(\boldsymbol{q}) - \delta L(\boldsymbol{q})\} + [\delta L(\boldsymbol{q})] \tag{3.2}$$

and discretize explicitly in time the terms in curly brackets and implicitly those in square brackets. The parameter δ is introduced in Eq.(3.2) to obtain a unified formalism for semi-implicit discretizations, for $\delta = 1$, and fully explicit ones, for $\delta = 0$.

As was done in [12, 18] we now consider a generic K step discretization of (3.2) of the form

$$\boldsymbol{q}^{n+1} = \sum_{k=0}^{K-1} \alpha_k \boldsymbol{q}^{n-k} + \chi \Delta t \sum_{k=0}^{K-1} \beta_k [S(\boldsymbol{q}^{n-k}) - \delta L(\boldsymbol{q}^{n-k})] + \chi \Delta t \delta L(\boldsymbol{q}^{n+1}), \tag{3.3}$$

where Δt is the time step, assumed to be constant for simplicity, and $\boldsymbol{q}^n$ denotes the solution at time level $n\Delta t$, for $n = 0, 1, \ldots$ To simplify the discussion of the semi-implicit formulation, let us now introduce the following variables

$$\boldsymbol{q}_{tt} = \boldsymbol{q}^{n+1} - \sum_{k=0}^{K-1} \beta_k \boldsymbol{q}^{n-k}, \quad \widehat{\boldsymbol{q}} = \boldsymbol{q}^E - \sum_{k=0}^{K-1} \beta_k \boldsymbol{q}^{n-k}, \quad \boldsymbol{q}^E = \sum_{k=0}^{K-1} \alpha_k \boldsymbol{q}^{n-k} + \chi \Delta t \sum_{k=0}^{K-1} \beta_k S(\boldsymbol{q}^{n-k})$$

that then allows us to write Eq. (3.2) as

$$\boldsymbol{q}_{tt} = \widehat{\boldsymbol{q}} + \lambda L(\boldsymbol{q}_{tt}) \tag{3.4}$$

where $\lambda = \chi \Delta t \delta$. For example, the coefficients for the BDF2 method are $\alpha_0 = 4/3$, $\alpha_1 = -1/3$, $\chi = 2/3$, $\beta_0 = 2$, and $\beta_1 = -1$ (see [15] for BDF methods of orders one through six).

The crux of the semi-implicit method, as is evident in Eq. (3.2), is the derivation of the linear operator L. The success of the SI method depends on this operator because it must be chosen such that the fastest waves in the system are retained, albeit in their linearized form. If the correct operator L is not obtained, the SI method will not work. Fortunately, deriving the linear operator is rather straightforward. We follow a similar approach used to split the variables into a hydrostatically-balanced reference state and the perturbation from this state; in other words, we define the variables as $\boldsymbol{q} = \boldsymbol{q}_0(z) + \boldsymbol{q}(\boldsymbol{x}, t)$.

3.1. Boundary Conditions. In this paper, we only consider two types of boundary conditions: no-flux (i.e., reflecting) and non-reflecting boundary conditions. For the no-flux boundary conditions, we apply the condition $\boldsymbol{n}_\Gamma \cdot \boldsymbol{u} = 0$ where $\boldsymbol{n}_\Gamma$ is the outward pointing normal vector of the boundary Γ. Since $\boldsymbol{u}$ and $\boldsymbol{n}_\Gamma$ both live in R^2 then we can define an augmented normal vector $\widehat{\boldsymbol{n}}_\Gamma = (0, \boldsymbol{n}_\Gamma^{\mathcal{T}}, 0)^{\mathcal{T}} \in R^4$ that then allows us to satisfy no-flux boundary conditions as follows: $\widehat{\boldsymbol{n}}_\Gamma \cdot \boldsymbol{q} = 0$. For explicit time-integration methods, one can apply all boundary conditions in an *a posteriori* fashion but this is not correct for an implicit method; for such methods, all boundary conditions need to be applied differently. We apply the boundary conditions through Lagrange multipliers as follows:

$$\frac{\partial \boldsymbol{q}}{\partial t} = S(\boldsymbol{q}) + \tau_{nf} \widehat{\boldsymbol{n}}_\Gamma + \tau_{nr}(\boldsymbol{q} - \boldsymbol{q}_b) \tag{3.5}$$

where τ_{nf} and τ_{nr} are the Lagrange multipliers for the *no-flux* and *non-reflecting* boundary conditions, respectively, and $\boldsymbol{q}_b$ is the free-stream (boundary) values of the state variable $\boldsymbol{q}$.

It turns out that to impose the non-reflecting boundary conditions given above in a strong sense, one can write the semi-discrete (in time) equations as follows

$$\boldsymbol{q}_{tt} = \alpha \left(\widehat{\boldsymbol{q}} + \lambda L(\boldsymbol{q}_{tt})\right) + \beta \widehat{\boldsymbol{q}}_b$$

where α and β are Newtonian relaxation coefficients that drive the solution towards the boundary reference value such that $\alpha \to 1$, $\beta \to 0$ in the interior and $\alpha \to 0$, $\beta \to 1$ as the non-reflecting boundaries are approached; this boundary condition is applied to the entire solution vector $\boldsymbol{q}$. Specifically, we define

$$\beta = \left(\frac{z - z_s}{z_t - z_s}\right)^4 \quad \text{and} \quad \alpha = 1 - \beta$$

with $z_s = 12km$, z_t is the top of the model, and $z \in [z_s, z_t]$, otherwise $\beta = 0$. A similar approach is used for the lateral boundaries where for the left boundary we define $x_s^{\text{left}} = 20km$ and $x_t^{\text{left}} = x_{min}$ and for the right $x_s^{\text{right}} = x_{max} - 20km$ and $x_t^{\text{right}} = x_{max}$; these boundary conditions are only used for the mountain test (case 4).

In contrast, for the no-flux boundaries, the boundary condition need only be applied to the velocity field $\boldsymbol{u}$. In this case, we rewrite the momentum equations as

$$\boldsymbol{U}_{tt} = \alpha \left(\widehat{\boldsymbol{U}} + \lambda L(\boldsymbol{q}_{tt})\right) + \beta \boldsymbol{U}_b + \tau_{nf} \boldsymbol{n}_\Gamma.$$

Taking the scalar product of this equation with $\boldsymbol{n}_\Gamma$ and rearranging results in the following equivalent system

$$\boldsymbol{U}_{tt} = \boldsymbol{P} \left[\alpha \left(\widehat{\boldsymbol{U}} + \lambda L(\boldsymbol{q}_{tt})\right) + \beta \boldsymbol{U}_b\right]$$

where $\boldsymbol{P}$ is the projection matrix

$$\boldsymbol{P} = \begin{pmatrix} 1 - n_x^2 & -n_x n_z \\ -n_x n_z & 1 - n_z^2 \end{pmatrix} \tag{3.6}$$

that imposes the no-flux boundary condition; note that we have dropped the subscript Γ from the normal vector $\boldsymbol{n}_\Gamma$ for convenience. It should be understood that $\boldsymbol{P}$ is only defined on Γ, in the interior domain, i.e., Ω - Γ, $\boldsymbol{P}$ simplifies to the identity matrix.

3.2. Definition of the Implicit Linear Problem. In this section, we address the precise definition, in the case of the various considered forms of the governing equations, of the linear operator L that has been introduced in Eq. (3.2) for the case of an abstract problem. In order to ensure stability, it is important that this operator includes the terms responsible for the fastest waves in the system, albeit in their linearized from. Once the operator L as been defined, the linear system to be solved at each time step is given by Eq. (3.4) in terms of the unknown $\boldsymbol{q}_{tt}$, from which the updated solution $\boldsymbol{q}^{n+1}$ can be readily obtained. For the 2D Euler equations, this requires the inversion of a $4N_p \times 4N_p$ matrix, where N_p denotes the total number of degrees of freedom for each scalar unknown in the problem. Such a system can be solved with a monolithic approach; however, a better strategy is often reformulating it into a smaller one with a technique known in the literature by many names, including block LU decomposition, collapsing the equations to a pseudo-Helmholtz operator form, or solving the *Schur* complement of the system. In the remainder of this section we construct the pseudo-Helmholtz operators for all the equation sets one at a time and shall refer to the full system as the *No Schur* form and the other as the *Schur* form. We will see that the Schur form invariably leads to an equation for a single pressure-like variable, requiring the inversion of an $N_p \times N_p$ matrix, considerably smaller compared to the matrix inverted in the monolithic approach. Since our discussion is independent from the chosen spatial discretization, we refer here to the time semi-discretized problem; the fully discrete problem is readily obtained by substitution of the continuous differential operator with the discrete ones. One final note is in order: for all of our simulations we use GMRES as our nonsymmetric iterative solver with Jacobi preconditioning (see [8] for a description). In future work, we will explore the effects of various preconditioners, e.g., overlapping Schwarz [7].

3.2.1. SE1. For the equation set SE1 we follow [24, 3, 36, 5, 35] and define the linear operator

$$L(\boldsymbol{q}) = - \begin{pmatrix} w\frac{d\pi_0}{dz} + (\gamma - 1)\pi_0 \boldsymbol{\nabla} \cdot \boldsymbol{u} \\ c_p\theta_0 \boldsymbol{\nabla}\pi' - g\frac{\theta'}{\theta_0}\boldsymbol{k} \\ w\frac{d\theta_0}{dz} \end{pmatrix}. \tag{3.7}$$

Note that in Eq.(3.7) we rely on the same reference state π_0, θ_0 introduced in Sect. 2. This is convenient, since it avoids introducing additional reference profiles, but it is not necessary, and in principle any known profile could be used in Eq. (3.7). Substituting Eq.(3.7) into Eq. (3.4), yields

$$\pi_{tt} = \alpha \left(\widehat{\pi} - \lambda w_{tt}\frac{d\pi_0}{dz} - \lambda(\gamma - 1)\pi_0 \boldsymbol{\nabla} \cdot \boldsymbol{u}_{tt} \right) + \beta\widehat{\pi}_b \tag{3.8}$$

$$\boldsymbol{u}_{tt} = \alpha \left(\widehat{\boldsymbol{u}} - \lambda c_p\theta_0 \boldsymbol{\nabla}\pi_{tt} + \lambda g\frac{\theta_{tt}}{\theta_0}\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{u}}_b \tag{3.9}$$

$$\theta_{tt} = \alpha \left(\widehat{\theta} - \lambda w_{tt}\frac{d\theta_0}{dz} \right) + \beta\widehat{\theta}_b \tag{3.10}$$

where $\widehat{\boldsymbol{q}}_b = \boldsymbol{q}_b - \sum_{k=0}^{K-1} \alpha_k \boldsymbol{q}^{n-k}$, with $\boldsymbol{q}_b$ being the reference values of the non-reflecting boundary conditions (NRBC). Equations (3.8)-(3.10) represent the full system (i.e., the *No Schur* form) of SE1 of dimension $4N_p \times 4N_p$. However, let us now construct the *Schur* form of this system.

