SIAM J. SCI. COMPUT.
Vol. 32, No. 6, pp. 3394–3425

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 (implicit-explicit) 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, density current (i.e., Kelvin–Helmholtz instabilities), and mountain test cases; the latter test case requires the implementation of nonreflecting 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 nonreflecting boundary conditions (NRBCs). It is shown that the NRBCs exert a strong impact on the accuracy and efficiency of the models.

Key words. compressible flow, element-based Galerkin methods, Euler, implicit-explicit, Lagrange, Legendre, Navier–Stokes, nonhydrostatic, spectral elements, time-integration

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

DOI. 10.1137/090775889

1. Introduction.

It can be argued that the single most important property of an operational nonhydrostatic mesoscale atmospheric model 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 multicore 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

*Received by the editors November 3, 2009; accepted for publication (in revised form) August 30, 2010; published electronically November 30, 2010.

http://www.siam.org/journals/sisc/32-6/77588.html

†Department of Applied Mathematics, Naval Postgraduate School, Monterey, CA 93943 (fxgirald@nps.edu). The research of this author was supported by the Office of Naval Research Program element PE-0602435N.

‡Departamento de Ecuaciones Diferenciales y Análisis Numérico, Universidad de Sevilla, Facultad de Matemáticas, Universidad de Sevilla, C/Tarfia s/n, 41012 Sevilla, Spain (mrestelli@gmail.com). The research of this author was supported by the Office of Naval Research Visiting Science Program and Junta de Andalucía Grant P07-FQM-02538.

§Zuse Institute Berlin (ZIB), Takustrasse 7, 14195 Berlin, Germany (matthias.laeuter@awi.de). The research of this author was supported by the Office of Naval Research Visiting Science Program and German Research Foundation grant LA2455/1-1.

element-based Galerkin (EBG) methods such as spectral element (SE) and discontinuous Galerkin (DG) methods (see [17] and [29] 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 [5, 3, 1], the spectral model found in [9], and our SE and DG models found in [17] and [29]. Finite element (FE), SE, and DG methods are preferable to spectral methods because they can be shown to scale on massively parallel computers; this is not the case with spectral methods due to the all-to-all communication required. One of the biggest advantages that FE, SE, and DG methods have over the FD method is that no terrain following coordinates of the type introduced in [12] needs to be included in the governing equations. Of course, the orography (e.g., mountains) has to be accounted for in some manner, but 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 [15, 20, 27, 18]) 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 [19, 14]). Another advantage that FE, SE, and DG methods have over the FD and FV methods is that high-order solutions (greater than third 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 [13]. 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 section 4 use tenth-order polynomials per element. Although lower-order (fourth) can certainly be used, we have chosen high-order as a worst case scenario with respect to efficiency; we discuss this point below.

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 to 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 (limited area) nonhydrostatic weather models use split-explicit methods [25], 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; some exceptions are the global models of the United Kingdom Meteorological Office [7], the global environmental multiscale (GEM) model of Canada [40], and the Arome model of Meteo-France [9]. However, there are many more models that use the split-explicit approach including the operational models of

the U.S. Navy [21], the National Center for Atmospheric Research [24], Pennsylvania State University/National Center for Atmospheric Research [39], U.S. National Center for Environmental Prediction [22], German Weather Service [31], and the Japanese Meteorological Agency [30], 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 [38]. The goal of this paper is to confirm that semi-implicit methods, indeed, compete with explicit methods; semi-implicit methods are currently under scrutiny by many operational weather and climate research centers, especially for petascale to exascale computing where it is believed that the iterative solvers used in semi-implicit methods may not be able to achieve the linear scalability expected with explicit methods.

To construct semi-implicit formulations, i.e., implicit-explicit (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 toward 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 (SE methods); however, our results should hold for all other spatial discretization methods, the only exceptions are Godunov methods that require special treatment in the construction of the Schur form (see [29] for issues facing these methods); we will extend these results to Godunov-type methods in a future paper. The next step will be to compare the semi-implicit formulations in all directions (as we have done here) with the 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 section 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 section 4 we present the results for all five semi-implicit models using three 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 section 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 nonconservative form using Exner pressure, velocity, and potential temperature,
2. (set 2NC) the nonconservative 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 nonconservative 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 atmo-

spheric 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 by 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. In what follows, we shall use the notation SE1, SE2NC, etc., where SE is an abbreviation for spectral element. We use this naming convention in order to be able to compare the results of this paper with those of our paper [17], where we used both SE and DG forms of the equations; however, it should be understood that the emphasis of what follows is on the time-integration and resulting semi-implicit formulations.

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 nonconservation form. We thus consider the system

$$\begin{aligned} \frac{\partial \pi}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\pi + (\gamma-1)\pi\boldsymbol{\nabla}\cdot\boldsymbol{u} &= 0, \\ \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} + c_p\theta\boldsymbol{\nabla}\pi + g\boldsymbol{k} &= \mu\nabla^2\boldsymbol{u}, \\ \frac{\partial \theta}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta &= \mu\nabla^2\theta, \end{aligned} \tag{2.1}$$

where the solution vector is $(\pi, \boldsymbol{u}^{\mathcal{T}}, \theta)^{\mathcal{T}}$, $\pi = (\frac{P}{P_A})^{R/c_p}$ 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 $\boldsymbol{\nabla} = (\frac{\partial}{\partial x}, \frac{\partial}{\partial z})^{\mathcal{T}}$, the gravitational constant g, the gas constant $R = c_p - c_v$, the specific heats for constant pressure c_p and constant volume 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 (2.1) as

$$\begin{aligned} \frac{\partial \pi'}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\pi' + w\frac{d\pi_0}{dz} + (\gamma-1)(\pi' + \pi_0)\boldsymbol{\nabla}\cdot\boldsymbol{u} &= 0, \\ \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\boldsymbol{u} + c_p\theta\boldsymbol{\nabla}\pi' - g\frac{\theta'}{\theta_0}\boldsymbol{k} &= \mu\nabla^2\boldsymbol{u}, \\ \frac{\partial \theta'}{\partial t} + \boldsymbol{u}\cdot\boldsymbol{\nabla}\theta' + w\frac{d\theta_0}{dz} &= \mu\nabla^2\theta \end{aligned} \tag{2.2}$$

that has been expanded and simplified in order to enforce hydrostatic balance 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 as defined in [35]). 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 (see [17] for a discussion of the complication of using the true viscous stresses for the other equation sets and the reason for choosing simpler forms).

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

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) &= 0, \\ \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho} \boldsymbol{\nabla} P + g \boldsymbol{k} &= \mu \nabla^2 \boldsymbol{u}, \\ \frac{\partial \theta}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta &= \mu \nabla^2 \theta, \end{aligned} \tag{2.3}$$

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 temperature $\theta(\boldsymbol{x}, t) = \theta_0(z) + \theta'(\boldsymbol{x}, t)$, where the reference values are in hydrostatic balance, i.e., $\frac{dP_0}{dz} = -\rho_0 g$, allows us to rewrite (2.3) as

$$\begin{aligned} \frac{\partial \rho'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho' + w \frac{d\rho_0}{dz} + (\rho' + \rho_0) \boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0, \\ \frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho' + \rho_0} \boldsymbol{\nabla} P' + \frac{\rho'}{\rho' + \rho_0} g \boldsymbol{k} &= \mu \nabla^2 \boldsymbol{u}, \\ \frac{\partial \theta'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \theta' + w \frac{d\theta_0}{dz} &= \mu \nabla^2 \theta. \end{aligned}$$

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

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} &= 0, \\ \frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P \boldsymbol{\mathcal{I}}_2 \right) + \rho g \boldsymbol{k} &= \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\boldsymbol{U}}{\rho} \right), \\ \frac{\partial \Theta}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\Theta \boldsymbol{U}}{\rho} \right), &= \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\Theta}{\rho} \right), \end{aligned} \tag{2.4}$$

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 $\boldsymbol{\mathcal{I}}_2$ 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 temperature 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 (2.4) as

$$\begin{aligned}
\frac{\partial \rho'}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} &= 0, \\
\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P' \boldsymbol{\mathcal{I}}_2 \right) + \rho' g \boldsymbol{k} &= \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\boldsymbol{U}}{\rho} \right), \\
\frac{\partial \Theta'}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\Theta \boldsymbol{U}}{\rho} \right) &= \boldsymbol{\nabla} \cdot \left(\mu \rho \boldsymbol{\nabla} \frac{\Theta}{\rho} \right).
\end{aligned}$$

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

$$\begin{aligned}
\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} &= 0, \\
\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P \boldsymbol{\mathcal{I}}_2 \right) + \rho g \boldsymbol{k} &= \boldsymbol{\nabla} \cdot \boldsymbol{F}^{\text{visc}}_{\boldsymbol{u}}, \\
\frac{\partial E}{\partial t} + \boldsymbol{\nabla} \cdot \left[\frac{(E+P)\boldsymbol{U}}{\rho} \right] &= \boldsymbol{\nabla} \cdot \boldsymbol{F}^{\text{visc}}_{e},
\end{aligned} \tag{2.5}$$

