

Exact spatial and temporal balance of energy exchanges within a horizontally explicit/vertically implicit non-hydrostatic atmosphere

David Lee^{a,*}, Artur Palha^b

^a*Bureau of Meteorology, Melbourne, Australia*

^b*Faculty of Aerospace Engineering, Delft University of Technology, Kluyverweg 2, 2629 HS Delft, The Netherlands*

Abstract

A new horizontally explicit/vertically implicit (HEVI) time splitting scheme for atmospheric modelling is introduced, for which the horizontal divergence terms are applied within the implicit vertical substep. The new HEVI scheme is implemented in conjunction with a mixed mimetic spectral element spatial discretisation and semi-implicit vertical time stepping scheme that both preserve the skew-symmetric structure of the non-canonical Hamiltonian form of the equations of motion. Within this context the new HEVI scheme allows for the exact balance of all energetic exchanges in space and time. However since the choice of horizontal fluxes for which this balance is satisfied is not consistent with the horizontal velocity at the end of the time level the scheme still admits a temporal energy conservation error. Linearised eigenvalue analysis shows that similar to a fully implicit method, the new HEVI scheme is neutrally stable for all buoyancy modes, and unlike a second order trapezoidal HEVI scheme is stable for all acoustic modes below a certain horizontal CFL number. The scheme is validated against standard test cases for both planetary and non-hydrostatic regimes. For the planetary scale baroclinic instability test case, the new formulation exhibits a secondary oscillation in the potential to kinetic energy power exchanges, with a temporal frequency approximately four times that exhibited by a horizontally third order, vertically second order trapezoidal scheme. For the non-hydrostatic test case, the vertical upwinding of the potential temperature diagnostic equation is shown to reduce spurious oscillations without altering the energetics of the solution, since this upwinding is performed in an energetically consistent manner. For this test case, which is configured on an affine geometry, the exact balance of energy exchanges allows the model to run stably without any form of dissipation.

Horizontally explicit/vertically implicit (HEVI) schemes are a popular approach to time stepping in atmospheric models. This is on account of the importance of removing the CFL limit imposed by the explicit representation of vertical dynamics, and the desire to avoid the computational expense of a three dimensional implicit solve at each nonlinear iteration. Numerous HEVI schemes have been implemented [1–5] in previous models. These schemes are typically motivated by a desire to improve upon the stability and dispersion properties of the second order trapezoidal HEVI scheme through the use of additional sub steps within each time step [6, 7]. The application of additional substeps qualitatively improves the stepping scheme by ensuring the linear stability of the horizontal buoyancy and acoustic models below some CFL limit.

In the present article a new HEVI scheme is introduced, motivated by a desire to preserve the exact balance of energetic exchanges. The scheme is implemented within the context of a mixed mimetic spectral element spatial discretisation [8] and energy conserving implicit vertical time stepping method [9] that together allow for the exact balance of all energy exchanges in space and time. This is achieved by evaluating the horizontal mass and temperature fluxes within the vertically implicit sub-step, after first computing a provisional horizontal velocity from which to derive a second order temporal representation of the horizontal fluxes. Unlike other HEVI schemes, the new method is composed of only two horizontal and one vertical substep. In contrast to the standard second order trapezoidal splitting scheme, linear analysis shows the new method to be unconditionally stable for all acoustic modes below some horizontal CFL number, and all buoyancy modes independent of CFL number. While the energetic exchanges

*Corresponding author. Tel. +61 452 262 804.

Email address: davelee2804@gmail.com (David Lee)

are all exactly balanced, since the provisional horizontal velocity does not exactly match the final horizontal velocity at the end of the time step, the scheme admits a horizontal energy conservation error, as is the case for explicit schemes used to solve compressible flow problems.

In section 1 the continuous 3D compressible Euler equations for atmospheric flows are introduced in variational form. The discrete analogue of these equations and a discussion of the energy conserving time stepping are then presented in section 2, and the new HEVI splitting scheme is described in section 3. Section 4 describes the energetically consistent stabilisation of spurious oscillations in potential temperature via the upwinding of the corresponding test space. The reproduction of standard test cases is described in section 5. Finally the conclusions are discussed in section 6.