We can now substitute Eq. (3.10) into Eq. (3.9) to get

$$\boldsymbol{u}_{tt} = \boldsymbol{C}_1 \left[\alpha \left(\widehat{\boldsymbol{u}} - \lambda c_p\theta_0 \boldsymbol{\nabla}\pi_{tt} + \lambda\frac{g}{\theta_0}\widehat{\theta}\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{u}}_b \right] \tag{3.11}$$

where

$$\boldsymbol{C}_1 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_1} \end{pmatrix} \tag{3.12}$$

with

$$c_1 = 1 + (\alpha\lambda)^2 \frac{g}{\theta_0}\frac{d\theta_0}{dz}. \tag{3.13}$$

Let us rewrite Eq. (3.11) as follows

$$\boldsymbol{u}_{tt} = \boldsymbol{C}_1 \left[\alpha \left(\widehat{\boldsymbol{u}} - \lambda c_p\theta_0 \boldsymbol{\nabla}\pi_{tt} + \lambda\frac{g}{\theta_0}\widehat{\theta}\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{u}}_b \right] + \tau_{nf}\boldsymbol{n} \tag{3.14}$$

To satisfy the no-flux boundary conditions, we simply replace $\boldsymbol{C}_1$ with $\boldsymbol{P}_1$ such that Eq. (3.14) with $\boldsymbol{n}$ and rearranging gives

$$\boldsymbol{u}_{tt} = \boldsymbol{P}_1 \left[\alpha \left(\widehat{\boldsymbol{u}} - \lambda c_p\theta_0 \boldsymbol{\nabla}\pi_{tt} + \lambda\frac{g}{\theta_0}\widehat{\theta}\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{u}}_b \right] \tag{3.15}$$

where $\boldsymbol{P}_1 = \boldsymbol{P}\boldsymbol{C}_1$ with $\boldsymbol{P}$ defined in Eq. (3.6). We can now substitute Eq. (3.15) into Eq. (3.8) to get

$$\pi_{tt} - (\alpha\lambda)^2\frac{d\pi_0}{dz}\boldsymbol{k} \cdot (\boldsymbol{P}_1 c_p\theta_0 \boldsymbol{\nabla}\pi_{tt}) - (\alpha\lambda)^2(\gamma - 1)\pi_0 \boldsymbol{\nabla} \cdot (\boldsymbol{P}_1 c_p\theta_0 \boldsymbol{\nabla}\pi_{tt}) = \alpha\widehat{\pi} + \beta\widehat{\pi}_b$$
$$- \alpha\lambda\frac{d\pi_0}{dz}\boldsymbol{k} \cdot \left[\boldsymbol{P}_1\alpha \left(\widehat{\boldsymbol{u}} + \lambda\frac{g}{\theta_0}\boldsymbol{k} \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right) \right) + \boldsymbol{P}_1\beta\widehat{\boldsymbol{u}}_b \right]$$
$$- \alpha\lambda(\gamma - 1)\pi_0 \boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_1\alpha \left(\widehat{\boldsymbol{u}} + \lambda\frac{g}{\theta_0}\boldsymbol{k} \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right) \right) + \boldsymbol{P}_1\beta\widehat{\boldsymbol{u}}_b \right] \tag{3.16}$$

which is the *Schur* form of SE1 and is of dimension $N_p \times N_p$. Note that this is a pseudo-Helmholtz equation for π_{tt} and can be solved quite readily by any nonsymmetric iterative solver. Note further that the solution of this linear problem satisfies both non-reflecting and no-flux boundary conditions. Upon getting a solution for π_{tt} from Eq. (3.16) we can then solve for $\boldsymbol{u}_{tt}$ using Eq. (3.15). To solve for θ_{tt} we next solve Eq. (3.10). Once $\boldsymbol{q}_{tt}$ is known, we then extract the solution $\boldsymbol{q}^{n+1}$ using Eq. (3.4).

3.2.2. SE2NC. The linear operator for SE2NC is

$$L(\boldsymbol{q}) = - \begin{pmatrix} w\frac{d\rho_0}{dz} + \rho_0 \boldsymbol{\nabla} \cdot \boldsymbol{u} \\ \frac{1}{\rho_0}\boldsymbol{\nabla}P' + g\frac{\rho'}{\rho_0}\boldsymbol{k} \\ w\frac{d\theta_0}{dz} \end{pmatrix}$$

with the pressure defined as

$$P' = \frac{\gamma P_0}{\rho_0}\rho' + \frac{\gamma P_0}{\theta_0}\theta'.$$

Applying the semi-implicit method to SE2NC yields

$$\rho_{tt} = \alpha \left(\widehat{\rho} - \lambda w_{tt}\frac{d\rho_0}{dz} - \lambda\rho_0 \boldsymbol{\nabla} \cdot \boldsymbol{u}_{tt} \right) + \beta\widehat{\rho}_b \tag{3.17}$$

$$\boldsymbol{u}_{tt} = \alpha \left(\widehat{\boldsymbol{u}} - \lambda\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} - \lambda g\frac{\rho_{tt}}{\rho_0}\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{u}}_b \tag{3.18}$$

$$\theta_{tt} = \alpha \left(\widehat{\theta} - \lambda w_{tt}\frac{d\theta_0}{dz} \right) + \beta\widehat{\theta}_b \tag{3.19}$$

$$P_{tt} = G_0\rho_{tt} + H_0\theta_{tt}. \tag{3.20}$$

where $G_0 = \frac{\gamma P_0}{\rho_0}$ and $H_0 = \frac{\gamma P_0}{\theta_0}$; the system represented by Eqs. (3.17)-(3.20) is the *No Schur* form of SE2NC. Substituting Eq. (3.19) into Eq. (3.20) yields

$$\rho_{tt} = \frac{1}{G_0} \left[P_{tt} - H_0\alpha \left(\widehat{\theta} - \lambda w_{tt}\frac{d\theta_0}{dz} \right) - H_0\beta\widehat{\theta}_b \right]. \tag{3.21}$$

We can now substitute Eq. (3.21) into Eq. (3.18) in order to express the momentum as a function of pressure only. Upon applying this substitution, we get

$$\boldsymbol{u}_{tt} = \boldsymbol{P}_{2NC} \left[(\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda\frac{gH_0}{\rho_0 G_0} \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right)\boldsymbol{k} - \alpha\lambda\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} - \alpha\lambda\frac{g}{\rho_0 G_0}P_{tt}\boldsymbol{k} \right] \tag{3.22}$$

where

$$\boldsymbol{C}_{2NC} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_{2NC}} \end{pmatrix} \tag{3.23}$$

with

$$c_{2NC} = 1 + (\alpha\lambda)^2 \frac{g}{\theta_0}\frac{d\theta_0}{dz}, \tag{3.24}$$

and $\boldsymbol{P}_{2NC} = \boldsymbol{P}\boldsymbol{C}_{2NC}$, where we have included the no-flux boundary conditions through the projection matrix $\boldsymbol{P}$.

Substituting Eqs. (3.17) and (3.19) into Eq. (3.20) yields

$$P_{tt} = G_0 \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right) + H_0 \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right) - \alpha\lambda F_0 w_{tt} - \alpha\lambda\rho_0 G_0 \boldsymbol{\nabla} \cdot \boldsymbol{u}_{tt} \tag{3.25}$$

where $F_0 = G_0 \frac{d\rho_0}{dz} + H_0 \frac{d\theta_0}{dz}$. The last step is to substitute Eq. (3.22) into Eq. (3.25) that yields

$$
\begin{aligned}
P_{tt} &- (\alpha\lambda)^2 F_0 \boldsymbol{k} \cdot \left[\boldsymbol{P}_{2NC} \left(\frac{1}{\rho_0} \boldsymbol{\nabla} P_{tt} + \frac{g}{\rho_0 G_0} P_{tt} \boldsymbol{k} \right) \right] \\
&- (\alpha\lambda)^2 G_0 \rho_0 \boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_{2NC} \left(\frac{1}{\rho_0} \boldsymbol{\nabla} P_{tt} + \frac{g}{\rho_0 G_0} P_{tt} \boldsymbol{k} \right) \right] \\
&= G_0 \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right) + H_0 \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right) \\
&- \alpha\lambda F_0 \boldsymbol{k} \cdot \left[\boldsymbol{P}_{2NC} \left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda \frac{gH_0}{\rho_0 G_0} \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right) \boldsymbol{k} \right) \right] \\
&- \alpha\lambda G_0 \rho_0 \boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_{2NC} \left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda \frac{gH_0}{\rho_0 G_0} \left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b \right) \boldsymbol{k} \right) \right].
\end{aligned}
\tag{3.26}
$$

and is the *Schur* form of SE2NC and is of dimension $N_p \times N_p$.

3.2.3. SE2C. The linear operator for SE2C is

$$
L(\boldsymbol{q}) = - \begin{pmatrix} \boldsymbol{\nabla} \cdot \boldsymbol{U} \\ \boldsymbol{\nabla} P' + g\rho' \boldsymbol{k} \\ \boldsymbol{\nabla} \cdot \left(\frac{\Theta_0}{\rho_0} \boldsymbol{U} \right) \end{pmatrix}
$$

with the pressure linearized as follows

$$
P' = \frac{\gamma P_0}{\Theta_0} \Theta'.
$$

Upon applying the semi-implicit method to SE2C and letting $F_0 = \frac{\gamma P_0}{\Theta_0}$ and $G_0 = \frac{\Theta_0}{\rho_0}$ we get

$$
\rho_{tt} = \alpha \left(\widehat{\rho} - \lambda \boldsymbol{\nabla} \cdot \boldsymbol{U}_{tt} \right) + \beta\widehat{\rho}_b
\tag{3.27}
$$

$$
\boldsymbol{U}_{tt} = \alpha \left(\widehat{\boldsymbol{U}} - \lambda \boldsymbol{\nabla} P_{tt} - \lambda g \rho_{tt} \boldsymbol{k} \right) + \beta\widehat{\boldsymbol{U}}_b
\tag{3.28}
$$

$$
\Theta_{tt} = \alpha \left(\widehat{\Theta} - \lambda \boldsymbol{\nabla} \cdot (G_0 \boldsymbol{U}_{tt}) \right) + \beta\widehat{\Theta}_b
\tag{3.29}
$$

$$
P_{tt} = F_0 \Theta_{tt}.
\tag{3.30}
$$

Equations (3.27)-(3.30) represent the full system of SE2C (i.e., the *No Schur* form). Let us now derive the Schur form.

Let us first substitute Eq. (3.29) into Eq. (3.30) to get

$$
P_{tt} = F_0 \alpha \left(\widehat{\Theta} - \lambda \boldsymbol{\nabla} \cdot (G_0 \boldsymbol{U}_{tt}) \right) + F_0 \beta\widehat{\Theta}_b.
\tag{3.31}
$$

Multiplying Eq. (3.27) by G_0 and subtracting from Eq. (3.29) to eliminate the term $G_0 \boldsymbol{\nabla} \cdot \boldsymbol{U}_{tt}$ yields

$$
\Theta_{tt} - G_0 \rho_{tt} = \left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b \right) - G_0 \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right) - \alpha\lambda W_{tt} \frac{dG_0}{dz}.
\tag{3.32}
$$

Substituting Eq. (3.30) into Eq. (3.32), to eliminate Θ_{tt}, gives

$$
\rho_{tt} = \frac{1}{F_0 G_0} P_{tt} + \alpha\lambda W_{tt} \frac{1}{G_0} \frac{dG_0}{dz} - \frac{1}{G_0} \left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b \right) + \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right).
\tag{3.33}
$$

Note that plugging Eq. (3.33) into Eq. (3.28) allows us to solve for $\boldsymbol{U}_{tt}$ as a function of P_{tt} such as

$$
\boldsymbol{U}_{tt} = \boldsymbol{P}_{2C} \left[\left(\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b \right) - \alpha\lambda \boldsymbol{\nabla} P_{tt} - \alpha\lambda \frac{g}{F_0 G_0} P_{tt} \boldsymbol{k} - \alpha\lambda g \boldsymbol{k} \left(\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right) - \frac{1}{G_0} \left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b \right) \right) \right]
\tag{3.34}
$$

where $\boldsymbol{P}_{2C} = \boldsymbol{P}\boldsymbol{C}_{2C}$ with

$$C_{2C} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_{2C}} \end{pmatrix} \qquad (3.35)$$

and $c_2 = 1 + (\alpha\lambda)^2 \frac{g}{G_0} \frac{dG_0}{dz}$ where $\frac{1}{G_0}\frac{dG_0}{dz} = \left(\frac{1}{\Theta_0}\frac{d\Theta_0}{dz} - \frac{1}{\rho_0}\frac{d\rho_0}{dz} \right)$. Finally, substituting Eq. (3.34) into Eq. (3.31) yields

$$P_{tt} - (\alpha\lambda)^2 F_0 \boldsymbol{\nabla} \cdot \left[G_0 \boldsymbol{P}_{2C} \left(\boldsymbol{\nabla} P_{tt} + \frac{g}{F_0 G_0} P_{tt} \boldsymbol{k} \right) \right] = F_0 \left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b \right)$$
$$- \alpha\lambda F_0 \boldsymbol{\nabla} \cdot \left[G_0 \boldsymbol{P}_{2C} \left(\left(\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b \right) - \alpha\lambda g \boldsymbol{k} \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right) + \alpha\lambda\frac{g}{G_0}\boldsymbol{k} \left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b \right) \right) \right]$$

which is a pseudo-Helmholtz equation for P_{tt} and is the *Schur* form of SE2C.