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:

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

and

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

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 (2.5) as

$$\begin{aligned}
\frac{\partial \rho'}{\partial t} + \boldsymbol{\nabla} \cdot \boldsymbol{U} &= 0, \\
\frac{\partial \boldsymbol{U}}{\partial t} + \boldsymbol{\nabla} \cdot \left(\frac{\boldsymbol{U} \otimes \boldsymbol{U}}{\rho} + P' \boldsymbol{\mathcal{I}}_2 \right) + \rho' g \boldsymbol{k} &= \boldsymbol{\nabla} \cdot \boldsymbol{F}^{\text{visc}}_{\boldsymbol{u}}, \\
\frac{\partial E'}{\partial t} + \boldsymbol{\nabla} \cdot \left[\frac{(E+P)\boldsymbol{U}}{\rho} \right] &= \boldsymbol{\nabla} \cdot \boldsymbol{F}^{\text{visc}}_{e}.
\end{aligned}$$

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

$$\begin{aligned}
\frac{\partial \rho}{\partial t} + \boldsymbol{\nabla} \cdot (\rho \boldsymbol{u}) &= 0, \\
\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho} \boldsymbol{\nabla} P + g \boldsymbol{k} &= \mu \nabla^2 \boldsymbol{u}, \\
\frac{\partial P}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} P + \gamma P \boldsymbol{\nabla} \cdot \boldsymbol{u} &= \mu \gamma \frac{P}{\theta} \nabla^2 \theta,
\end{aligned} \tag{2.6}$$

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 (2.6) as

$$\begin{aligned}
\frac{\partial \rho'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \rho' + w \frac{d\rho_0}{dz} + (\rho' + \rho_0) \boldsymbol{\nabla} \cdot \boldsymbol{u} &= 0, \\
\frac{\partial \boldsymbol{u}}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} \boldsymbol{u} + \frac{1}{\rho' + \rho_0} \boldsymbol{\nabla} P' + \frac{\rho'}{\rho' + \rho_0} g \boldsymbol{k} &= \mu \nabla^2 \boldsymbol{u}, \\
\frac{\partial P'}{\partial t} + \boldsymbol{u} \cdot \boldsymbol{\nabla} P' + w \frac{dP_0}{dz} &= \mu \gamma \frac{P}{\theta} \nabla^2 \theta.
\end{aligned}$$

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 SE 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 FD and SE methods. For the construction of semi-implicit methods for Godunov-type methods, such as the FV and DG methods, see Restelli and Giraldo [29], 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 SE discretization for the equations described herein, see [17]. 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 $L(\boldsymbol{q})$ that approximates $S(\boldsymbol{q})$ and contains the terms responsible for the acoustic and gravity waves (the precise form of which will be defined in section 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 (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 [14, 16], we now consider a generic K-step (multistep method) 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

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

that then allows us to write (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 second-order backward difference formula (BDF2) method, assuming constant time-stepping, are $\alpha_0 = 4/3$, $\alpha_1 = -1/3$, $\chi = 2/3$, $\beta_0 = 2$, and $\beta_1 = -1$ (see [18] for BDF-K methods of orders one through six); in this work, we use BDF2 exclusively, i.e., K = 2. Ideally one would like to balance the errors between space and time (and boundary conditions) as we show in [28] for a simple equation. However, in this work, we use BDF2 only as a proxy for measuring the efficiency of multistep IMEX methods. We cannot use higher-order BDFs because they are not A-stable (e.g., see [18]). Clearly, using a second-order method in time with a tenth-order method in space means that the time-integrator will dominate the model error.

The crux of the semi-implicit method, as is evident in (3.2), is the derivation of the linear operator L. The success of the semi-implicit 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 semi-implicit 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 consider only two types of boundary conditions: no-flux (i.e., reflecting) and nonreflecting boundary conditions (NRBCs). 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 $\widehat{\boldsymbol{n}}_\Gamma = \left(0, \boldsymbol{n}_\Gamma^{\mathcal{T}}, 0\right)^{\mathcal{T}} \in R^4$ 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 NRBCs, respectively, and $\boldsymbol{q}_b$ is the free-stream (boundary) values of the state variable $\boldsymbol{q}$.

It turns out that, to impose the NRBCs given above in a strong sense, one can write the semidiscrete (in time) equations as

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

where α and β are Newtonian relaxation coefficients that drive the solution toward the boundary reference value such that $\alpha \to 1$, $\beta \to 0$ in the interior and $\alpha \to 0$, $\beta \to 1$ as the nonreflecting boundaries are approached; this boundary condition is applied to the entire solution vector $\boldsymbol{q}$. In general, we define these parameters as

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

and specifically for the linear hydrostatic mountain (case 3, where we refer to this NRBC as a *sponge*), we define $z_s = 12km$, 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, we define $x_s^{\text{right}} = x_{\max} - 20km$ and $x_t^{\text{right}} = x_{\max}$.

In contrast, for the no-flux boundaries, the boundary condition need be applied only 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 the 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 defined only 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 (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 has been defined, the linear system to be solved at each time step is given by (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 two-dimensional compressible Navier–Stokes 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 to reformulate 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 (monolithic) 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, 16 times smaller compared to the matrix inverted in the monolithic approach (see, e.g., [2] for why it must be pressure and not some other variable). Since our discussion is independent from the chosen spatial discretization, we refer here to the time semidiscretized problem; the fully discrete problem is 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 [10] for a description). In future work, we will explore the effects of various preconditioners.

As an example to show the construction of the Schur and No Schur forms, we begin with SE1; for the rest of the equation sets, we refer the reader to the appendices. For SE1 we follow [26, 4, 37, 6, 36] 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 (3.7), we rely on the same reference state π_0, θ_0 introduced in section 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 (3.7). Substituting (3.7) into (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 NRBCs. 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 (3.10) into (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 (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 C_1 with P_1 such that (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 (3.6). We can now substitute (3.15) into (3.8) to get

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

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 by any nonsymmetric iterative solver. Note further that the solution of this linear problem satisfies both nonreflecting and no-flux boundary conditions. Upon getting a solution for π_{tt} from (3.16), we can then solve for $\boldsymbol{u}_{tt}$ using (3.15). To solve for θ_{tt}, we next solve (3.10). Once $\boldsymbol{q}_{tt}$ is known, we then extract the solution $\boldsymbol{q}^{n+1}$ using (3.4).

4. Results. In this section we validate the five semi-implicit models on a test case suite of three problems using, for the spatial discretization, the SE method. For the definitions of the test cases, as well as for the details of the spatial discretization, we refer the reader to [17].

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 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 that, 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, take 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 deviations as follows. We define change in mass as

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

Similarly, we define the change in energy as

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

and 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 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 (for case 1 the solution is Boussinesq, and for case 3 it is based on linear theory and is valid only far away from the nonreflecting boundaries).

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 (in serial mode) with a clock speed of 2.8 GHz 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

$$\mathcal{C}_{sound,advect} = \max \left(\frac{C_{sound,advect} \Delta t}{\Delta s} \right),$$

where $\mathcal{C}_{sound}$ and $\mathcal{C}_{advect}$ are the Courant numbers with and without the sound waves and the characteristic speeds are defined as $C_{sound} = U + \sqrt{a}$ and $C_{advect} = U$, with $U = |\boldsymbol{n} \cdot \boldsymbol{u}|$ being the velocity in the direction $\boldsymbol{n}$ (where $\boldsymbol{n}$ is measured along the edges of the subcells formed by the high-order grid points), a is the sound speed, and $\Delta s = \sqrt{\Delta x^2 + \Delta z^2}$ 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 semianalytic 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 it 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 [8, 28]), but we are still far from implementing such methods into Navier–Stokes models. Unfortunately, sponge-based NRBCs are those typically used today in operational 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 section 4.1.3, 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 (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

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

where $N_s = 40{,}000$ are the number of sample points used to construct the RMS errors. Finally, in addition to the diagnostics used for the other test cases, we will also consider the momentum flux as [33]

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

where $\rho_0(z)$ is the reference density as a function of height. From linear theory, the analytic hydrostatic momentum flux is given as [33]

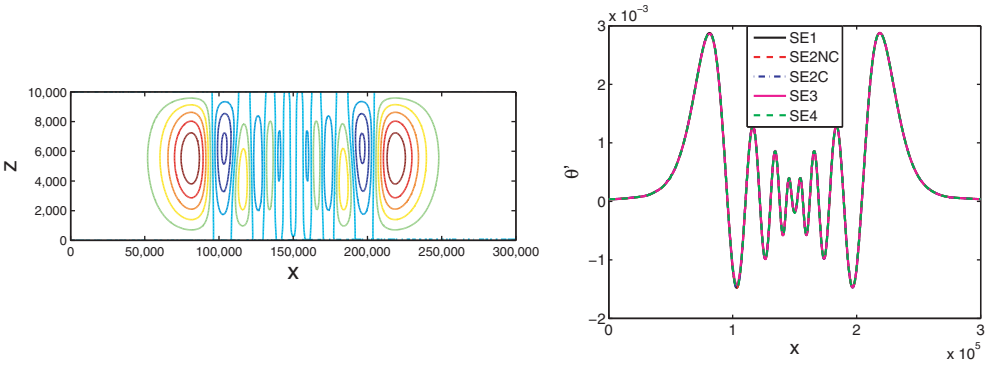
$$m^H(z) = -\frac{\pi_c}{4} \rho_s u_s \mathcal{N} h_c^2,$$

where the superscript H signifies *hydrostatic*, ρ_s and u_s are the reference density and horizontal velocity values at the surface, respectively, $\mathcal{N}$ 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 three test cases. We begin with the inertia-gravity waves, followed by the density current, and, finally, the linear hydrostatic mountain wave. Although each of the five semi-implicit models is derived from different equations (using different prognostic variables), it should be mentioned that the Schur forms 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.