1. Skew-symmetric variational structure of the compressible Euler equations

The compressible Euler equations for atmospheric flows may be expressed in skew-symmetric form for the evolution of the velocity, \mathbf{u} , density, ρ , and density weighted potential temperature, Θ , as:

$$\frac{\partial \mathbf{u}}{\partial t} = -\mathbf{q} \times \mathbf{U} - \nabla \Phi - \theta \nabla \Pi \quad (1a)$$

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot \mathbf{U} \quad (1b)$$

$$\frac{\partial \Theta}{\partial t} = -\nabla \cdot (\theta \mathbf{U}) \quad (1c)$$

where $\mathbf{q} := (\nabla \times \mathbf{u} + \mathbf{f})/\rho$ is the potential vorticity (and $\mathbf{f} = f\mathbf{e}_z$ is the Coriolis term), $\mathbf{U} := \rho \mathbf{u}$ is the mass flux, $\Phi := \frac{1}{2} \mathbf{u} \cdot \mathbf{u} + \rho g z$ is the Bernoulli function (with g the gravitational acceleration and z the height), $\theta := \Theta/\rho$ is the potential temperature, and $\Pi := c_p (R\Theta/p_0)^{\frac{R}{c_v}}$ is the Exner pressure (where c_v and c_p are the specific heats at constant volume and temperature respectively, R is the ideal gas constant and p_0 is the reference pressure). The corresponding Hamiltonian (total energy) is given as:

$$H = \int_{\Omega} \underbrace{\frac{1}{2} \rho \mathbf{u} \cdot \mathbf{u}}_K + \underbrace{\rho g z}_P + \underbrace{\frac{c_v}{c_p} \Theta \Pi}_{I} d\Omega, \quad (2)$$

where K is the kinetic energy, P is the potential energy, and I is the internal energy. In this work the domain, Ω is configured as either the volume on the surface of a sphere in physical units, or a high resolution Cartesian geometry which is periodic in the two horizontal dimensions. The variational derivatives of the Hamiltonian with respect to the dependent variables $(\mathbf{u}, \rho, \Theta)$ are

$$\frac{\delta H}{\delta \mathbf{u}} = \rho \mathbf{u} =: \mathbf{U}, \quad (3a)$$

$$\frac{\delta H}{\delta \rho} = \frac{1}{2} \mathbf{u} \cdot \mathbf{u} + g z =: \Phi, \quad (3b)$$

$$\frac{\delta H}{\delta \Theta} = c_p \left(\frac{R\Theta}{p_0} \right)^{\frac{R}{c_v}} =: \Pi, \quad (3c)$$

which give the mass flux, \mathbf{U} , the Bernoulli function, Φ and the Exner pressure, Π respectively.

The weak form of (1) is given for $\mathbf{u}, \mathbf{U} \in H(\text{div}, \Omega)$ and $\rho, \Theta, \Phi, \Pi \in L^2(\Omega)$ as:

$$\left\langle \boldsymbol{\beta}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle = -\langle \boldsymbol{\beta}, \mathbf{q} \times \mathbf{U} \rangle + \langle \nabla \cdot \boldsymbol{\beta}, \Phi \rangle + \langle \nabla \cdot \mathbb{P}^{\text{div}} [\boldsymbol{\theta} \boldsymbol{\beta}], \Pi \rangle, \quad \forall \boldsymbol{\beta} \in H(\text{div}, \Omega) \quad (4a)$$

$$\left\langle \gamma, \frac{\partial \rho}{\partial t} \right\rangle = -\langle \gamma, \nabla \cdot \mathbf{U} \rangle, \quad \forall \gamma \in L^2(\Omega) \quad (4b)$$