3.2.4. SE3. The linear operator for SE3 is

$$L(\boldsymbol{q}) = - \begin{pmatrix} \boldsymbol{\nabla} \cdot \boldsymbol{U} \\ \boldsymbol{\nabla} P' + \rho' g \boldsymbol{k} \\ \boldsymbol{\nabla} \cdot (h_0 \boldsymbol{U}) \end{pmatrix}$$

with the pressure defined as

$$P' = (\gamma - 1)(E' - \rho'\phi)$$

and $h_0 = \frac{E_0 + P_0}{\rho_0}$ is the reference enthalpy where E_0, P_0, and ρ_0 are the hydrostatically-balanced reference total energy, pressure, and density. Upon applying the semi-implicit method to SE3 we arrive at the following semi-discrete problem

$$\rho_{tt} = \alpha \left(\widehat{\rho} - \lambda\boldsymbol{\nabla} \cdot \boldsymbol{U}_{tt} \right) + \beta\widehat{\rho}_b \qquad (3.36)$$

$$\boldsymbol{U}_{tt} = \alpha \left(\widehat{\boldsymbol{U}} - \lambda\boldsymbol{\nabla} P_{tt} - \lambda\rho_{tt}g\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{U}}_b \qquad (3.37)$$

$$E_{tt} = \alpha \left(\widehat{E} - \lambda\boldsymbol{\nabla} \cdot (h_0 \boldsymbol{U}_{tt}) \right) + \beta\widehat{E}_b \qquad (3.38)$$

$$P_{tt} = (\gamma - 1)(E_{tt} - \phi\rho_{tt}). \qquad (3.39)$$

The system represented by Eqs. (3.36)-(3.39) is the *No Schur* form of SE3. Let us now derive the Schur form.

Substituting Eqs. (3.36) and (3.38) into Eq. (3.39) yields

$$P_{tt} = (\gamma - 1)\left[(\alpha\widehat{E} + \beta\widehat{E}_b) - \alpha\lambda h_0 \boldsymbol{\nabla} \cdot \boldsymbol{U}_{tt} - \alpha\lambda\boldsymbol{\nabla} h_0 \cdot \boldsymbol{U}_{tt} \right] - \phi(\gamma - 1)\left[(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) - \alpha\lambda\boldsymbol{\nabla} \cdot \boldsymbol{U}_{tt} \right].$$
$$(3.40)$$

Multiplying Eq. (3.36) by h_0 and subtracting from Eq. (3.38) to eliminate the term $h_0 \boldsymbol{\nabla} \cdot \boldsymbol{U}_{tt}$ yields

$$E_{tt} - h_0\rho_{tt} = (\alpha\widehat{E} + \beta\widehat{E}_b) - h_0(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) - \alpha\lambda W_{tt}\frac{dh_0}{dz}. \qquad (3.41)$$

Next, substituting Eq. (3.41) into Eq. (3.39), to eliminate E_{tt}, and rearranging gives

$$\rho_{tt} = \frac{1}{h_0 - \phi} \left[\frac{1}{(\gamma - 1)}P_{tt} + \alpha\lambda W_{tt}\frac{dh_0}{dz} - (\alpha\widehat{E} + \beta\widehat{E}_b) + h_0(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right] \qquad (3.42)$$

which can now be substituted into Eq. (3.37) and solved for $\boldsymbol{U}_{tt}$ to yield

$$\boldsymbol{U}_{tt} = \boldsymbol{P}_3 \left[(\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b) - \alpha\lambda\boldsymbol{\nabla} P_{tt} - \alpha\lambda\frac{g}{(\gamma - 1)(h_0 - \phi)}P_{tt}\boldsymbol{k} - \alpha\lambda\frac{g}{h_0 - \phi}\left(h_0(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) - (\alpha\widehat{E} + \beta\widehat{E}_b) \right)\boldsymbol{k} \right]$$
$$(3.43)$$

where

$$C_3 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_3} \end{pmatrix} \tag{3.44}$$

with

$$c_3 = 1 + (\alpha\lambda)^2 \frac{g}{h_0 - \phi} \frac{dh_0}{dz}. \tag{3.45}$$

Finally, substituting Eq. (3.43) into Eq. (3.40) yields

$$P_{tt} - (\alpha\lambda)^2(\gamma - 1)(h_0 - \phi)\boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_3 \left(\boldsymbol{\nabla} P_{tt} + \frac{g}{(\gamma - 1)(h_0 - \phi)} P_{tt}\boldsymbol{k} \right) \right]$$

$$- (\alpha\lambda)^2(\gamma - 1)\boldsymbol{\nabla} h_0 \cdot \left[\boldsymbol{P}_3 \left(\boldsymbol{\nabla} P_{tt} + \frac{g}{(\gamma - 1)(h_0 - \phi)} P_{tt}\boldsymbol{k} \right) \right]$$

$$= (\gamma - 1) \left[(\alpha\widehat{E} + \beta\widehat{E}_b) - \phi(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right]$$

$$- \alpha\lambda(\gamma - 1)(h_0 - \phi)\boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_3 \left((\alpha\widehat{U} + \beta\widehat{U}_b) - \alpha\lambda\frac{gh_0}{h_0 - \phi}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b)\boldsymbol{k} + \alpha\lambda\frac{g}{h_0 - \phi}(\alpha\widehat{E} + \beta\widehat{E}_b)\boldsymbol{k} \right) \right]$$

$$- \alpha\lambda(\gamma - 1)\boldsymbol{\nabla} h_0 \cdot \left[\boldsymbol{P}_3 \left((\alpha\widehat{U} + \beta\widehat{U}_b) - \alpha\lambda\frac{gh_0}{h_0 - \phi}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b)\boldsymbol{k} + \alpha\lambda\frac{g}{h_0 - \phi}(\alpha\widehat{E} + \beta\widehat{E}_b)\boldsymbol{k} \right) \right]$$

which is a pseudo-Helmholtz equation for P_{tt} and is the *Schur* form of SE3.

3.2.5. SE4. The linear operator for SE4 is

$$L(\boldsymbol{q}) = - \begin{pmatrix} w\frac{d\rho_0}{dz} + \rho_0\boldsymbol{\nabla} \cdot \boldsymbol{u} \\ \\ \frac{1}{\rho_0}\boldsymbol{\nabla} P' + g\frac{\rho'}{\rho_0}\boldsymbol{k} \\ \\ w\frac{dP_0}{dz} + \gamma P_0\boldsymbol{\nabla} \cdot \boldsymbol{u} \end{pmatrix}.$$

Upon applying the semi-implicit method to SE4 we arrive at the following semi-discrete problem

$$\rho_{tt} = \alpha \left(\widehat{\rho} - \lambda w_{tt}\frac{d\rho_0}{dz} - \lambda\rho_0\boldsymbol{\nabla} \cdot \boldsymbol{u}_{tt} \right) + \beta\widehat{\rho}_b \tag{3.46}$$

$$\boldsymbol{u}_{tt} = \alpha \left(\widehat{\boldsymbol{u}} - \lambda\frac{1}{\rho_0}\boldsymbol{\nabla} P_{tt} - \lambda g\frac{\rho_{tt}}{\rho_0}\boldsymbol{k} \right) + \beta\widehat{\boldsymbol{u}}_b \tag{3.47}$$

$$P_{tt} = \alpha \left(\widehat{P} - \lambda w_{tt}\frac{dP_0}{dz} - \lambda\gamma P_0\boldsymbol{\nabla} \cdot \boldsymbol{u}_{tt} \right) + \beta\widehat{P}_b. \tag{3.48}$$

The system described by Eqs. (3.46)-(3.48) is the *No Schur* form of SE4. Let us now derive the Schur form.

Multiplying Eq. (3.46) by γP_0 and subtracting Eq. (3.48) multiplied by ρ_0 and rearranging yields

$$\rho_{tt} = \frac{1}{\gamma P_0} \left[\rho_0 P_{tt} - \rho_0 \left(\alpha\widehat{P} + \beta\widehat{P}_b \right) + \gamma P_0 \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b \right) + \alpha\lambda w_{tt} \left(\rho_0\frac{dP_0}{dz} - \gamma P_0\frac{d\rho_0}{dz} \right) \right]. \tag{3.49}$$

Substituting Eq. (3.49) into Eq. (3.47) yields

$$\boldsymbol{u}_{tt} = \boldsymbol{P}_4 \left[(\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda g\boldsymbol{k} \left(\frac{1}{\gamma P_0}(\alpha\widehat{P} + \beta\widehat{P}_b) - \frac{1}{\rho_0}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right) - \alpha\lambda\frac{1}{\rho_0}\boldsymbol{\nabla} P_{tt} - \alpha\lambda\frac{g}{\gamma P_0}P_{tt}\boldsymbol{k} \right] \tag{3.50}$$

where $\boldsymbol{P}_4 = \boldsymbol{P}\boldsymbol{C}_4$ and

$$C_4 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_4} \end{pmatrix} \tag{3.51}$$

with

$$c_4 = 1 + (\alpha\lambda)^2 g \left(\frac{1}{\gamma P_0} \frac{dP_0}{dz} - \frac{1}{\rho_0} \frac{d\rho_0}{dz} \right) \tag{3.52}$$

Plugging Eq. (3.50) into Eq. (3.48) yields

$$P_{tt} - (\alpha\lambda)^2 \frac{dP_0}{dz} \boldsymbol{k} \cdot \left[\boldsymbol{P}_4 \left(\frac{1}{\rho_0} \boldsymbol{\nabla} P_{tt} + \frac{g}{\gamma P_0} P_{tt} \boldsymbol{k} \right) \right] - (\alpha\lambda)^2 \gamma P_0 \boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_4 \left(\frac{1}{\rho_0} \boldsymbol{\nabla} P_{tt} + \frac{g}{\gamma P_0} P_{tt} \boldsymbol{k} \right) \right]$$

$$= (\alpha\widehat{P} + \beta\widehat{P}_b) - \alpha\lambda \frac{dP_0}{dz} \boldsymbol{k} \cdot \left[\boldsymbol{P}_4 \left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda g \boldsymbol{k} \left(\frac{1}{\gamma P_0}(\alpha\widehat{P} + \beta\widehat{P}_b) - \frac{1}{\rho_0}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right) \right) \right]$$

$$- \alpha\lambda\gamma P_0 \boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_4 \left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda g \boldsymbol{k} \left(\frac{1}{\gamma P_0}(\alpha\widehat{P} + \beta\widehat{P}_b) - \frac{1}{\rho_0}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right) \right) \right]$$

which is a pseudo-Helmholtz equation for P_{tt} and is the *Schur* form of SE4.

4. Results. In this section we validate the five semi-implicit models on a test case suite of four problems using, for the spatial discretization, the spectral element method. For the definitions of the test cases as well as for the details of the spatial discretization we refer the reader to [14].

For our comparisons, we identify the following three criteria: discrete conservation properties, accuracy, and efficiency. Since each of these criteria can be expressed in more than one metric, we need first of all to clarify what we mean by each of them. With the term *discrete conservation* properties we mean the ability of the numerical method to reproduce the integral balance equations of the continuous problem, which in the case of an isolated system reduce to conservation of flow integrals and in the case of a system with mass or energy exchange with the environment takes the form of a balance between boundary fluxes and variation of the system mass or energy. In analyzing our results, we have to distinguish two classes of numerical models: those for which discrete conservation properties can be shown by construction, and those for which this is not possible. In the first case, the experimental datum concerning conservation serves as a confirmation that the expected balance is satisfied up to machine precision; in the second case, it provides a fundamental error indicator because it is a quantitative measure of the deviation of the numerical solution from the analytic one. In practice, we will thus provide the mass and energy losses as follows. We define mass loss as

$$\text{Mass Loss} = \frac{\mathcal{M}(t) - \mathcal{M}(0)}{\mathcal{M}(0)} \quad \text{where} \quad \mathcal{M}(t) = \int_\Omega \rho(\boldsymbol{x}, t) d\Omega.$$

Similarly, we define the energy loss as

$$\text{Energy Loss} = \frac{\mathcal{E}(t) - \mathcal{E}(0)}{\mathcal{E}(0)} \quad \text{where} \quad \mathcal{E}(t) = \int_\Omega E(\boldsymbol{x}, t) d\Omega,$$

E is the density total energy of the system.

Concerning accuracy, we should mention that a significant difficulty in testing mesoscale models is the lack of nontrivial analytic solutions, so that we assess the accuracy of our results by comparing the results of various equation sets, both qualitatively and quantitatively against each other and with reference solutions published in the literature.

Providing a reliable assessment of the efficiency of our five implementations of the Navier-Stokes equations is not obvious, since it can be implementation and problem dependent. To solve this difficulty we compare the effort required by the solution of the semi-implicit system and the wallclock time of our experiments using comparable stopping criteria for the iterative solvers for all our codes and by making sure that the five Fortran 90 implementations are as similar as possible. We use wallclock time in seconds where all simulations are performed on an Apple Xserve with a clock speed of 2.8GHz on Intel Xeon Processors. In addition, we use the Courant number as a measure of the size of the time-steps that can be achieved with the semi-implicit method. We define the Courant number as

$$\text{Courant number} = \max\left(\frac{C\Delta t}{\Delta s} \right)$$

where $C = |U + \sqrt{a}|$ is the characteristic speed, $U = \boldsymbol{n} \cdot \boldsymbol{u}$ is the velocity in the direction $\boldsymbol{n}$, a is the sound speed, $\Delta s = \sqrt{\Delta x^2 + \Delta z^2}$ is the grid spacing.