We add 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 underresolved simulations, and the comparisons that we now report are representative of differences of the five models in both the well, and underresolved regimes; here we refer mostly to the conservation measures which do not change with varying resolutions.

4.1.1. Case 1: Inertia-gravity waves. Figure 4.1(a) shows the potential temperature perturbation contours after 3000 seconds, and Figure 4.1(b) shows the one-dimensional profile along z = 5,000 meters for all five models. Figure 4.1(b) shows



(a) (b)

FIG. 4.1. *Case* 1: *Inertia-gravity waves. Potential temperature perturbation after* 3000 *seconds for* 250*-meter resolution and tenth-order polynomials. Figure* (a) *shows the total domain using contour values between* -0.0015 *and* 0.003 *with a contour interval of* 0.0005 *for SE*2*NC, and Figure* (b) *shows the profiles along* 5,000*-meter height for all five models.*

TABLE 1

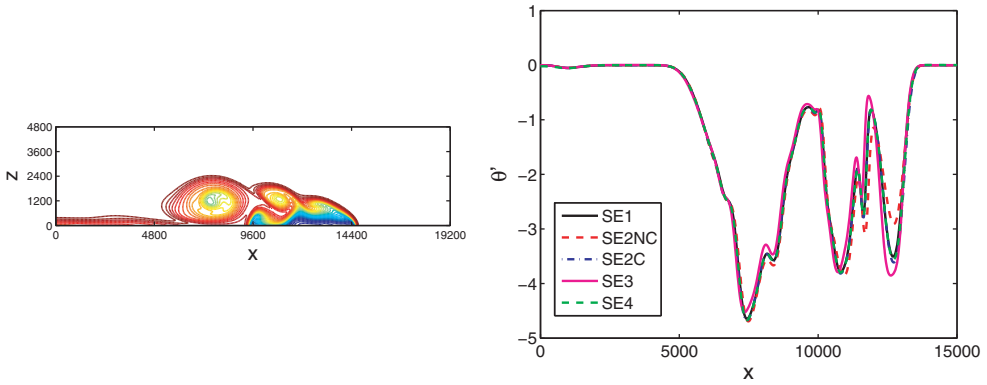
Case 1: *Inertia-gravity waves. Comparison of the five models studied for* 125*-meter resolution and tenth-order polynomials after* 2500 *seconds using* $\Delta t = 0.5$ *seconds* ($\mathcal{C}_{sound} = 3.15$ *and* $\mathcal{C}_{advect} = 0.17$).

	SE1	SE2NC	SE2C	SE3	SE4
Δ_{mass}	3.23×10^{-11}	4.89×10^{-13}	4.39×10^{-13}	3.36×10^{-13}	5.15×10^{-13}
Δ_{energy}	1.48×10^{-8}	1.61×10^{-13}	9.86×10^{-14}	1.10×10^{-13}	5.46×10^{-7}
GMRES Iterations	5	5	5	5	5
WallClock Time	2749	3051	3346	3238	2997

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.

Skamarock and Klemp [32] give an analytic solution for this test, but unfortunately it is valid only for the Boussinesq linearized problem that, while useful for qualitative comparisons, cannot be used to compute error norms since we use the fully compressible nonlinear equations. We use the same contouring interval used in [32], 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 $2.80 \times 10^{-3} \leq \theta' \leq -1.51 \times 10^{-3}$. In addition, by comparing our semi-implicit results to the results in [17] 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.

The main differences of interest among the five semi-implicit models are in the mass and energy conservation measures and in the efficiency (i.e., wallclock time) of the models (see Table 1). In terms of mass conservation, all models perform 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 surprisingly well, given the fact that it is not in strict conservation form. On



(a) (b)

FIG. 4.2. *Case* 2: *Density current. Potential temperature perturbations after* 900 *seconds with* 20*-meter resolution and tenth-order polynomials. Figure* (a) *shows the total domain using contour values between* -9 *to* 0 *with a contour interval of* 0.25 *for SE2NC. Figure* (b) *shows the profiles along* $z = 1200$ *meters for all five models.*

TABLE 2

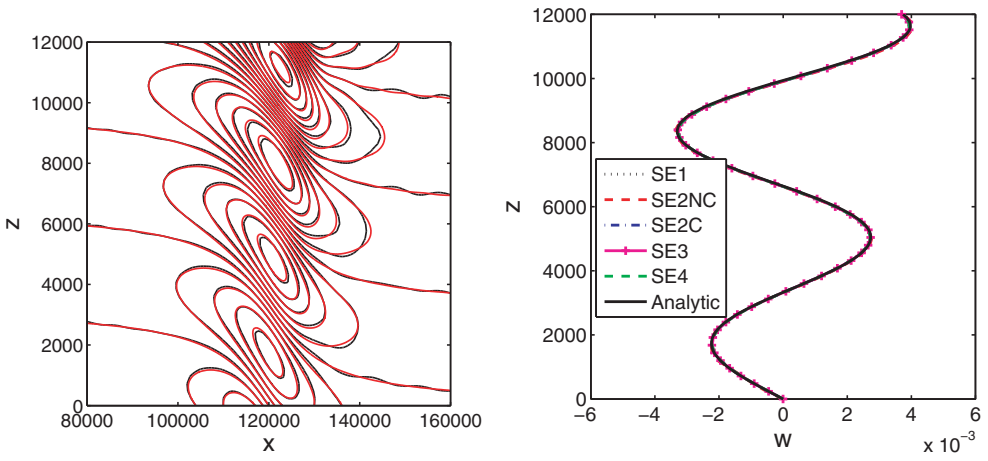
Case 2: *Density current. Comparison of the five models studied for* 20*-meter resolution and tenth-order polynomials after* 900 *seconds using* $\Delta t = 0.03125$ *seconds* ($\mathcal{C}_{sound} = 1.24$ *and* $\mathcal{C}_{advect} = 0.09$).

	SE1	SE2NC	SE2C	SE3	SE4
$\theta'_{\max}$	2.11×10^{-5}	3.76×10^{-6}	7.68×10^{-6}	3.30×10^{-3}	7.33×10^{-3}
$\theta'_{\min}$	-8.84	-8.70	-8.90	-9.09	-8.89
Δ_{mass}	2.33×10^{-5}	1.36×10^{-12}	1.78×10^{-13}	3.63×10^{-12}	3.39×10^{-14}
Δ_{energy}	2.34×10^{-4}	1.05×10^{-6}	1.92×10^{-5}	4.86×10^{-12}	7.65×10^{-4}
GMRES Iterations	2	2	2	2	2
WallClock Time	33673	37832	40244	38220	36885

the other hand, sets SE1 and SE4 are not expected to conserve energy at all, and they exhibit this weakness quite strongly. 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, and 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 nonconservation form (e.g., for the conservation forms, the flux tensor contains cubic terms instead of the quadratic terms found in the nonconseration form).

4.1.2. Case 2: Density current. In Figure 4.2(a) we plot the contours of potential temperature perturbation and in Figure 4.2(b) the one-dimensional profile of the potential temperature perturbation along $z = 1200$ meters for all five models. The three negative wells in Figure 4.2(b) correspond to the three distinct Kelvin–Helmholtz instability waves clearly visible in Figure 4.2(a). It is clear from Figure 4.2(b) that there are small differences among the five models. In order to discern the differences among the five models, let us now review Table 2.