$$\left\langle \sigma, \frac{\partial \Theta}{\partial t} \right\rangle = -\langle \sigma, \nabla \cdot \mathbb{P}^{\text{div}} [\theta \mathbf{U}] \rangle, \quad \forall \sigma \in L^2(\Omega) \quad (4c)$$

where we have invoked inner product notation as $\langle a, b \rangle = \int_{\Omega} abd\Omega$, the integration-by-parts relation $\langle \boldsymbol{\beta}, \nabla \alpha \rangle = -\langle \nabla \cdot \boldsymbol{\beta}, \alpha \rangle$ (assuming periodic or homogeneous boundary conditions) in (4), and introduced the L^2 projection to $H(\text{div}, \Omega)$

$$\mathbb{P}^{\text{div}} : [L^2(\Omega)]^3 \rightarrow H(\text{div}, \Omega). \quad (5)$$

Specifically for $\mathbf{v} \in [L^2(\Omega)]^3$, the projection $\mathbb{P}^{\text{div}}[\mathbf{v}] \in H(\text{div}, \Omega)$ is given by

$$\langle \boldsymbol{\beta}, \mathbb{P}^{\text{div}}[\mathbf{v}] \rangle = \langle \boldsymbol{\beta}, \mathbf{v} \rangle, \quad \forall \boldsymbol{\beta} \in H(\text{div}, \Omega). \quad (6)$$

Energy conservation is assured for the choice of $\boldsymbol{\beta} = \mathbf{U}$, $\gamma = \Phi$, $\sigma = \Pi$ by summing all equations in (4), such that

$$\left\langle \frac{\delta H}{\delta \mathbf{u}}, \frac{\partial \mathbf{u}}{\partial t} \right\rangle + \left\langle \frac{\delta H}{\delta \rho}, \frac{\partial \rho}{\partial t} \right\rangle + \left\langle \frac{\delta H}{\delta \Theta}, \frac{\partial \Theta}{\partial t} \right\rangle = \frac{dH}{dt} = 0. \quad (7)$$

where the key ingredient for energy conservation is the skew-symmetry of (4).

2. Discrete formulation

In this section the mixed mimetic spectral element spatial discretisation used in the model is introduced, as well as an analysis of the geometric properties of the energy conserving time integration scheme.

2.1. Spatial discretisation

For the purposes of this article the salient feature of the mixed mimetic spectral element method is that it allows for the preservation of the skew-symmetric structure of the compressible Euler equations, and thus the conservation of energy and energetic exchanges in the discrete form. For a more detailed discussion the reader is referred to previous work on the use of this method for geophysical flow modelling [8, 10, 11], as well as more foundational works on the subject [12–15]. In order to begin this discussion we introduce the finite dimensional subspaces $\mathcal{P}_h \subset H^1(\Omega)$, $\mathcal{W}_h \subset H(\text{curl}, \Omega)$, $\mathcal{U}_h \subset H(\text{div}, \Omega)$ and $\mathcal{Q}_h \subset L^2(\Omega)$. These subspaces are spanned by basis functions

$$\text{span}\{\psi_1, \dots, \psi_{d_p}\} = \mathcal{P}_h, \quad \text{span}\{\alpha_1, \dots, \alpha_{d_w}\} = \mathcal{W}_h, \quad \text{span}\{\beta_1, \dots, \beta_{d_u}\} = \mathcal{U}_h, \quad \text{and} \quad \text{span}\{\gamma_1, \dots, \gamma_{d_Q}\} = \mathcal{Q}_h, \quad (8)$$

with d_p , d_w , d_u , and d_Q , the number of degrees of freedom of the spaces \mathcal{P}_h , \mathcal{W}_h , \mathcal{U}_h , and \mathcal{Q}_h , respectively. The basis functions themselves are constructed by tensor product combinations of nodal and edge polynomials [12, 16]. As usual, elements of these spaces can be represented as a linear combination of the basis functions, for example, for $p_h \in \mathcal{P}_h$, $\mathbf{w}_h \in \mathcal{W}_h$, $\mathbf{u}_h \in \mathcal{U}_h$, and $c_h \in \mathcal{Q}_h$ we have that

$$p_h = \sum_{i=1}^{d_p} \hat{p}_i \psi_i, \quad \mathbf{w}_h = \sum_{i=1}^{d_w} \hat{w}_i \alpha_i, \quad \mathbf{u}_h = \sum_{i=1}^{d_u} \hat{u}_i \beta_i, \quad \text{and} \quad c_h = \sum_{i=1}^{d_Q} \hat{c}_i \gamma_i. \quad (9)$$