A separate note is finally required for the last of our test cases, namely the hydrostatic mountain flow. This case differs from the others because of the presence of NRBCs, the availability of a semi-analytic solution and the use of additional diagnostics. The presence of NRBCs, in particular, poses some problems in determining the conservation properties and accuracy of the numerical solution. Since the problem is posed on an open domain, we should expect conservation of mass and energy in the form of an integral flux balance; the NRBCs, however, mathematically represent a source/sink term artificially introduced into the computational domain, that inevitably destroys the integral balance. The solution that we choose here is to restrict the integral flux balance to the inner domain where the sponge term vanishes.

Clearly, this notion of conservation can be satisfactory for processes that are entirely contained in the inner domain, but can be unsatisfactory for processes significantly affected by flows through the boundaries of the domain. In particular, the sponge layer makes it impossible to build a model that is conservative with respect to fluxes prescribed on the boundary. To overcome this problem, we have begun work on the construction of high-order NRBCs that can be used with high-order spatial and temporal discretizations (see [6] and [26]) but we are still far away from implementing such methods into Navier-Stokes models. Unfortunately, sponge-based NRBCs are those typically used today in industrial-type nonhydrostatic atmospheric models.

The fact that the analytic solution is known for the problem defined in an infinite domain while the numerical method solves the problem in a limited domain with the addition of the sponge terms (or, in other words, the fact that the NRBCs are not exact in modeling the infinite domain and not even high-order) prevents the model from converging to the analytic solution with the theoretical order of accuracy. In fact, as is shown in Sec. 4.1.4, all our simulations converge to a solution that is close but distinct from the analytic one which we can interpret as the solution of the modified problem "Navier-Stokes equations with NRBCs" defined by Eq. (3.5). In order to quantify this deviation and to compare with other results in the literature, we define the root-mean-square error as

$$\|\boldsymbol{q}\|_{RMS} = \sqrt{\sum_{i=1}^{N_p} \left(\boldsymbol{q}^{\text{numerical}} - \boldsymbol{q}^{\text{analytic}}\right)^2 / N_p}$$

where $N_p = 40,000$. The semi-analytic solutions are computed via Matlab routines that are available upon request. Finally, in addition to the diagnostics used for the other test cases, we will also consider the momentum flux as [32]

$$m(z) = \int_{-\infty}^{+\infty} \rho_0(z) u(x,z) w(x,z) dx$$

where $\rho_0(z)$ is the reference density as a function of height. From linear theory, the analytic hydro-static momentum flux is given as [32]

$$m^H(z) = -\frac{\pi_c}{4} \bar{\rho}_s \bar{u}_s \mathcal{N} h_c^2$$

where the superscript H signifies *hydrostatic*, $\bar{\rho}_s$ and $\bar{u}_s$ are the reference density and horizontal velocity values at the surface, $\mathcal{N}$ is the Brunt-Väisälä frequency, and h_c is the height of the mountain. We shall use the normalized momentum flux, $m(z)/m^H(z)$, as a metric to test for convergence to steady-state.

4.1. Comparison of All Five Semi-Implicit Models. In this section we summarize the results of the five semi-implicit models using the Schur form for each of the four test cases. We begin with the inertia-gravity waves followed by the rising thermal bubble, the density current, and, finally, the linear hydrostatic mountain wave. Although each of the five semi-implicit models are derived from different equations (using different prognostic variables) it should be mentioned that the Schur form of all five models are very similar since they all reduce to a scalar second-order equation

for the pressure. In fact, the eigenvalue spectrum (spectral radius) and condition number (with respect to the 2-norm) of the linear matrix arising from the semi-implicit implementation are, for all intents and purposes, identical for all five equation sets. We mention this here only to emphasize that the difference in the number of GMRES iterations per time-step for each of the models is not a function of the condition number nor the spectral radius but of some other mechanism that we try to identify below.

One final note about the results below: while we show results for very specific resolutions (in this case, the flow is well-resolved) we have also analyzed under-resolved simulations and the comparisons that we now report are representative of differences of the five models in both the well and under-resolved regimes; here we refer mostly to the conservation measures which do not change with varying resolutions.

a)

b)
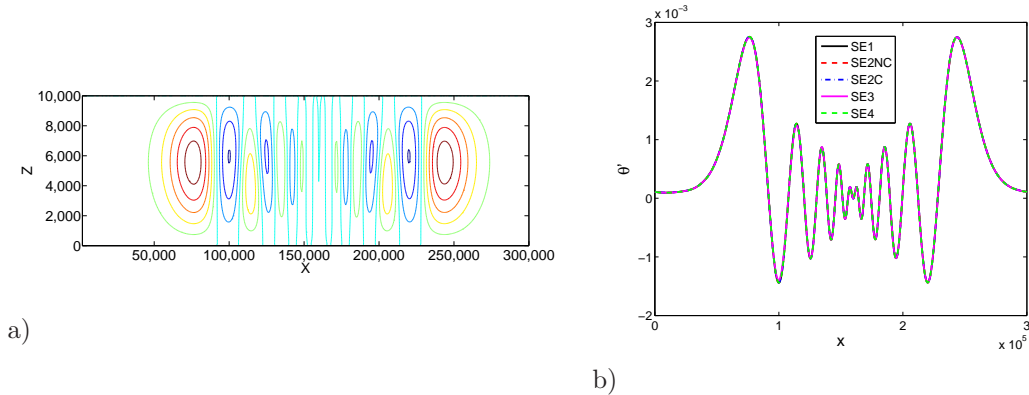
FIGURE 4.1. *Case 1: Inertia-Gravity Waves. Potential temperature perturbation after 3000 seconds for 250 meter resolution and 10th order polynomials. Figure a) shows the total domain using contour values between -0.0015 and 0.003 with a contour interval of 0.0005 and Figure b) shows the profiles along 5,000 meter height for all five models.*

4.1.1. Case 1: Inertia-Gravity Waves. Figure 4.1a shows the potential temperature perturbation contours after 3000 seconds and Fig. 4.1b shows the one-dimensional profile along z=5,000 meters for all five models. Figure 4.1b shows that all five models yield identical solutions; this is especially of interest since the models use different equation sets. The second result worth noting is that the profiles are perfectly symmetric about the position x=160,000 meters. Note that there is a mean horizontal flow in this problem, which tests the ability of the algorithm to preserve the proper phase speeds.

	SE1	SE2NC	SE2C	SE3	SE4
π'_{max}	1.06×10^{-6}	1.06×10^{-6}	1.10×10^{-6}	1.05×10^{-6}	1.04×10^{-6}
π'_{min}	-8.25×10^{-7}	-8.25×10^{-7}	-8.56×10^{-7}	-8.27×10^{-7}	-8.16×10^{-7}
u_{max}	1.07×10^{-2}	1.07×10^{-2}	1.07×10^{-2}	1.06×10^{-2}	1.07×10^{-2}
u_{min}	-1.06×10^{-2}	-1.06×10^{-2}	-1.06×10^{-2}	-1.06×10^{-2}	-1.06×10^{-2}
w_{max}	2.85×10^{-3}	2.85×10^{-3}	2.84×10^{-3}	2.84×10^{-3}	2.85×10^{-3}
w_{min}	-2.42×10^{-3}	-2.42×10^{-3}	-2.42×10^{-3}	-2.42×10^{-3}	-2.42×10^{-3}
θ'_{max}	2.80×10^{-3}	2.80×10^{-3}	2.80×10^{-3}	2.80×10^{-3}	2.80×10^{-3}
θ'_{min}	-1.51×10^{-3}	-1.51×10^{-3}	-1.51×10^{-3}	-1.51×10^{-3}	-1.51×10^{-3}
Mass Loss	3.09×10^{-11}	9.38×10^{-13}	1.85×10^{-12}	1.45×10^{-12}	2.11×10^{-12}
Energy Loss	2.45×10^{-8}	9.53×10^{-14}	6.12×10^{-14}	1.71×10^{-13}	8.89×10^{-7}
GMRES Iterations	5	5	5	5	5
WallClock Time	397	480	514	500	454

TABLE I

Case 1: Inertia-Gravity Wave. Comparison of the five models studied for 250 meter resolution and 10th order polynomials after 3000 seconds using $\Delta t = 1$ second (Courant Number = 3.15).

Skamarock and Klemp [31] give an analytic solution for this test but, unfortunately, it is only valid for the linearized problem, that while useful for qualitative comparisons, cannot be used to compute error norms since we use the fully nonlinear equations. We use the same contouring interval used in [31] and our results match very well. Specifically, their values are in the range $2.82 \times 10^{-3} \leq \theta' \leq -1.49 \times 10^{-3}$ whereas ours are $2.80 \times 10^{-3} \leq \theta' \leq -1.51 \times 10^{-3}$, which we show in Table I. In addition, comparing our semi-implicit results to the results in [14] for the explicit version of our models we find that they match almost exactly, in spite of the fact that here we now use much larger time-steps.

Table I shows that the five models give exactly the same results; the only outlier is SE2C that gives slightly different results. Note, however, that these differences are in the eighth decimal place. The main differences of interest are in the mass and energy conservation measures and in the efficiency (i.e., wallclock time) of the models. In terms of mass conservation, all models perform quite well except for SE1; this equation set is not expected to formally conserve mass. In terms of energy conservation, sets SE2NC, SE2C, and SE3 perform very well; sets SE1 and SE4 do not perform very well. It is not surprising that SE3 and SE2C achieve good energy conservation measures since they are in complete conservation form; however, SE2NC performs surprising well given the fact that it is not in strict conservation form. On the other hand, sets SE1 and SE4 are not expected to conserve energy at all and they exhibit this weakness quite strongly here. In terms of efficiency from best to worst the order is: SE1, SE4, SE2NC, SE3, and SE2C; the average number of GMRES iterations per step is the same for all five models thus the efficiency differences are due to differences in number of operations required by the equations themselves. SE1 and SE4 do not have an equation of state and therefore require fewer operations per time-step. The fully conservative models SE2C and SE3 have a larger operation count than the other models. This is the case because for the conservation forms, taking the divergence of the flux tensor requires more operations than merely taking the derivatives of the non-conservation form.
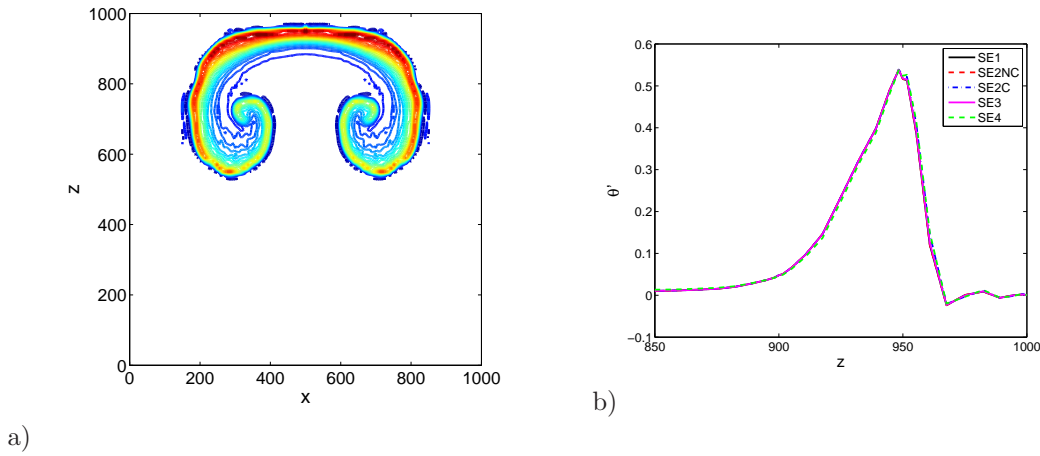

a)

b)

FIGURE 4.2. *Case 2: Rising Thermal Bubble. Potential temperature perturbation after 700 seconds for 5 meter resolution and 10th order polynomials. Figure a) shows the total domain using contour values between 0 to 0.525 with an interval of 0.025. Figure b) shows the profiles along x=500 meters for all five models.*

4.1.2. Case 2: Rising Thermal Bubble. Figure 4.2a shows the potential temperature perturbation contours after 700 seconds and Fig. 4.2b shows the one-dimensional profile along x=500 meters for all five models. This case has no analytic solution but the resulting dynamics are sufficiently simple to be able to predict its proper evolution. The initial condition consists of a warm bubble perturbation. Because of the effects of gravity, it begins to rise and shears along the way up. Since the maximum initial value of the bubble is $\theta' = 0.5$ then one expects to see a value of 0.5 as the maximum perturbation. Figure 4.2b shows that the maximum peak of the bubble is near 0.5, and on the leeward side ($950 < z < 1000$), the values fall slightly below 0 (see Table II).