While Table 2 shows that there is close agreement among all five models, it does show that the maximum and minimum values for potential temperature do vary. Re-



(a) (b)

FIG. 4.3. *Case* 3: *Linear hydrostatic mountain. Vertical velocity after* 30 *hours with* 1200*-meter (in x) and* 240*-meter (in z) resolution and tenth-order polynomials. Figure* (a) *shows contour values between* -0.005 *to* $+0.005$ *for the numerical (solid black line) and analytic (dashed line) solutions for SE2NC. Figure* (b) *shows the profiles along* $x = 120{,}000$ *meters for all five models and the analytic solution.*

call 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 paper by Straka et al. [35]; 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 of 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 operational atmospheric models (i.e., these operators are not consistent with the true Navier–Stokes viscous stresses).

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 with the corresponding 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, SE3, and SE2C. This ordering conforms to the number of operation counts because the number of GMRES iterations is identical for all of the models.

4.1.3. Case 3: Linear hydrostatic mountain. This case is different from the previous two in that (i) it has a steady-state analytic solution and (ii) it requires the implementation of NRBCs. The previous two test cases used either no-flux (reflecting) or periodic boundary conditions. Figure 4.3(a) shows that the numerical (black solid) and analytic (dashed) values for the vertical velocity compare very well. Figure 4.3(b) shows that the analytic and numerical vertical velocity solutions for all five models match identically; these values are sampled along the center of the mountain

TABLE 3

Case 3: *Linear hydrostatic mountain. Root-mean-square errors for the four variables for* 1200-*meter (in x) and* 240-*meter (in z) resolution and tenth-order polynomials for all five models using* $\Delta t = 1.5$ *seconds (Courant number* $= 1.25$*).*

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 4

Case 3: *Linear hydrostatic mountain. Comparison of the five models studied for* 1200-*meter (in x) and* 240-*meter (in z) resolution and tenth-order polynomials after* 30 *hours using* $\Delta t = 1.5$ *seconds (*$\mathcal{C}_{sound} = 1.25$ *and* $\mathcal{C}_{advect} = 0.07$*).*

	SE1	SE2NC	SE2C	SE3	SE4
Δ_{mass}	1.23×10^{-8}	1.98×10^{-9}	2.24×10^{-7}	1.21×10^{-8}	1.22×10^{-8}
Δ_{energy}	3.29×10^{-8}	3.47×10^{-9}	7.18×10^{-8}	9.05×10^{-8}	2.03×10^{-7}
GMRES Iterations	11	8	12	12	12
WallClock Time	5211	4731	6409	6121	6057

(x direction) for various values of z (vertical). Note that the actual computational domain is much larger than that shown in Figure 4.3(a). In fact, the domain shown in Figure 4.3(a) is the domain used to compute the root-mean-square errors; the portion of the domain not shown is, in fact, where the sponge layer is nonzero ($\beta > 0$).

In Table 3 we show the maximum and minimum values for all four variables for the five models for 10, 20, and 30 hours. The values for all five models are identical, clearly illustrating that all five models have converged to the identical steady-state solution. If we showed only the results after 30 hours, then one could argue that the reason why all of the models agree so closely is because they all converge to the same solution. However, the results in Table 3 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 NRBCs. 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.

Now turning to Table 4 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 in operational nonhydrostatic atmospheric models. To overcome the first problem (accuracy), we have begun work on the construction of high-order NRBCs that

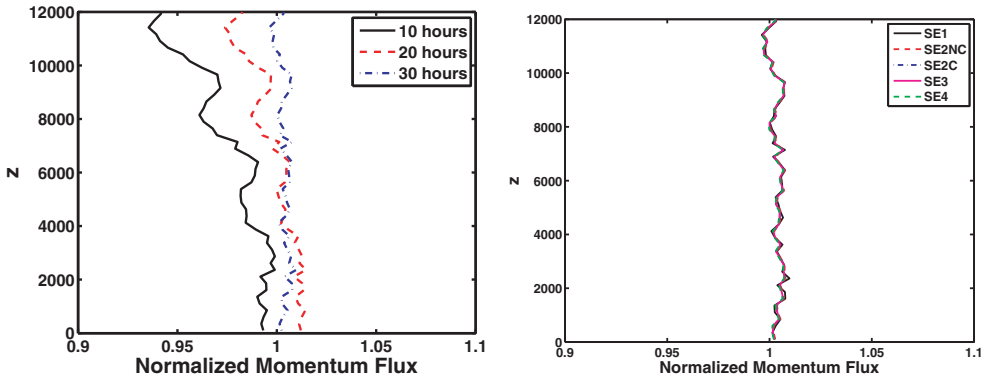


FIG. 4.4. *Case* 3: *Linear hydrostatic mountain. Normalized momentum flux for* 1200*-meter (in x) and* 240*-meter (in z) resolution and tenth-order polynomials for* (a) *SE*2*NC at times* 10*,* 20*, and* 30 *hours and* (b) *for all five models at* 30 *hours.*

can be used with high-order spatial and temporal discretizations (see [8, 28]), but we are still far 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 [19, 14] 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.

For this test, 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 8 GMRES iterations per time-step compared to 11 for SE1 and 12 for SE4 that then allows SE2NC to run faster; it is not obvious why SE2NC requires fewer GMRES iterations than the other models for this test.

In Figure 4.4(a) we plot the normalized momentum flux at various times in the integration, and in Figure 4.4(b) we show the normalized momentum flux for all five models after 30 hours. Figure 4.4(a) shows that the simulations have reached steady state after 30 hours since the normalized momentum flux achieves the theoretical value of one throughout the vertical column. Figure 4.4(b) shows that the normalized momentum flux values are essentially identical for all five models and are very good (compared to the linear solution defined in [33]).

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 [34] that we used previously with our explicit Navier–Stokes models [17]. 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 only SE2NC since it represents the median of all of the models in terms of efficiency and conservation measures.

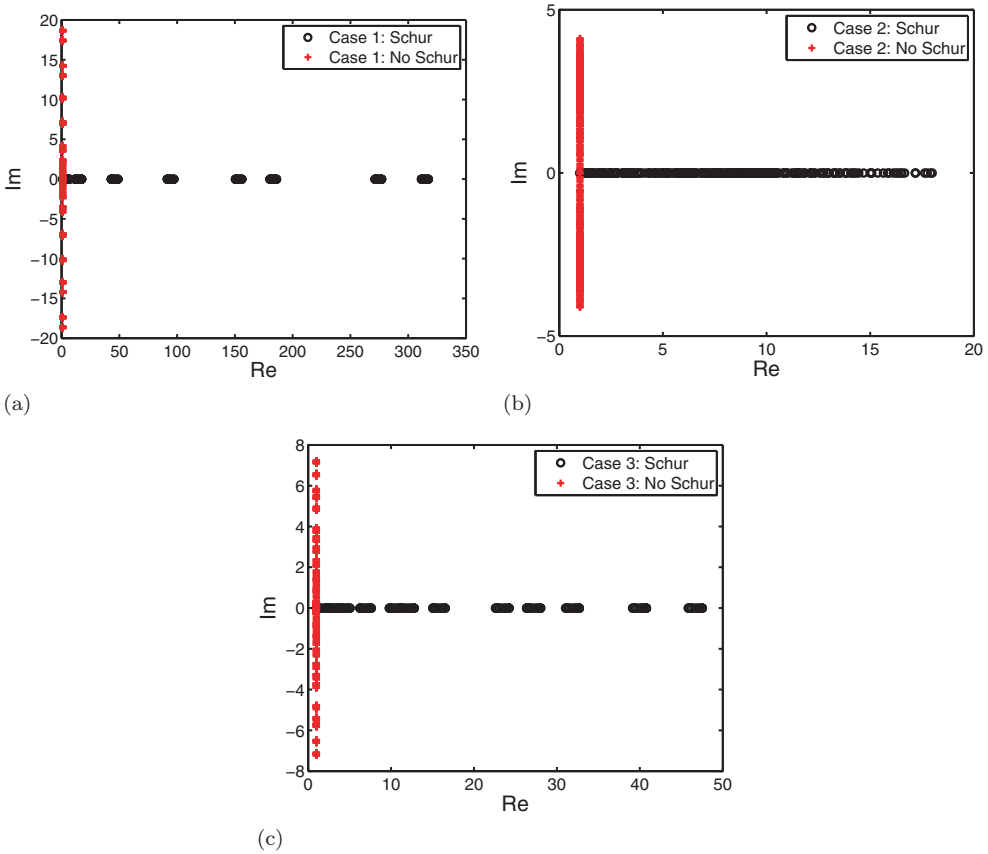


FIG. 4.5. *The eigenspectra of SE2NC for* (a) *case* 1 *with* $N_p = 297$, (b) *case* 2 *with* $N_p = 297$, *and* (c) *case* 3 *with* $N_p = 325$ *for the Schur (black circles along horizontal) and No Schur (crosses along vertical) forms of the semi-implicit method.*

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 (A.1)–(A.4) which, assuming N_P grid points, represents a $16N_p^2$ matrix problem. In contrast, the Schur form is defined by (A.10) and represents an 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 different eigenspectra. To get a sense of this difference, we show the eigenspectra for the No Schur and Schur forms in Figure 4.5 for all three test cases. For case 3, 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) discrete differential operators. While we will not address this here, this information gives insight into how to construct optimal preconditioners.

In Figures 4.6, 4.7, and 4.8, 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

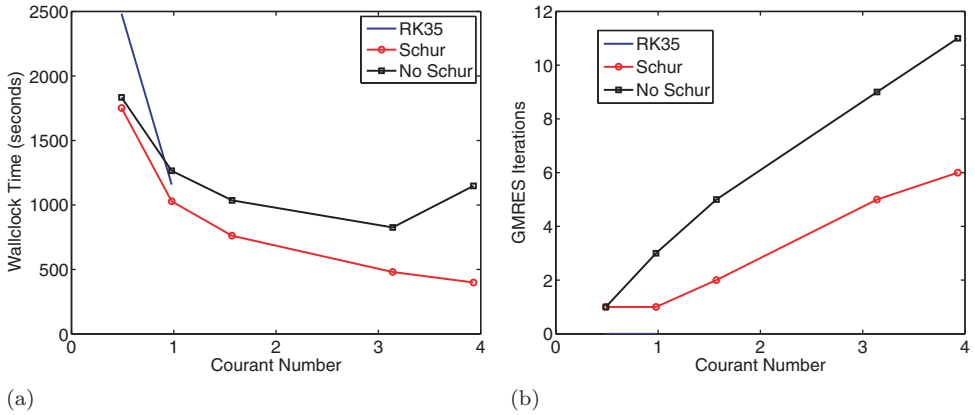


FIG. 4.6. *Case* 1: *Inertia-gravity waves. The* (a) *wallclock time and* (b) *number of GMRES iterations as functions of the Courant number for* 250*-meter resolution with tenth-order polynomials after* 3000 *seconds. The explicit Runge–Kutta method (RK*35*) is compared with the semi-implicit methods with and without the Schur complements (Schur and No Schur, respectively).*

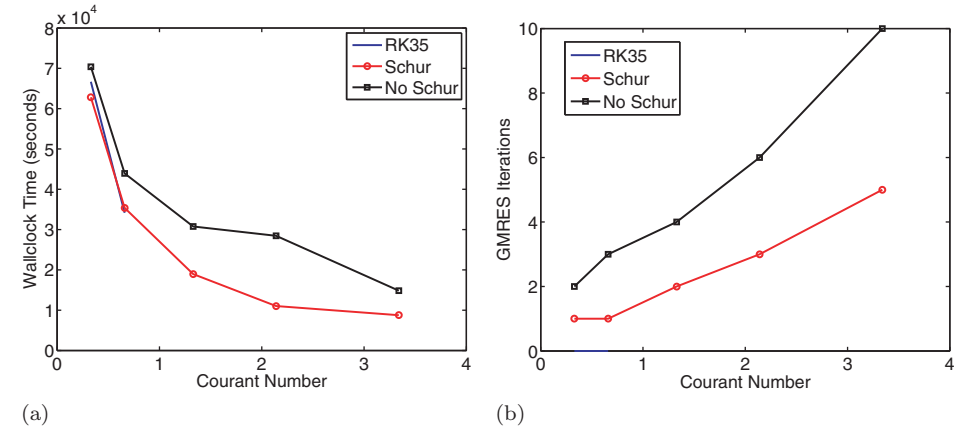


FIG. 4.7. *Case* 2: *Density current. The* (a) *wallclock time and* (b) *number of GMRES iterations as functions of the Courant number for* 20*-meter resolution with tenth-order polynomials after* 900 *seconds. The explicit Runge–Kutta method (RK*35*) is compared with the semi-implicit methods with and without the Schur complements (Schur and No Schur, respectively).*

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.

4.2.1. Case 1: Inertia-gravity waves. Figure 4.6 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 Figure 4.6(a) 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

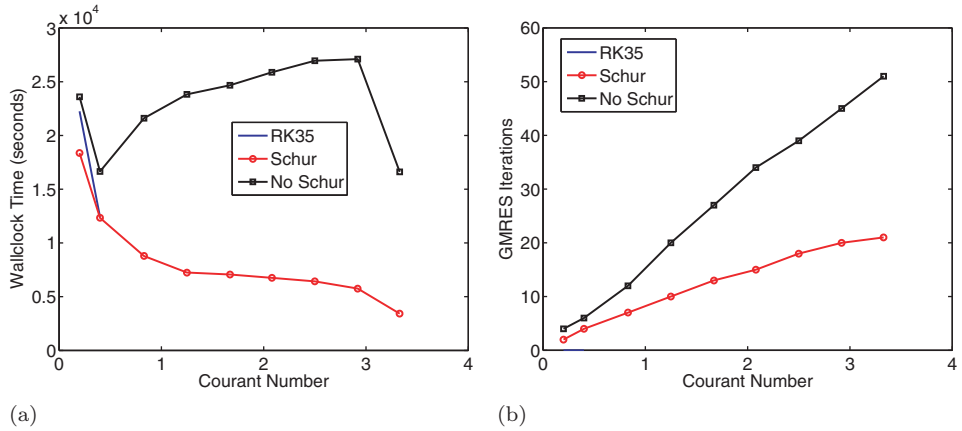


FIG. 4.8. *Case* 3: *Linear hydrostatic mountain. The* (a) *wallclock time and* (b) *number of GMRES iterations as functions of the Courant number for* 1200*-meter (in* x*) and* 240*-meter (in* z*) resolution with tenth-order polynomials after* 30 *hours. The explicit Runge–Kutta method (RK*35*) is compared with the semi-implicit methods with and without the Schur complements Schur and No Schur, respectively).*

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 the average number of GMRES iterations required per time-step. Figure 4.6(b) 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).

4.2.2. Case 2: Density current. In Figure 4.7(a) (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 true for the Schur form for all the cases we studied. (We do not include a discussion on preconditioning, but this is important in maintaining the efficiency of the semi-implicit method; otherwise, at some time-step value, the number of GMRES iterations will become too large, and the method loses its efficiency gains.) In the previous test, 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 one. (The maximum Courant numbers are found empirically.) 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 4.7(b) (right panel) shows that the number of GMRES iterations increases at an accelerated rate for the No Schur form but increases only 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.

4.2.3. Case 3: Linear hydrostatic mountain. In Figure 4.8(a) (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 4.8(b) (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 has already climbed to 20 and continues 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 are offset by a larger implicit solver iteration count. In contrast, the Schur form exhibits efficiency gains for increasing Courant number, even though 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. The difference between this test and the others is that NRBCs are employed. It should be noted that this test is very typical of the class of problems that must be run efficiently in operational-type 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. Condition for extracting the Schur complement. Our aim in this section is not to provide a classical stability analysis of the semi-implicit method (e.g., see [2] for such an analysis) but rather to show that, although we use five different variations of the governing equations, they all satisfy the identical condition for extracting the Schur complement.

To show that the semi-implicit form of all five equation sets is stable requires going back to the original time-integration statement of the problem, that is,

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

Recall that in an IMEX approach, we treat the nonlinear terms $N(\boldsymbol{q})$ explicitly and the linear terms $L(\boldsymbol{q})$ implicitly. 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 contains only the advective waves that are assumed to be slower than the acoustic and gravity waves (since we are assuming that the flow remains subsonic).

Thus, in order to maintain stability, we require that the Courant number associated with the advective waves satisfies the CFL condition of standard explicit time-integrators (in this case, we are using the explicit BDF2; see [23]). 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 $\mathcal{Z} = \lambda \Delta t$ in the left-hand plane, where λ is the eigenvalue of the right-hand-side operator that includes every term in the equations

except for the time-derivative and Δt is the time-step) and L-stable (the amplification function goes to zero for $\mathcal{Z} \to -\infty$). This means that as long as we adhere to a *modified* 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., [11] where it is shown how the implicit linear part modifies the explicit stability region of IMEX-BDF2). For the Schur form, we have to perform further analysis.

Assuming that we are adhering to the (modified) explicit CFL limit of the slow-moving waves, the only possibility for instabilities to occur 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 (3.12). This is also true for SE2NC (see (A.7)), SE2C (see (B.9)), SE3 (see (C.9)), and SE4 (see (D.6)). 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 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 $T_0 = \theta_0 \pi_0$ and using the definition of hydrostatic balance $\frac{d\pi_0}{dz} = -\frac{g}{c_p \theta_0}$ allows us to write the above expression as

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

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

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

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 (4.1), yields

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

which is identical to the expression for SE1, SE2NC, SE2C, and SE3; thus c_4 can never be zero. Therefore, as one would expect, the stability condition for all the models is 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 nonsingular, but the proof of stability is more complicated.