Figure 4.2b shows a profile of the thermal waves along $x = 500$ for z values from 850 to 1000.

This figure shows that all five simulations yield nearly identical results and Fig. 4.2a shows that they are all symmetric about x=500 meters.

	SE1	SE2NC	SE2C	SE3	SE4
π'_{max}	5.67×10^{-6}	5.64×10^{-6}	5.67×10^{-6}	5.65×10^{-6}	5.80×10^{-6}
π'_{min}	-1.32×10^{-5}	-1.32×10^{-5}	-1.32×10^{-5}	-1.32×10^{-5}	-1.32×10^{-5}
u_{max}	2.01	2.01	2.01	2.01	2.02
u_{min}	-2.01	-2.01	-2.01	-2.01	-2.02
w_{max}	2.55	2.55	2.55	2.55	2.56
w_{min}	-1.95	-1.95	-1.95	-1.95	-1.96
θ'_{max}	0.54	0.54	0.54	0.54	0.54
θ'_{min}	-0.09	-0.09	-0.09	-0.09	-0.08
Mass Loss	3.51×10^{-9}	8.50×10^{-13}	1.20×10^{-13}	9.90×10^{-13}	3.80×10^{-12}
Energy Loss	4.10×10^{-6}	1.20×10^{-9}	3.50×10^{-9}	2.60×10^{-11}	1.10×10^{-5}
GMRES Iterations	33	34	33	34	33
WallClock Time	3454	3597	3590	3738	3511

TABLE II

Case 2: Rising Thermal Bubble. Comparison of the five models studied for 5 meter resolution and 10th order polynomials after 700 seconds using $\Delta t = 0.125$ seconds (Courant Number = 18.61).

Table II shows the maximum and minimum values for the four variables for all five models. The models give virtually identical results for all the variables; the outlier is SE4 which gives slightly different values, most notably in the second decimal place of the velocity fields. In terms of mass conservation, SE1 is by far the worst. However, the fact that SE1 does not conserve mass is expected since this equation set cannot formally conserve either mass or energy. In terms of energy conservation, SE1 and SE4 are the worst, the other three models do very well in both conservation of mass and energy, especially SE3. In terms of efficiency (wallclock time), from best to worst are: SE1, SE4, SE2C, SE2NC, and SE3. Since the conservation forms require a larger operation count, we would expect SE2NC to be faster than SE2C but we see that for this test case SE2NC averaged 34 GMRES iterations per time-step while SE2C averaged 33; this is the reason why SE2C is faster than SE2NC. The difference in wallclock time between SE2NC and SE2C is minimal but the difference between SE1 and SE3 is almost 300 seconds!
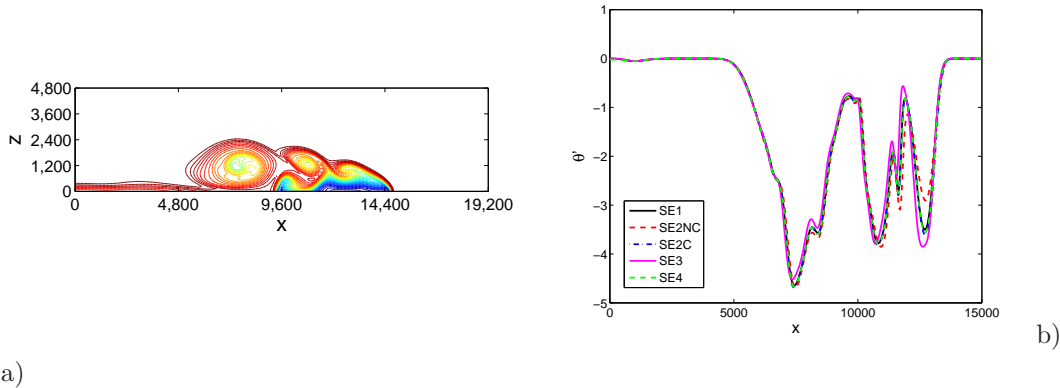

a)

b)

FIGURE 4.3. *Case 3: Density Current. Potential temperature perturbations after 900 seconds with 25 meter resolution and 8th order polynomials. Figure a) shows the total domain using contour values between -9 to 0 with a contour interval of 0.25. Figure b) shows the profiles along z=1200 meters for all five models.*

4.1.3. Case 3: Density Current. In Fig. 4.3a we plot the contours of potential temperature perturbation and in Fig. 4.3b the one-dimensional profile of the potential temperature perturbation along $z = 1200$ meters for all five models. The three negative wells in Fig. 4.3b correspond to the

three distinct Kelvin-Helmholtz instability waves (or rotors) clearly visible in Fig. 4.3a. It is not very clear from Figure 4.3b that there are differences between the five models. Instead, this figure shows that the models agree very well. In order to discern the differences between the five models, let us now review Table III.

	SE1	SE2NC	SE2C	SE3	SE4
π'_{max}	7.74×10^{-4}	7.78×10^{-4}	7.78×10^{-4}	7.74×10^{-4}	7.76×10^{-4}
π'_{min}	-1.67×10^{-3}	-1.61×10^{-3}	-1.66×10^{-3}	-1.69×10^{-3}	-1.66×10^{-3}
u_{max}	33.50	33.39	33.36	33.32	33.45
u_{min}	-14.76	-14.47	-14.78	-15.04	-14.79
w_{max}	12.15	12.04	12.25	12.40	12.16
w_{min}	-15.56	-15.32	-15.61	-15.85	-15.57
θ'_{max}	1.47×10^{-4}	1.67×10^{-5}	9.50×10^{-5}	3.30×10^{-3}	7.33×10^{-3}
θ'_{min}	-8.84	-8.70	-8.90	-9.09	-8.89
Mass Loss	2.33×10^{-5}	4.58×10^{-12}	1.95×10^{-13}	3.13×10^{-13}	5.41×10^{-12}
Energy Loss	2.34×10^{-4}	1.05×10^{-6}	1.92×10^{-5}	5.49×10^{-12}	7.64×10^{-4}
GMRES Iterations	3	3	3	4	3
WallClock Time	11073	12252	12665	12757	11991

TABLE III

Case 3: Density Current. Comparison of the five models studied for 25 meter resolution and 8th order polynomials after 900 seconds using $\Delta t = 0.08$ seconds (Courant Number = 2.14).

While Table III shows that there is close agreement between all five models, it does show that the maximum and minimum values for all four variables do vary. Recall that this is the only case with viscosity and that only SE3 uses the true viscous stresses whereas the remaining four models use slightly modified diffusion operators in order to agree with the formulations presented in the Straka et al. paper [34]; in other words, each equation set uses a slightly different viscous operator and thereby each simulation represents the solution of a different governing equation, therefore, one should not expect to arrive at the same results for all the models. We present this test case because it exhibits a classical wave found in atmospheric modeling applications, namely, Kelvin-Helmholtz instabilities. Furthermore, diffusion operators of the type that we use here are representative of the kinds of diffusion mechanisms used today in industrial-type atmospheric models.

In terms of mass conservation, once again we see that SE1 is the worst with the other four models performing well and the two conservation forms (SE2C and SE3) performing best. In terms of energy conservation, only SE3 performs superbly; since SE3 uses the true Navier-Stokes equations along with the proper viscous stresses, this set is able to conserve both mass and energy even with the presence of viscosity. This is one big advantage of this equation set. In terms of efficiency we see that once again the ordering from best to worst is: SE1, SE4, SE2NC, SE2C, and SE3. This ordering conforms to the number of operation counts because the number of GMRES iterations is identical for four of the models (at 3) with one being different (4 for SE3).

4.1.4. Case 4: Linear Hydrostatic Mountain. This case is different from the previous three in that: i) it has a steady-state analytic solution and ii) it requires the implementation of non-reflecting boundary conditions. The previous three test cases used either no-flux (reflecting) or periodic boundary conditions. Figure 4.4 shows that the numerical (solid) and analytic (dashed) values for both horizontal and vertical velocities compare very well. Note that the actual computational domain is much larger than that shown in the figure. We show the domain that we used to compute the root-mean-square errors; the portion of the domain not shown is in fact where the sponge layer is non-zero ($\beta > 0$).

In Table IV we show the maximum and minimum values for all four variables for the five models after 30 hours. The values for all five models are identical, clearly illustrating that all five models have converged to the identical steady-state solution. Furthermore, Table V shows that indeed the RMS errors for all five models are virtually identical for the four variables; this is true at the three times reported (after 10, 20, and 30 hours). If we only showed the results after 30 hours, then one could argue that the reason why all the models agree so closely is because they all converge to the
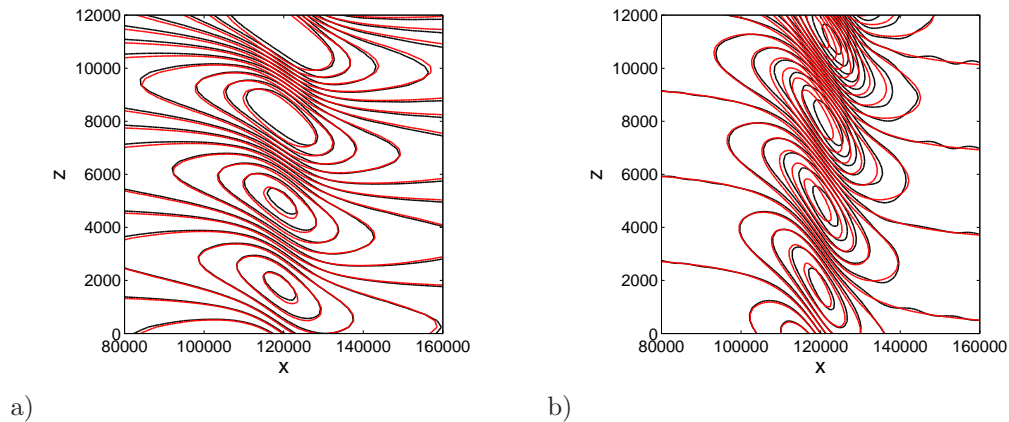
a)   b)

FIGURE 4.4. *Case 4: Linear Hydrostatic Mountain. The numerical solution (solid line) and analytic solution (dashed line) after 30 hours for 1200 meter (in x) and 240 meter (in z) resolution and 10th order polynomials. For the horizontal velocity (figure a) the contour values are between -0.025 to +0.025 with a contour interval of 0.005. For the vertical velocity (figure b) the contour values are between -0.005 to +0.005 with a contour interval of 0.0005.*

	SE1	SE2NC	SE2C	SE3	SE4
π'_{max}	2.05×10^{-6}	2.05×10^{-6}	2.05×10^{-6}	2.05×10^{-6}	2.05×10^{-6}
π'_{min}	-2.17×10^{-6}	-2.16×10^{-6}	-2.17×10^{-6}	-2.17×10^{-6}	-2.16×10^{-6}
u_{max}	4.72×10^{-2}	4.72×10^{-2}	4.72×10^{-2}	4.72×10^{-2}	4.72×10^{-2}
u_{min}	-4.01×10^{-2}	-4.01×10^{-2}	-4.01×10^{-2}	-4.01×10^{-2}	-4.01×10^{-2}
w_{max}	4.28×10^{-3}	4.28×10^{-3}	4.28×10^{-3}	4.28×10^{-3}	4.28×10^{-3}
w_{min}	-5.28×10^{-3}	-5.28×10^{-3}	-5.28×10^{-3}	-5.28×10^{-3}	-5.28×10^{-3}
θ'_{max}	2.68×10^{-2}	2.68×10^{-2}	2.68×10^{-2}	2.68×10^{-2}	2.68×10^{-2}
θ'_{min}	-3.73×10^{-2}	-3.73×10^{-2}	-3.73×10^{-2}	-3.73×10^{-2}	-3.73×10^{-2}
Mass Loss	1.23×10^{-8}	1.22×10^{-8}	1.20×10^{-8}	1.21×10^{-8}	1.22×10^{-8}
Energy Loss	3.29×10^{-8}	9.05×10^{-8}	9.06×10^{-8}	9.05×10^{-8}	2.03×10^{-7}
GMRES Iterations	11	10	12	12	12
WallClock Time	5954	5801	6861	6667	6543

TABLE IV

Case 4: Linear Hydrostatic Mountain. Comparison of the five models studied for 1200 meter (in x) and 240 meter (in z) resolution and 10th order polynomials after 30 hours using $\Delta t = 1.5$ seconds (Courant Number = 1.25).

same solution. However, the results in Table V show that there is more to it than that. For instance, the fact that all the models agree at all three times reported indicates that the models are being forced to yield this identical solution state. The only difference between this test case and all the others is the use of non-reflecting boundary conditions (NRBC). This result clearly indicates that it is the use of these NRBCs that is forcing the solution state regardless of the equation set being used. Revisiting Table IV once more and looking specifically at the mass and energy conservation it becomes immediately obvious that all five models are behaving identically even with respect to their conservation measures. Therefore, the NRBCs are not only imposing the solution state but are also affecting the conservation measures of the models and preventing the formally conservative SE3 from conserving to machine precision. This test case emphasizes the need for better NRBCs that are high order accurate and conservative; unfortunately, NRBCs such as the ones we use here are used today in all industrial-type nonhydrostatic atmospheric models. To overcome the first problem (accuracy), we have begun work on the construction of high-order NRBCs that can be used with high-order spatial and temporal discretizations (see [6] and [26]) but we are still far away from implementing such methods into Navier-Stokes models. The second problem is more complicated to overcome. While in the present work we describe nonhydrostatic mesoscale (limited area) models, these models will eventually also be used for global nonhydrostatic models (i.e., three-dimensional models on the

sphere, see [16] and [12] for a hydrostatic version of such a model). In global mode, the NRBCs along the lateral boundaries are eliminated by the periodicity of the sphere and, if the top NRBCs are replaced by reflecting boundary conditions then the conservation properties of the model will be retained; conservation of both mass and energy are vital for accurately modeling atmospheric processes at very long time-scales such as those typically run for climate change predictions.