5. Conclusions. We have presented semi-implicit formulations of five different forms of the compressible NSE used in nonhydrostatic atmospheric modeling. These equations have typically been solved either explicitly or semi-implicitly along only the vertical direction; the only exceptions are the three models mentioned in the Introduction, where two are global models and the other is a spectral model. 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 of 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 able to show only in the special case when either reflecting (no-flux) or periodic boundary conditions were used. The true advantage of the semi-implicit formulation can be realized only 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) 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 are not as efficient as those that are in nonconservation 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. It should be mentioned that, for equations that are not strictly energy conserving, slight modifications can be made to make them so (e.g., writing the advection terms in rotation form). Furthermore, the currently used NRBCs adversely affect both accuracy and conservation, which motivates the need for better NRBCs 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 that 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. For the two-dimensional Navier–Stokes equations, the No Schur form is 16 times larger than the Schur form; in three dimensions, this increases to a factor of 25. 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 varies 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 IMEX Runge–Kutta methods and fully implicit 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.

Appendix A. 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

$$\rho_{tt} = \alpha\left(\widehat{\rho} - \lambda w_{tt}\frac{d\rho_0}{dz} - \lambda\rho_0\boldsymbol{\nabla}\cdot\boldsymbol{u}_{tt}\right) + \beta\widehat{\rho}_b, \tag{A.1}$$

$$\boldsymbol{u}_{tt} = \alpha\left(\widehat{\boldsymbol{u}} - \lambda\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} - \lambda g\frac{\rho_{tt}}{\rho_0}\boldsymbol{k}\right) + \beta\widehat{\boldsymbol{u}}_b, \tag{A.2}$$

$$\theta_{tt} = \alpha\left(\widehat{\theta} - \lambda w_{tt}\frac{d\theta_0}{dz}\right) + \beta\widehat{\theta}_b, \tag{A.3}$$

$$P_{tt} = G_0\rho_{tt} + H_0\theta_{tt}, \tag{A.4}$$

where $G_0 = \frac{\gamma P_0}{\rho_0}$ and $H_0 = \frac{\gamma P_0}{\theta_0}$; the system represented by (A.1)–(A.4) is the No

Schur form of SE2NC. Substituting (A.3) into (A.4) yields

$$\rho_{tt} = \frac{1}{G_0}\left[P_{tt} - H_0\alpha\left(\widehat{\theta} - \lambda w_{tt}\frac{d\theta_0}{dz}\right) - H_0\beta\widehat{\theta}_b\right]. \tag{A.5}$$

We can now substitute (A.5) into (A.2) in order to express the momentum as a function only of pressure. Upon applying this substitution, we get

$$\boldsymbol{u}_{tt} = \boldsymbol{P}_{2NC}\left[(\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda\frac{gH_0}{\rho_0 G_0}\left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b\right)\boldsymbol{k} - \alpha\lambda\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} - \alpha\lambda\frac{g}{\rho_0 G_0}P_{tt}\boldsymbol{k}\right], \tag{A.6}$$

where

$$\boldsymbol{C}_{2NC} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_{2NC}} \end{pmatrix} \tag{A.7}$$

with

$$c_{2NC} = 1 + (\alpha\lambda)^2\frac{g}{\theta_0}\frac{d\theta_0}{dz}, \tag{A.8}$$

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 (A.1) and (A.3) into (A.4) yields

$$P_{tt} = G_0\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right) + H_0\left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b\right) - \alpha\lambda F_0 w_{tt} - \alpha\lambda\rho_0 G_0\boldsymbol{\nabla}\cdot\boldsymbol{u}_{tt} \tag{A.9}$$

where $F_0 = G_0\frac{d\rho_0}{dz} + H_0\frac{d\theta_0}{dz}$. The last step is to substitute (A.6) into (A.9) that yields

$$\begin{aligned}
&P_{tt} - (\alpha\lambda)^2 F_0\boldsymbol{k}\cdot\left[\boldsymbol{P}_{2NC}\left(\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} + \frac{g}{\rho_0 G_0}P_{tt}\boldsymbol{k}\right)\right] \\
&\quad - (\alpha\lambda)^2 G_0\rho_0\boldsymbol{\nabla}\cdot\left[\boldsymbol{P}_{2NC}\left(\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} + \frac{g}{\rho_0 G_0}P_{tt}\boldsymbol{k}\right)\right] \\
&= G_0\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right) + H_0\left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b\right) \\
&\quad - \alpha\lambda F_0\boldsymbol{k}\cdot\left[\boldsymbol{P}_{2NC}\left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda\frac{gH_0}{\rho_0 G_0}\left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b\right)\boldsymbol{k}\right)\right] \\
&\quad - \alpha\lambda G_0\rho_0\boldsymbol{\nabla}\cdot\left[\boldsymbol{P}_{2NC}\left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda\frac{gH_0}{\rho_0 G_0}\left(\alpha\widehat{\theta} + \beta\widehat{\theta}_b\right)\boldsymbol{k}\right)\right]
\end{aligned} \tag{A.10}$$

which is a pseudo-Helmholtz equation for P_{tt} and is the Schur form of SE2NC.

Appendix B. 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

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

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

$$\rho_{tt} = \alpha\left(\widehat{\rho} - \lambda \boldsymbol{\nabla}\cdot \boldsymbol{U}_{tt}\right) + \beta\widehat{\rho}_b, \tag{B.1}$$

$$\boldsymbol{U}_{tt} = \alpha\left(\widehat{\boldsymbol{U}} - \lambda\boldsymbol{\nabla} P_{tt} - \lambda g \rho_{tt}\boldsymbol{k}\right) + \beta\widehat{\boldsymbol{U}}_b, \tag{B.2}$$

$$\Theta_{tt} = \alpha\left(\widehat{\Theta} - \lambda\boldsymbol{\nabla}\cdot\left(G_0\boldsymbol{U}_{tt}\right)\right) + \beta\widehat{\Theta}_b, \tag{B.3}$$

$$P_{tt} = F_0\Theta_{tt}. \tag{B.4}$$

Equations (B.1)–(B.4) represent the full system of SE2C (i.e., the No Schur form). Let us now derive the Schur form.

Let us first substitute (B.3) into (B.4) to get

$$P_{tt} = F_0\alpha\left(\widehat{\Theta} - \lambda\boldsymbol{\nabla}\cdot\left(G_0\boldsymbol{U}_{tt}\right)\right) + F_0\beta\widehat{\Theta}_b. \tag{B.5}$$

Multiplying (B.1) by G_0 and subtracting from (B.3) to eliminate the term $G_0\boldsymbol{\nabla}\cdot\boldsymbol{U}_{tt}$ yields

$$\Theta_{tt} - G_0\rho_{tt} = \left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b\right) - G_0\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right) - \alpha\lambda W_{tt}\frac{dG_0}{dz}. \tag{B.6}$$

Substituting (B.4) into (B.6), to eliminate Θ_{tt}, gives

$$\rho_{tt} = \frac{1}{F_0G_0}P_{tt} + \alpha\lambda W_{tt}\frac{1}{G_0}\frac{dG_0}{dz} - \frac{1}{G_0}\left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b\right) + \left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right). \tag{B.7}$$

Note that substituting (B.7) into (B.2) allows us to solve for $\boldsymbol{U}_{tt}$ as a function of P_{tt} such as

$$\begin{aligned}\boldsymbol{U}_{tt} = \boldsymbol{P}_{2C}\Bigg[&\left(\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b\right) - \alpha\lambda\boldsymbol{\nabla}P_{tt} - \alpha\lambda\frac{g}{F_0G_0}P_{tt}\boldsymbol{k}\\ &- \alpha\lambda g\boldsymbol{k}\left(\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right) - \frac{1}{G_0}\left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b\right)\right)\Bigg],\end{aligned} \tag{B.8}$$

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

$$\boldsymbol{C}_{2C} = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_{2C}} \end{pmatrix}, \tag{B.9}$$

and $c_2 = 1 + (\alpha\lambda)^2\frac{g}{G_0}\frac{dG_0}{dz}$, where $\frac{1}{G_0}\frac{dG_0}{dz} = (\frac{1}{\Theta_0}\frac{d\Theta_0}{dz} - \frac{1}{\rho_0}\frac{d\rho_0}{dz})$. Finally, substituting (B.8) into (B.5) yields

$$\begin{aligned}&P_{tt} - (\alpha\lambda)^2F_0\boldsymbol{\nabla}G_0\cdot\left[\boldsymbol{P}_{2C}\left(\boldsymbol{\nabla}P_{tt} + \frac{g}{F_0G_0}P_{tt}\boldsymbol{k}\right)\right]\\ &\quad - (\alpha\lambda)^2F_0G_0\boldsymbol{\nabla}\cdot\left[\boldsymbol{P}_{2C}\left(\boldsymbol{\nabla}P_{tt} + \frac{g}{F_0G_0}P_{tt}\boldsymbol{k}\right)\right] = F_0\left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b\right)\\ &-\alpha\lambda F_0\boldsymbol{\nabla}\cdot\left[G_0\boldsymbol{P}_{2C}\left(\left(\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b\right) - \alpha\lambda g\boldsymbol{k}\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right) + \alpha\lambda\frac{g}{G_0}\boldsymbol{k}\left(\alpha\widehat{\Theta} + \beta\widehat{\Theta}_b\right)\right)\right],\end{aligned}$$

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

Appendix C. 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

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

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, respectively. Upon applying the semi-implicit method to SE3, we arrive at the following semi-discrete problem:

$$\rho_{tt} = \alpha\left(\widehat{\rho} - \lambda\boldsymbol{\nabla}\cdot\boldsymbol{U}_{tt}\right) + \beta\widehat{\rho}_b, \tag{C.1}$$

$$\boldsymbol{U}_{tt} = \alpha\left(\widehat{\boldsymbol{U}} - \lambda\boldsymbol{\nabla}P_{tt} - \lambda\rho_{tt}g\boldsymbol{k}\right) + \beta\widehat{\boldsymbol{U}}_b, \tag{C.2}$$

$$E_{tt} = \alpha\left(\widehat{E} - \lambda\boldsymbol{\nabla}\cdot(h_0\boldsymbol{U}_{tt})\right) + \beta\widehat{E}_b, \tag{C.3}$$

$$P_{tt} = (\gamma-1)\left(E_{tt} - \phi\rho_{tt}\right). \tag{C.4}$$

The system represented by (C.1)–(C.4) is the No Schur form of SE3. Let us now derive the Schur form.

Substituting (C.1) and (C.3) into (C.4) yields

$$\begin{aligned} P_{tt} =& (\gamma-1)\left[(\alpha\widehat{E} + \beta\widehat{E}_b) - \alpha\lambda h_0\boldsymbol{\nabla}\cdot\boldsymbol{U}_{tt} - \alpha\lambda\boldsymbol{\nabla}h_0\cdot\boldsymbol{U}_{tt}\right] \\ & - \phi(\gamma-1)\left[(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) - \alpha\lambda\boldsymbol{\nabla}\cdot\boldsymbol{U}_{tt}\right]. \end{aligned} \tag{C.5}$$

Multiplying (C.1) by h_0 and subtracting from (C.3) to eliminate the term $h_0\boldsymbol{\nabla}\cdot\boldsymbol{U}_{tt}$ yields

$$E_{tt} - h_0\rho_{tt} = (\alpha\widehat{E} + \beta\widehat{E}_b) - h_0(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) - \alpha\lambda W_{tt}\frac{dh_0}{dz}. \tag{C.6}$$

Next, substituting (C.6) into (C.4) to eliminate E_{tt} and rearranging gives

$$\rho_{tt} = \frac{1}{h_0-\phi}\left[\frac{1}{(\gamma-1)}P_{tt} + \alpha\lambda W_{tt}\frac{dh_0}{dz} - (\alpha\widehat{E} + \beta\widehat{E}_b) + h_0(\alpha\widehat{\rho} + \beta\widehat{\rho}_b)\right], \tag{C.7}$$

which can now be substituted into (C.2) and solved for $\boldsymbol{U}_{tt}$ to yield

$$\begin{aligned} \boldsymbol{U}_{tt} = \boldsymbol{P}_3\Bigg[&(\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b) - \alpha\lambda\boldsymbol{\nabla}P_{tt} - \alpha\lambda\frac{g}{(\gamma-1)(h_0-\phi)}P_{tt}\boldsymbol{k} \\ & - \alpha\lambda\frac{g}{h_0-\phi}\left(h_0(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) - (\alpha\widehat{E} + \beta\widehat{E}_b)\right)\boldsymbol{k}\Bigg], \end{aligned} \tag{C.8}$$

where

$$\boldsymbol{C}_3 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_3} \end{pmatrix} \tag{C.9}$$

with

$$c_3 = 1 + (\alpha\lambda)^2 \frac{g}{h_0 - \phi} \frac{dh_0}{dz}. \tag{C.10}$$

Finally, substituting (C.8) into (C.5) yields

$$\begin{aligned}
&P_{tt} - (\alpha\lambda)^2(\gamma-1)(h_0-\phi)\boldsymbol{\nabla}\cdot\left[\boldsymbol{P}_3\left(\boldsymbol{\nabla}P_{tt} + \frac{g}{(\gamma-1)(h_0-\phi)}P_{tt}\boldsymbol{k}\right)\right] \\
&\quad - (\alpha\lambda)^2(\gamma-1)\boldsymbol{\nabla}h_0\cdot\left[\boldsymbol{P}_3\left(\boldsymbol{\nabla}P_{tt} + \frac{g}{(\gamma-1)(h_0-\phi)}P_{tt}\boldsymbol{k}\right)\right] \\
&= \quad (\gamma-1)\left[(\alpha\widehat{E} + \beta\widehat{E}_b) - \phi(\alpha\widehat{\rho} + \beta\widehat{\rho}_b)\right] \\
&\quad - \alpha\lambda(\gamma-1)(h_0-\phi)\boldsymbol{\nabla}\cdot\left[\boldsymbol{P}_3\left((\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b) - \alpha\lambda\frac{gh_0}{h_0-\phi}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b)\boldsymbol{k}\right.\right. \\
&\qquad\qquad \left.\left. + \alpha\lambda\frac{g}{h_0-\phi}(\alpha\widehat{E} + \beta\widehat{E}_b)\boldsymbol{k}\right)\right] \\
&\quad - \alpha\lambda(\gamma-1)\boldsymbol{\nabla}h_0\cdot\left[\boldsymbol{P}_3\left((\alpha\widehat{\boldsymbol{U}} + \beta\widehat{\boldsymbol{U}}_b) - \alpha\lambda\frac{gh_0}{h_0-\phi}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b)\boldsymbol{k}\right.\right. \\
&\qquad\qquad \left.\left. + \alpha\lambda\frac{g}{h_0-\phi}(\alpha\widehat{E} + \beta\widehat{E}_b)\boldsymbol{k}\right)\right],
\end{aligned}$$

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

Appendix D. 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:

$$\rho_{tt} = \alpha\left(\widehat{\rho} - \lambda w_{tt}\frac{d\rho_0}{dz} - \lambda\rho_0\boldsymbol{\nabla}\cdot\boldsymbol{u}_{tt}\right) + \beta\widehat{\rho}_b, \tag{D.1}$$

$$\boldsymbol{u}_{tt} = \alpha\left(\widehat{\boldsymbol{u}} - \lambda\frac{1}{\rho_0}\boldsymbol{\nabla}P_{tt} - \lambda g\frac{\rho_{tt}}{\rho_0}\boldsymbol{k}\right) + \beta\widehat{\boldsymbol{u}}_b, \tag{D.2}$$

$$P_{tt} = \alpha\left(\widehat{P} - \lambda w_{tt}\frac{dP_0}{dz} - \lambda\gamma P_0\boldsymbol{\nabla}\cdot\boldsymbol{u}_{tt}\right) + \beta\widehat{P}_b. \tag{D.3}$$

The system described by (D.1)–(D.3) is the No Schur form of SE4. Let us now derive the Schur form.

Multiplying (D.1) by γP_0, subtracting (D.3) multiplied by ρ_0, and rearranging yields

$$\rho_{tt} = \frac{1}{\gamma P_0}\left[\rho_0 P_{tt} - \rho_0\left(\alpha\widehat{P} + \beta\widehat{P}_b\right) + \gamma P_0\left(\alpha\widehat{\rho} + \beta\widehat{\rho}_b\right) + \alpha\lambda w_{tt}\left(\rho_0\frac{dP_0}{dz} - \gamma P_0\frac{d\rho_0}{dz}\right)\right]. \tag{D.4}$$

Substituting (D.4) into (D.2) yields

$$\boldsymbol{u}_{tt} = \boldsymbol{P}_4 \Bigg[(\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda g\boldsymbol{k} \left(\frac{1}{\gamma P_0}(\alpha\widehat{P} + \beta\widehat{P}_b) - \frac{1}{\rho_0}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right) - \alpha\lambda \frac{1}{\rho_0}\boldsymbol{\nabla} P_{tt} - \alpha\lambda \frac{g}{\gamma P_0} P_{tt}\boldsymbol{k} \Bigg], \tag{D.5}$$

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

$$\boldsymbol{C}_4 = \begin{pmatrix} 1 & 0 \\ 0 & \frac{1}{c_4} \end{pmatrix} \tag{D.6}$$

with

$$c_4 = 1 + (\alpha\lambda)^2 g \left(\frac{1}{\gamma P_0}\frac{dP_0}{dz} - \frac{1}{\rho_0}\frac{d\rho_0}{dz} \right). \tag{D.7}$$