In terms of efficiency, the ordering from best to worst is: SE2NC, SE1, SE4, SE3, and SE2C. For an equal number of GMRES iterations, SE2NC requires more floating point operations than both SE1 and SE4 due to the fact that an equation of state has to be solved (and this equation is exponential). However, for this test case, SE2NC needs an average of 10 GMRES iterations per time-step compared to 11 for SE1 and 12 for SE4 that then allows SE2NC to run faster.

Time	Variable	SE1	SE2NC	SE2C	SE3	SE4
10 hours	π'	1.56×10^{-7}	1.56×10^{-7}	1.56×10^{-7}	1.56×10^{-7}	1.56×10^{-7}
	u	2.99×10^{-3}	2.99×10^{-3}	2.99×10^{-3}	2.99×10^{-3}	3.00×10^{-3}
	w	1.90×10^{-4}	1.90×10^{-4}	1.90×10^{-4}	1.90×10^{-4}	1.91×10^{-4}
	θ'	2.46×10^{-3}	2.46×10^{-3}	2.46×10^{-3}	2.46×10^{-3}	2.46×10^{-3}
20 hours	π'	8.87×10^{-8}	8.86×10^{-8}	8.84×10^{-8}	8.85×10^{-8}	8.90×10^{-8}
	u	1.68×10^{-3}	1.68×10^{-3}	1.68×10^{-3}	1.68×10^{-3}	1.69×10^{-3}
	w	1.88×10^{-4}	1.88×10^{-4}	1.88×10^{-4}	1.88×10^{-4}	1.89×10^{-4}
	θ'	1.32×10^{-3}	1.32×10^{-3}	1.32×10^{-3}	1.32×10^{-3}	1.32×10^{-3}
30 hours	π'	6.74×10^{-8}	6.74×10^{-8}	6.72×10^{-8}	6.73×10^{-8}	6.79×10^{-8}
	u	1.27×10^{-3}	1.27×10^{-3}	1.27×10^{-3}	1.27×10^{-3}	1.28×10^{-3}
	w	1.88×10^{-4}	1.88×10^{-4}	1.87×10^{-4}	1.87×10^{-4}	1.89×10^{-4}
	θ'	8.99×10^{-4}	8.99×10^{-4}	8.99×10^{-4}	8.99×10^{-4}	8.98×10^{-4}

TABLE V

Case 4: Linear Hydrostatic Mountain. Root-Mean-Square errors for the four variables for 1200 meter (in x) and 240 meter (in z) resolution and 10th order polynomials for all five models using $\Delta t = 1.5$ seconds (Courant Number = 1.25).
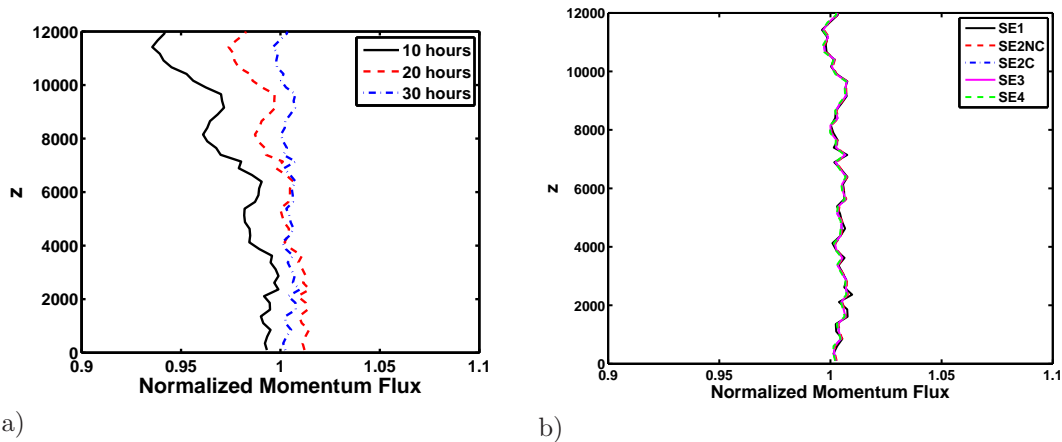

a)   b)

FIGURE 4.5. *Case 6: Linear Hydrostatic Mountain. Normalized momentum flux for 1200 meter (in x) and 240 meter (in z) resolution and 10th order polynomials for a) SE2NC at times 10, 20, and 30 hours, and b) for all five models at 30 hours.*

In Figure 4.5a we plot the normalized momentum flux at various times in the integration and in Fig. 4.5b we show the normalized momentum flux for all five models after 30 hours. Figure 4.5a shows that the simulations have reached steady-state after 30 hours; the RMS errors continue to change beyond this time but they change very little and we thereby assume this to be steady-state. Figure 4.5b shows that the normalized momentum flux values are essentially identical for all five models and are very good, that is, the values are almost equal to 1 everywhere which is the correct

theoretical result based on linear theory (see [32]).

4.2. Efficiency of the Semi-Implicit Time-Integrator. In this section we study the efficiency of the semi-implicit time-integrator compared to a fast explicit time-integrator, namely, the RK35 method [33] that we used previously with our explicit Navier-Stokes models [14]. In addition, we compare the semi-implicit method both with and without the *Schur* complement to see how much of an efficiency gain one gets. For this study we use SE2NC only since it represents the median of all the models in terms of efficiency and conservation measures.

In Figs. 4.7, 4.8, 4.9, and 4.10 the left panel (a) shows the wallclock time as a function of Courant number and the right panel (b) shows the average number of GMRES iterations required per time-step as a function of Courant number. Even though we list RK35 in this figure as well, the number of GMRES iterations per time-step is zero for this method since it is a fully explicit method. In all of these efficiency tests, the maximum Courant number reported for RK35 is the maximum Courant number allowed by this method.

Before discussing the four test cases in detail it is important to point out once again the differences between the *Schur* and *No Schur* systems. For set SE2NC, the *No Schur* form is the system defined by Eqs. (3.17)-(3.20) which, assuming N_P grid points, represents a $16N_p^2$ matrix problem. In contrast, the *Schur* form is defined by Eq. (3.26) and represents a N_p^2 matrix problem. The differences between these two systems go further: for the *No Schur* form, the differential operators are all first order whereas for the *Schur* form they are second order; this means that the two systems will have very different eigenspectra. To get a sense of this difference, we show the eigenspectra for the *No Schur* and *Schur* forms in Fig. 4.6 with $N_p = 289$ for case 2 (which happens to be the worst case scenario in terms of matrix conditioning). The condition number for the No Schur form is $\kappa(A_{NS}) = 2.6 \times 10^6$ whereas for the Schur form it is $\kappa(A_S) = 2.1 \times 10^2$. Note that the eigenvalues for the *No Schur* form are all imaginary whereas for the *Schur* form they are all real; this is consistent with the eigenvalues of first order (imaginary) and second order (real) differentiation operators.
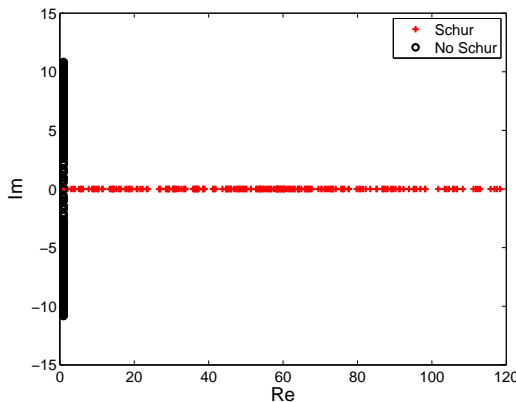

FIGURE 4.6. *The eigenspectra of SE2NC for case 2 with $N_p = 289$ for the Schur (red crosses) and No Schur (blue circles) forms of the semi-implicit method. The condition number for the* Schur *form is $\kappa(A_S) = 2.1 \times 10^2$ and for the* No Schur *form $\kappa(A_{NS}) = 2.6 \times 10^6$.*

4.2.1. Case 1: Inertia-Gravity Waves. Figure 4.7 shows that the efficiency (left panel) is linear for RK35 since doubling the Courant number yields a simulation that is twice as fast. In contrast, we see that the semi-implicit results are not linear due to the iterative solvers that may require a nonlinear increase in GMRES iterations with increased Courant number. In Fig. 4.7a we see that the *Schur* form semi-implicit method increases its efficiency with increasing Courant number. In contrast, the *No Schur* form semi-implicit method does not. In fact, the *No Schur* form reaches an optimal Courant number near 3 and increases in cost beyond this value. The difference in efficiency between the *Schur* and *No Schur* forms is partly due to the difference in the sizes of the matrix problem being solved (N_p^2 for Schur and $16N_p^2$ for No Schur) but also due to the difference in
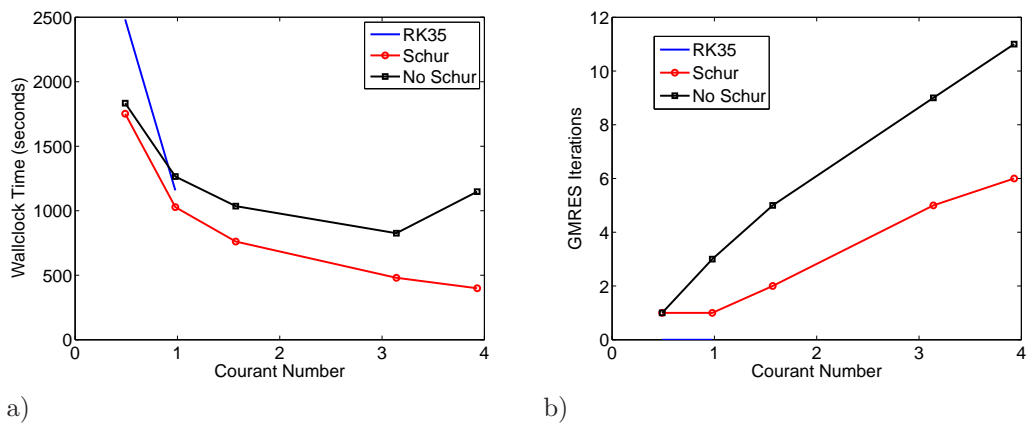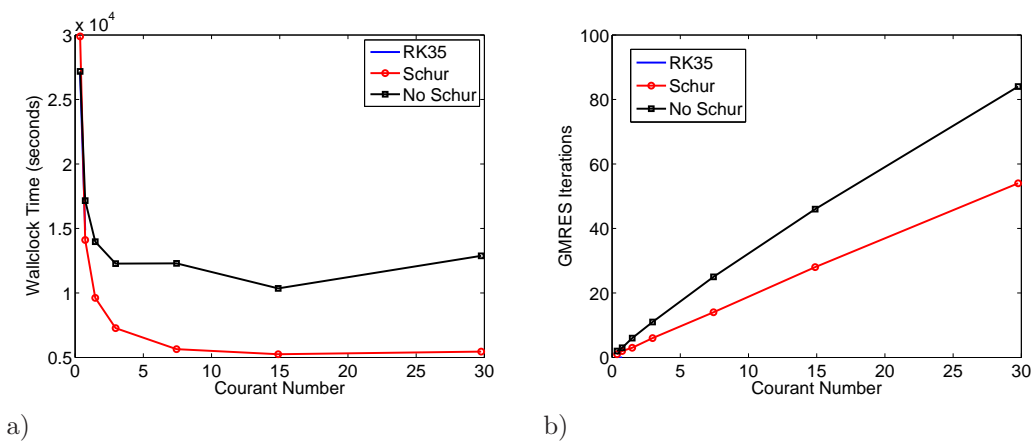
FIGURE 4.7. *Case 1: Inertia-Gravity Waves. The a) wallclock time and b) number of GMRES iterations as functions of the Courant number for 250 meter resolution with 10th order polynomials after 3000 seconds. The explicit Runge-Kutta method (RK35) is compared with the semi-implicit methods with and without the Schur complements (Schur and No Schur, respectively).*

the average number of GMRES iterations required per time-step. Figure 4.7b shows this difference and it is striking. Without a Schur complement (i.e., the *No Schur* form), the number of GMRES iterations increases linearly with increasing Courant number (i.e., time-step size).