Substituting (D.5) into (D.3) yields

$$\begin{aligned} & P_{tt} - (\alpha\lambda)^2 \frac{dP_0}{dz}\boldsymbol{k} \cdot \left[\boldsymbol{P}_4 \left(\frac{1}{\rho_0}\boldsymbol{\nabla} P_{tt} + \frac{g}{\gamma P_0} P_{tt}\boldsymbol{k} \right) \right] - (\alpha\lambda)^2 \gamma P_0 \boldsymbol{\nabla} \\ & \cdot \left[\boldsymbol{P}_4 \left(\frac{1}{\rho_0}\boldsymbol{\nabla} P_{tt} + \frac{g}{\gamma P_0} P_{tt}\boldsymbol{k} \right) \right] \\ & = (\alpha\widehat{P} + \beta\widehat{P}_b) - \alpha\lambda \frac{dP_0}{dz}\boldsymbol{k} \cdot \Bigg[\boldsymbol{P}_4 \Bigg((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) \\ & \qquad + \alpha\lambda g\boldsymbol{k} \left(\frac{1}{\gamma P_0}(\alpha\widehat{P} + \beta\widehat{P}_b) - \frac{1}{\rho_0}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right) \Bigg) \Bigg] \\ & - \alpha\lambda\gamma P_0 \boldsymbol{\nabla} \cdot \left[\boldsymbol{P}_4 \left((\alpha\widehat{\boldsymbol{u}} + \beta\widehat{\boldsymbol{u}}_b) + \alpha\lambda g\boldsymbol{k} \left(\frac{1}{\gamma P_0}(\alpha\widehat{P} + \beta\widehat{P}_b) - \frac{1}{\rho_0}(\alpha\widehat{\rho} + \beta\widehat{\rho}_b) \right) \right) \right], \end{aligned}$$

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

REFERENCES

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

[2] P. Benard, *Stability of semi-implicit and iterative centered-implicit time discretizations for various equation systems used in NWP*, Monthly Weather Rev., 131 (2000), pp. 2479–2491.

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

[4] D.M. Burridge, *A split semi-implict reformulation of the Bushby–Timpson* 10*-level model*, Quart. J. Roy. Meteorol. Soc., 101 (1975), pp. 777–792.

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

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

[7] T. Davies, M.J.P. Cullen, A.J. Malcolm, M.H. Mawson, A. Staniforth, A.A. White, and N. Wood, *A new dynamical core for the Met Office's global and regional modelling of the atmosphere*, Quart. J. Roy. Meteorol. Soc., 131 (2005), pp. 1759–1782.

[8] 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.

[9] V. Ducrocq, F. Bouttier, S. Malardel, T. Montmerle, and Y. Seity, *Le projet arome*, La Houille Blanche, 2005, pp. 39–43 (in French).
[10] P.F. Fischer, *Projection techniques for iterative solution of $Ax = b$ with successive right-hand sides*, Comput. Methods Appl. Mech. Engrg., 163 (1998), pp. 193–204.
[11] J. Frank, W. Hundsdorfer, and J.G. Verwer, *On the stability of implicit-explicit linear multistep methods*, Appl. Numer. Math., 25 (1997), pp. 193–205.
[12] T. Galchen and R.C.J. Somerville, *Use of a coordinate transformation for solution of Navier–Stokes equations*, J. Comput. Phys., 17 (1975), pp. 209–228.
[13] F.X. Giraldo, *The Lagrange–Galerkin spectral element method on unstructured quadrilateral grids*, J. Comput. Phys., 147 (1998), pp. 114–146.
[14] F.X. Giraldo, *Semi-implicit time-integrators for a scalable spectral element atmospheric model*, Quart. J. Roy. Meteorol. Soc., 131 (2005), pp. 2431–2454.
[15] F.X. Giraldo, *High-order triangle-based discontinuous Galerkin methods for hyperbolic equations on a rotating sphere*, J. Comput. Phys., 214 (2006), pp. 447–465.
[16] F.X. Giraldo, *Hybrid Eulerian–Lagrangian semi-implicit time-integrators*, Comput. Math. Appl., 52 (2006), pp. 1325–1342.
[17] 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. Comput. Phys., 227 (2008), pp. 3849–3877.
[18] F.X. Giraldo and M. Restelli, *High-order semi-implicit time-integrators for a triangular discontinuous Galerkin oceanic shallow water model*, Internat. J. Numer. Methods Fluids, 63 (2010), pp. 1077–1102.
[19] F.X. Giraldo and T.E. Rosmond, *A scalable spectral element Eulerian atmospheric model (SEE-AM) for NWP: Dynamical core tests*, Monthly Weather Rev., 132 (2004), pp. 133–153.
[20] F.X. Giraldo and T. Warburton, *A high-order triangular discontinuous Galerkin oceanic shallow water model*, Internat. J. Numer. Methods Fluids, 56 (2008), pp. 899–925.
[21] R.M. Hodur, *The Naval Research Laboratory's coupled ocean/atmosphere mesoscale prediction system (COAMPS)*, Monthly Weather Rev., 125 (1997), pp. 1414–1430.
[22] Z.I. Janjic, *A nonhydrostatic model based on a new approach*, Meteorol. Atmos. Phys., 82 (2003), pp. 271–285.
[23] G.E. Karniadakis, M. Israeli, and S.A. Orszag, *High-order splitting methods for the incompressible Navier–Stokes equations*, J. Comput. Phys., 97 (1991), pp. 414–443.
[24] J.B. Klemp, W.C. Skamarock, and J. Dudhia, *Conservative split-explicit time integration methods for the compressible nonhydrostatic equations*, Monthly Weather Rev., 135 (2007), pp. 2897–2913.
[25] J.B. Klemp and R.B. Wilhelmson, *Simulation of* 3*-dimensional convective storm dynamics*, J. Atmos. Sci., 35 (1978), pp. 1070–1096.
[26] M. Kwizak and A.J. Robert, *A semi–implicit scheme for grid point atmospheric models of the primitive equations*, Monthly Weather Rev., 99 (1971), pp. 32–36.
[27] M. Läuter, F.X. Giraldo, D. Handorf, and K. Dethloff, *A discontinuous Galerkin method for the shallow water equations in spherical triangular coordinates*, J. Comput. Phys., 227 (2008), pp. 10226–10242.
[28] 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, 47 (2010), pp. 289–298.
[29] M. Restelli and F.X. Giraldo, *A conservative discontinuous Galerkin semi-implicit formulation for the Navier–Stokes equations in nonhydrostatic mesoscale modeling*, SIAM J. Sci. Comput., 31 (2009), pp. 2231–2257.
[30] K. Saito, J. Ishida, K. Aranami, T. Hara, T. Segawa, M. Narita, and Y. Honda, *Nonhydrostatic atmospheric models and operational development at JMA*, J. Meteorol. Soc. Japan, 85B (2007), pp. 271–304.
[31] U. Schattler, G. Doms, and J. Steppeler, *Requirements and problems in parallel model development at DWD*, Scientific Programming, 8 (2000), pp. 13–22.
[32] W.C. Skamarock and J.B. Klemp, *Efficiency and accuracy of the Klemp–Wilhelmson time-splitting technique*, Monthly Weather Rev., 122 (1994), pp. 2623–2630.
[33] R.B. Smith, *The influence of mountains on the atmosphere*, Adv. Geophys., 221 (1979), pp. 87–230.
[34] R.J. Spiteri and S.J. Ruuth, *A new class of optimal high-order strong-stability-preserving time discretization methods*, SIAM J. Numer. Anal., 40 (2002), pp. 469–491.
[35] 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*, Internat. J. Numer. Methods Fluids, 17 (1993), pp. 1–22.

[36] M. Tanguay, A. Robert, and R. Laprise, *A semi-implicit semi-Lagrangian fully compressible regional forecast model*, Monthly Weather Rev., 118 (1990), pp. 1970–1980.
[37] M.C. Tapp and P.W. White, *A non-hydrostatic mesoscale model*, Quart. J. Roy. Meteorol. Soc., 102 (1976), pp. 277–296.
[38] S. Thomas, C. Girard, G. Doms, and U. Schattler, *Semi-implicit scheme for the DWD Lokal-Modell*, Meteorol. Atmos. Phys., 73 (2000), pp. 105–125.
[39] M. Xue, K.K. Droegemeier, and V. Wong, *The advanced regional prediction system (ARPS)—A multiscale nonhydrostatic atmosphere simulation and prediction model. Part* I: *Model dynamics and verification*, Meteorology and Atmospheric Physics, 75 (2000), pp. 161–193.
[40] K.S. Yeh, J. Cote, S. Gravel, A. Methot, A. Patoine, M. Roch, and A. Staniforth, *The CMC-MRB global environmental multiscale (GEM) model. Part* III: *Nonhydrostatic formulation*, Monthly Weather Rev., 130 (2002), pp. 339–356.