FIGURE 4.8. *Case 2: Rising Thermal Bubble. The a) wallclock time and b) number of GMRES iterations as functions of the Courant number for 5 meter resolution with 10th order polynomials after 700 seconds. The explicit Runge-Kutta method (RK35) is compared with the semi-implicit methods with and without the Schur complements (Schur and No Schur, respectively).*

4.2.2. Case 2: Rising Thermal Bubble. In Fig. 4.8a (left panel) the curve for RK35 is not visible because the maximum Courant number allowed by this method is less than one while the semi-implicit methods (Schur and No Schur) allow Courant numbers of 30. For Courant numbers approaching the maximum allowed by the RK35 method, both semi-implicit methods give better efficiency than the RK35 method.

The *No Schur* form yields an optimal efficiency near a Courant number of 15. For values larger than 15, the efficiency decreases. The *Schur* form, on the other hand, continues to increase in efficiency with increasing Courant number but it begins to plateau near Courant numbers of 15.

Figure 4.8b (right panel) shows that the number of GMRES iterations increases linearly with increasing Courant number for both the *Schur* and *No Schur* forms but that this rate is much larger for the *No Schur* form. At a Courant number of 30, the *Schur* and *No Schur* forms are approaching 60 and 80 GMRES iterations, respectively; both of these values are unacceptable and work continues

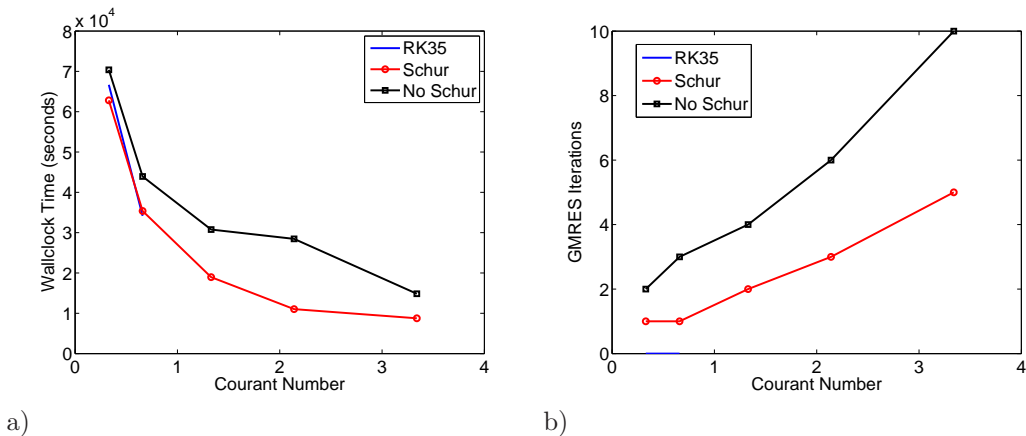on developing preconditioners that will require far fewer GMRES iterations.

a)   b)

FIGURE 4.9. *Case 3: Density Current. The a) wallclock time and b) number of GMRES iterations as functions of the Courant number for 25 meter resolution with 8th order polynomials after 900 seconds. The explicit Runge-Kutta method (RK35) is compared with the semi-implicit methods with and without the Schur complements (Schur and No Schur, respectively).*

4.2.3. Case 3: Density Current. In Fig. 4.9a (left panel) we see that for Courant numbers less than one, the efficiency of the *Schur* form is competitive with that of RK35 but that the *No Schur* form is not. For all the Courant numbers shown, the efficiency of the *No Schur* form continues to increase with increasing Courant number; this is always true for the *Schur* form. In the previous tests we saw that the *No Schur* form reached an optimal Courant number value whereas here it has not. So the question is: what accounts for this difference? Recall that this is the only test with viscosity (i.e., diffusion). In the current semi-implicit formulation we do not include the viscous operators in the linear implicit operators so that we must adhere to the explicit stability limit for diffusion. This is the reason why the maximum Courant numbers are smaller for this test than for the previous two. It should be pointed out that including the diffusion operator into the semi-implicit method is not at all problematic for the *No Schur* form but it is for the *Schur* form (for the *Schur* form one would have to invert a Helmholtz-type operator for both momentum and energy in order to construct the Schur form in terms of pressure). Thus, for the *No Schur* form we could include viscosity in the semi-implicit operators and perhaps see an increase in efficiency beyond Courant numbers of 3.

Figure 4.9b (right panel) shows that the number of GMRES iterations increases at an accelerated rate for the *No Schur* form but only increases linearly for the *Schur* form. The reason why both the *Schur* and *No Schur* forms yield comparable results is due to the small number of GMRES iterations required - these values are less than 10 iterations per time-step.

4.2.4. Case 4: Linear Hydrostatic Mountain. In Fig. 4.10a (left panel) we see that for Courant numbers less than one, the efficiency of the *Schur* form is competitive with that of RK35 but that the *No Schur* form is not. In fact, the *No Schur* form is not competitive at all (for any Courant number) with the explicit RK35. On the other hand, the *Schur* form is more efficient than the RK35 and this efficiency continues to increase as the Courant number is increased.

Figure 4.10b (right panel) tells us the reason for the *No Schur* form not being competitive, namely, the excessively large number of GMRES iterations. For the *No Schur* form, for Courant numbers beyond 1.5, the number of GMRES iterations have already climbed to 20 and continue to increase up to 50 for a Courant number of 3. Therefore, for the *No Schur* form, any efficiency gains offered by a larger time-step is offset by a larger implicit solver iteration count. In contrast, the *Schur* form exhibits efficiency gains for increasing Courant number. The number of GMRES iterations per time-step is much larger than for the other tests when we take into account the modest Courant numbers being used (case 2 requires much larger iteration counts but they correspond to Courant numbers near 30). The difference between this test and all the others is that non-reflecting boundary
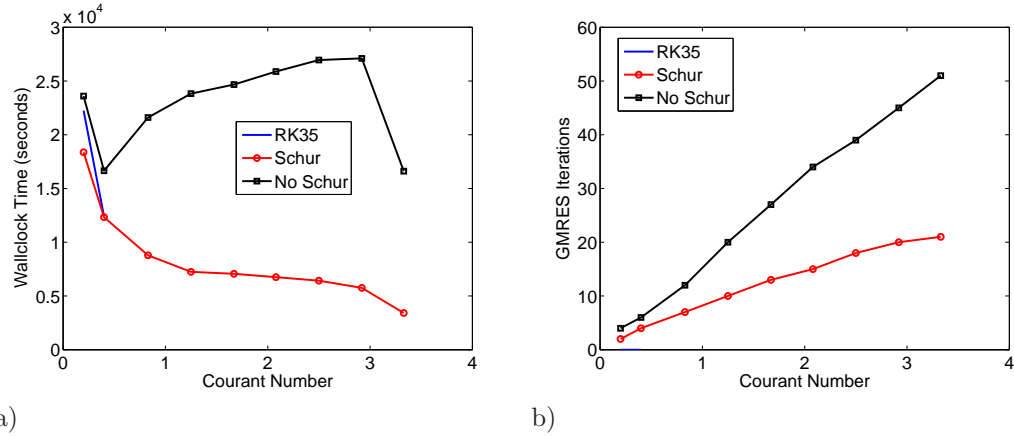
FIGURE 4.10. *Case 4: Linear Hydrostatic Mountain. The a) wallclock time and b) number of GMRES iterations as functions of the Courant number for 1200 meter (in x) and 240 meter (in z) resolution with 10th order polynomials after 30 hours. The explicit Runge-Kutta method (RK35) is compared with the semi-implicit methods with and without the Schur complements (Schur and No Schur, respectively).*

conditions (NRBCs) are employed. It should be noted that this test is very much indicative of the class of problems that must be run efficiently in nonhydrostatic mesoscale atmospheric modeling since almost all simulations require the use of NRBCs; efforts are currently underway to develop preconditioners that specifically target this class of boundary conditions.

4.3. Stability of the Semi-Implicit Method. To prove stability for the semi-implicit form of all five equation sets requires going back to the original time-integration statement of the problem which is

$$\frac{\partial \boldsymbol{q}}{\partial t} = \{N(\boldsymbol{q})\} + [L(\boldsymbol{q})].$$

Recall that here, we treat the nonlinear terms $N(\boldsymbol{q})$ explicitly and the linear terms $L(\boldsymbol{q})$ implicitly in an IMEX approach. At this point we assume a system of ordinary differential equations where the right-hand-side operators have already been discretized in space in a method of lines approach. Recall that we chose the linear operator to contain the fastest waves in the system, namely the acoustic and gravity (i.e., buoyancy) waves. Furthermore, recall that the nonlinear operator does not contain these waves any more since they have been subtracted. Thereby the nonlinear operator only contains the advective waves that are far slower than the acoustic or gravity waves.

Thus in order to maintain stability, we require that the Courant number associated with the advective waves satisfy the Courant-Friedrichs-Lewy (CFL) condition of standard explicit time-integrators (in this case we are using the explicit BDF2, see [21]). Since the linear operator is implicit, then the Courant number with respect to the acoustic and gravity waves is unlimited. In fact, we are using BDF2 for the implicit part that is both A-stable (stable for all values of z in the left-hand plane) and L-stable (the amplification function goes to zero for $z \to -\infty$). This means that as long as we adhere to the explicit CFL condition for the advective waves then we are guaranteed stability, this is certainly true for the full system, i.e., the *No Schur* form (see , e.g., [9]). For the *Schur* form we have to perform further analysis.

Assuming that we are adhering to the explicit CFL limit of the slow-moving waves, the only possibility of instabilities stems from the conversion of the full system (i.e., *No Schur*) to the reduced or *Schur* form. For example, for SE1, to extract the Schur complement requires the construction of the matrix $\boldsymbol{C}_1$ given in Eq. (3.12). This is also true for SE2NC (see Eq. (3.23)), SE2C (see Eq. (3.35)), SE3 (see Eq. (3.44)), and SE4 (see Eq. (3.51)). The only possibility for instabilities to occur is if these matrices become singular at any point (z).

4.3.1. SE1. For SE1, we see that this can only occur if and only if

$$c_1 \equiv 1 + (\alpha\lambda)^2 \frac{g}{\theta_0} \frac{d\theta_0}{dz} = 0.$$

Using the definition of the Brunt-Väisälä frequency

$$\mathcal{N}^2 = g \frac{d}{dz}(\ln \theta_0(z))$$

the stability condition can be rewritten in the form

$$1 + (\alpha\lambda)^2 \mathcal{N}^2 = 0 \quad \leftrightarrow \quad \mathcal{N}^2 = -\frac{1}{(\alpha\lambda)^2}.$$

With the assumption of a stable stratified reference atmosphere $\frac{d\theta_0}{dz} > 0$, this condition always fails.

4.3.2. SE2NC. The analysis for SE2NC is identical since $c_{2NC} = c_1$.

4.3.3. SE2C. For SE2C, instabilities can occur if and only if

$$c_{2C} \equiv 1 + (\alpha\lambda)^2 \frac{g}{G_0} \frac{dG_0}{dz} = 0.$$

Since $G_0 = \frac{\Theta_0}{\rho_0} = \theta_0$ then we see that c_{2C} can never be zero because it is the same expression as for SE1 and SE2NC.

4.3.4. SE3. For SE3 we need to show that for instabilities to arise, the following statement must be true

$$c_3 \equiv 1 + (\alpha\lambda)^2 \frac{g}{h_0 - \phi} \frac{dh_0}{dz} = 0.$$

Since $h_0 = c_p T_0 + \phi$ then this expression becomes

$$c_3 = 1 + (\alpha\lambda)^2 \frac{g}{c_p T_0} \left(c_p \frac{dT_0}{dz} + g \right).$$

Using the definition of potential temperature $T_0 = \theta_0 \pi_0$, expanding, and using the definition of hydrostatic balance $\frac{d\pi_0}{dz} = -\frac{g}{c_p \theta_0}$ allows us to write

$$c_3 = 1 + (\alpha\lambda)^2 \frac{g}{c_p \theta_0} \frac{d\theta_0}{dz}$$

thereby proving that c_3 is in fact equal to the terms for SE1, SE2NC, and SE2C.

4.3.5. SE4. For SE4 it is not clear that the same analysis holds since the expression that we have to analyze is

$$c_4 \equiv 1 + (\alpha\lambda)^2 \left(\frac{1}{\gamma P_0} \frac{dP_0}{dz} - \frac{1}{\rho_0} \frac{d\rho_0}{dz} \right). \tag{4.1}$$

However, writing the pressure as follows:

$$P_0(z) = P_A \left(\frac{\rho_0(z) R \theta_0(z)}{P_A} \right)^\gamma$$

and differentiating and rearranging yields

$$\frac{1}{\gamma P_0} \frac{dP_0}{dz} = \frac{1}{\theta_0} \frac{d\theta_0}{dz} + \frac{1}{\rho_0} \frac{d\rho_0}{dz}$$

that, when substituted into Eq. (4.1) yields

$$c_4 = 1 + (\alpha\lambda)^2 \frac{g}{\theta_0} \frac{d\theta_0}{dz}$$

that is identical to the expression for SE1, SE2NC, and SE2C and so c_4 can never be zero. Therefore, as one would expect, the stability condition for all the models are identical - one would expect this because although the five equation sets are written differently, they represent the same dynamical system and must have the same stability condition.

This brief analysis proves stability of the semi-implicit method using the Schur form, at least when a hydrostatically-balanced reference state $\boldsymbol{q}_0$ is used in the semi-implicit formulation. Other reference states can also be used as long as the matrix $\boldsymbol{C}$ remains non-singular but the proof of stability is more complicated.

5. Conclusions. We have presented semi-implicit formulations of five different forms of the compressible Navier-Stokes equations (NSE) used in nonhydrostatic atmospheric modeling. These equations have typically been solved either explicitly or semi-implicitly along the vertical direction only; the common reason given for not solving the equations semi-implicitly in all directions, as we have done here, is that this approach is not competitive with explicit forms. Our experiences have shown otherwise and in this work we show that this is in fact the case for all the equations being used today especially if the Schur complement is extracted. If the full system (i.e., the No Schur form) is solved instead, then it is still faster than fast explicit methods but only by a small margin; this, however, we were only able to show in the special case when either reflecting (no-flux) or periodic boundary conditions were used. The true advantage of the semi-implicit formulation can only be realized if the Schur complement is used; this we were able to show for all types of boundary conditions including the more realistic (in mesoscale nonhydrostatic atmospheric modeling) non-reflecting boundary conditions (NRBCs). In addition, we show that choosing one form of the NSE over another can be quite advantageous if the most efficient form of the equations is sought. Specifically, we found that the equation sets in conservation (flux-form) form are not as efficient as those that are in non-conservation form. While it is important to conserve as many variables as possible, we have found that those sets that use density as the mass variable conserve mass quite well; only set 3, which uses total energy, was able to conserve both mass and energy up to machine precision. Furthermore, the currently used NRBCs adversely affect both accuracy and conservation which motivates the need for better non-reflecting boundary conditions that are, at the very least, high-order. Set 1, which does not use density as its mass variable, was the worst in terms of mass conservation, regardless of the type of boundary conditions used.

Comparing the *Schur* and *No Schur* semi-implicit forms, we see that the *Schur* form is far more efficient than explicit methods and this efficiency increases with increasing Courant number (i.e., time-step). However, the *No Schur* form reaches an optimal time-step (Courant number) with the cost then increasing with increasing time-step. The reason that the *Schur* form beats the explicit method so easily whereas the *No Schur* form struggles is partly due to the different dimension sizes of the linear matrix problem that both methods solve. In the case of the *Schur* form this system is N_p^2 whereas for the *No Schur* form it is $16N_p^2$ where N_p denotes the number of gridpoints in the domain. The other reason is due to the difference in the number of GMRES iterations required by the two semi-implicit forms; the number of iterations vary from test case but the general trend observed is that, on average, the *No Schur* form requires almost twice as many GMRES iterations per time-step as the *Schur* form. This result shows that one must always seek the *Schur* complement form and we are currently working on generalizing this study to include semi-implicit (additive) Runge-Kutta methods into this framework. Furthermore, in the future we shall study the impact of various preconditioners to see if we can decrease the number of GMRES iterations for both the *Schur* and *No Schur* forms.

REFERENCES

[1] N. Ahmad and J. Lindeman, *Euler solutions using flux–based wave decomposition*, Int. J. Numer. Meth. Fl., 54 (2007), pp. 47–72.

[2] N. Botta, R. Klein, S. Langenberg, and S. Lützenkirchen, *Well balanced finite volume methods for nearly hydrostatic flows*, J. Comp. Phys., 196 (2004), pp. 539–565.

[3] D.M. Burridge, *A split semi-implicit reformulation of the Bushby–Timpson 10-level model*, Q. J. R. Meteorol. Soc., 101 (1975), pp. 777–792.

[4] R.L. Carpenter, K.K. Droegemeier, P.R. Woodward, and C.E. Hane, *Application of the piecewise parabolic method (PPM) to meteorological modeling*, Mon. Wea. Rev., 118 (1990), pp. 586–612.

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

[6] J. Dea, F.X. Giraldo, and B. Neta, *High–order Higdon non–reflecting boundary conditions for the linearized Euler equations:no mean flow case*, Wave Motion, 46 (2009), pp. 210–220.

[7] PF Fischer, *An overlapping Schwarz method for spectral element solution of the incompressible Navier-Stokes equations*, JOURNAL OF COMPUTATIONAL PHYSICS, 133 (1997), pp. 84–101.

[8] ———, *Projection techniques for iterative solution of Ax=b with successive right-hand sides*, COMPUTER METHODS IN APPLIED MECHANICS AND ENGINEERING, 163 (1998), pp. 193–204.

[9] J Frank, W Hundsdorfer, and JG Verwer, *On the stability of implicit-explicit linear multistep methods*, APPLIED NUMERICAL MATHEMATICS, 25 (1997), pp. 193–205.

[10] T Galchen and RCJ Somerville, *USE OF A COORDINATE TRANSFORMATION FOR SOLUTION OF NAVIER-STOKES EQUATIONS*, JOURNAL OF COMPUTATIONAL PHYSICS, 17 (1975), pp. 209–228.

[11] FX Giraldo, *The Lagrange-Galerkin spectral element method on unstructured quadrilateral grids*, JOURNAL OF COMPUTATIONAL PHYSICS, 147 (1998), pp. 114–146.

[12] F.X. Giraldo, *Semi-implicit time-integrators for a scalable spectral element atmospheric model*, Q. J. R. Meteorol. Soc., 131 (2005), pp. 2431–2454.

[13] ———, *High-order triangle-based discontinuous Galerkin methods for hyperbolic equations on a rotating sphere*, J. Comp. Phys., 214 (2006), pp. 447–465.

[14] F.X. Giraldo and M. Restelli, *A study of spectral element and discontinuous Galerkin methods for the Navier–Stokes equations in nonhydrostatic mesoscale atmospheric modeling: Equation sets and test cases*, J. Comp. Phys., 227 (2008), pp. 3849–3877.

[15] ———, *High-order semi-implicit time-integrators for a triangular discontinuous galerkin oceanic shallow water model*, Int. J. Numer. Meth. Fl., in press (2009).

[16] FX Giraldo and TE Rosmond, *A scalable spectral element eulerian atmospheric model (SEE-AM) for NWP: Dynamical core tests*, MONTHLY WEATHER REVIEW, 132 (2004), pp. 133–153.

[17] F.X. Giraldo and T. Warburton, *A high-order triangular discontinuous Galerkin oceanic shallow water model*, Int. J. Numer. Meth. Fl., 56 (2008), pp. 899–925.

[18] F. X. Giraldo, *Hybrid Eulerian-Lagrangian semi-implicit time-integrators*, COMPUTERS & MATHEMATICS WITH APPLICATIONS, 52 (2006), pp. 1325–1342.

[19] RM Hodur, *The Naval Research Laboratory's coupled ocean/atmosphere mesoscale prediction system (COAMPS)*, MONTHLY WEATHER REVIEW, 125 (1997), pp. 1414–1430.

[20] ZI Janjic, *A nonhydrostatic model based on a new approach*, METEOROLOGY AND ATMOSPHERIC PHYSICS, 82 (2003), pp. 271–285.

[21] G.E. Karniadakis, M. Israeli, and S.A. Orszag, *High–order splitting methods for the incompressible Navier–Stokes equations*, J. Comp. Phys., 97 (1991), pp. 414–443.

[22] J.B. Klemp, W.C. Skamarock, and J. Dudhia, *Conservative split-explicit time integration methods for the compressible nonhydrostatic equations*, Mon. Wea. Rev., 135 (2007), pp. 2897–2913.

[23] JB Klemp and RB Wilhelmson, *SIMULATION OF 3-DIMENSIONAL CONVECTIVE STORM DY-NAMICS*, JOURNAL OF THE ATMOSPHERIC SCIENCES, 35 (1978), pp. 1070–1096.

[24] M. Kwizak and A.J. Robert, *A semi–implicit scheme for grid point atmospheric models of the primitive equations*, Mon. Wea. Rev., 99 (1971), pp. 32–36.

[25] Matthias Läuter, Francis X. Giraldo, Dörthe Handorf, and Klaus Dethloff, *A discontinuous Galerkin method for the shallow water equations in spherical triangular coordinates*, JOURNAL OF COMPUTA-TIONAL PHYSICS, 227 (2008), pp. 10226–10242.

[26] J. Lindquist, B. Neta, and F.X. Giraldo, *A spectral element solution of the Klein–Gordon equation with high-order treatment of time and non-reflecting boundary*, Wave Motion, in press (2009).

[27] M. Restelli and F. X. Giraldo, *A conservative discontinuous galerkin semi-implicit formulation for the navier-stokes equations in nonhydrostatic mesoscale modeling*, SISC, (2009).

[28] Kazuo Saito, Jun-Ichi Ishida, Kohei Aranami, Tabito Hara, Tomonori Segawa, Masami Narita, and Yuuki Honda, *Nonhydrostatic atmospheric models and operational development at JMA*, JOURNAL OF THE METEOROLOGICAL SOCIETY OF JAPAN, 85B (2007), pp. 271–304.

[29] U. Schattler, G. Doms, and J. Steppeler, *Requirements and problems in parallel model development at DWD*, Scientific Programming, 8 (2000), pp. 13–22.

[30] ——— , *Requirements and problems in parallel model development at DWD*, Scientific Programming, 8 (2000), pp. 13–22.

[31] W.C. Skamarock and J.B. Klemp, *Efficiency and accuracy of the Klemp–Wilhelmson time-splitting technique*, Mon. Wea. Rev., 122 (1994), pp. 2623–2630.

[32] R.B. Smith, *The influence of mountains on the atmosphere*, Adv. Geophys., 221 (1979), pp. 87–.

[33] Raymond J. Spiteri and Steven J. Ruuth, *A new class of optimal high-order strong-stability-preserving time discretization methods*, SIAM J. Numer. Anal., 40 (2002), pp. 469–491.

[34] J.M. Straka, R.B. Wilhelmson, L.J. Wicker, J.R. Anderson, and K.K. Droegemeier, *Numerical solutions of a non–linear density current: A benchmark solution and comparisons*, Int. J. Numer. Meth. Fl., 17 (1993), pp. 1–22.

[35] M. Tanguay, A. Robert, and R. Laprise, *A semi-implicit semi-Lagrangian fully compressible regional forecast model*, Mon. Wea. Rev., 118 (1990), pp. 1970–1980.

[36] M.C. Tapp and P.W. White, *A non-hydrostatic mesoscale model*, Q. J. R. Meteorol. Soc., 102 (1976), pp. 277–296.

[37] S Thomas, C Girard, G Doms, and U Schattler, *Semi-implicit scheme for the DWD Lokal-Modell*, METEOROLOGY AND ATMOSPHERIC PHYSICS, 73 (2000), pp. 105–125.