We now introduce an equivalent, but more compact, notation for the expansion of a function as a linear combination of the basis functions

$$p_h = \psi_h \hat{p}_h, \quad \mathbf{w}_h = \alpha_h \hat{\mathbf{w}}_h, \quad \mathbf{u}_h = \beta_h \hat{\mathbf{u}}_h, \quad \text{and} \quad c_h = \gamma_h \hat{c}_h, \quad (10)$$

where $\hat{p}_h := [\hat{p}_1, \dots, \hat{p}_{d_p}]^\top$, $\hat{\mathbf{w}}_h := [\hat{w}_1, \dots, \hat{w}_{d_w}]^\top$, $\hat{\mathbf{u}}_h := [\hat{u}_1, \dots, \hat{u}_{d_u}]^\top$, and $\hat{c}_h := [\hat{c}_1, \dots, \hat{c}_{d_Q}]^\top$ are the column vectors containing the coefficients of the discrete representation of the associated functions, and $\psi_h := [\psi_1, \dots, \psi_{d_p}]$, $\alpha_h := [\alpha_1, \dots, \alpha_{d_w}]$, $\beta_h := [\beta_1, \dots, \beta_{d_u}]$, and $\gamma_h := [\gamma_1, \dots, \gamma_{d_Q}]$ are the row vectors containing the basis functions.

Remark 1. Throughout this work we will extensively use the basis functions ψ_h , α_h , β_h , and γ_h . For clarity in the notation, these symbols will be solely used to refer to the basis functions:

$$\begin{array}{lll} \psi_h := [\psi_1, \dots, \psi_{d_p}] & \text{such that} & \text{span}\{\psi_1, \dots, \psi_{d_p}\} = \mathcal{P}_h \\ \alpha_h := [\alpha_1, \dots, \alpha_{d_w}] & \text{such that} & \text{span}\{\alpha_1, \dots, \alpha_{d_w}\} = \mathcal{W}_h \\ \beta_h := [\beta_1, \dots, \beta_{d_u}] & \text{such that} & \text{span}\{\beta_1, \dots, \beta_{d_u}\} = \mathcal{U}_h \\ \gamma_h := [\gamma_1, \dots, \gamma_{d_Q}] & \text{such that} & \text{span}\{\gamma_1, \dots, \gamma_{d_Q}\} = \mathcal{Q}_h. \end{array} \quad (11)$$

These discrete function spaces satisfy a compatibility relation that is expressed as a discrete de Rham complex of the form

$$\mathbb{R} \longrightarrow \mathcal{P}_h \xrightarrow{\nabla} \mathcal{W}_h \xrightarrow{\nabla \times} \mathcal{U}_h \xrightarrow{\nabla \cdot} \mathcal{Q}_h \longrightarrow 0. \quad (12)$$

The differential operators ∇ , $\nabla \times$, and $\nabla \cdot$ are concisely represented at the discrete level by incidence matrices, $\mathbf{E}^{1,0}$, $\mathbf{E}^{2,1}$, $\mathbf{E}^{3,2}$, such that for $p_h \in \mathcal{P}_h$, $\mathbf{w}_h \in \mathcal{W}_h$, and $\mathbf{u}_h \in \mathcal{U}_h$ we have that

$$\nabla p_h = \alpha_h \mathbf{E}^{1,0} \hat{p}_h, \quad \nabla \times \mathbf{w}_h = \beta_h \mathbf{E}^{2,1} \hat{\mathbf{w}}_h, \quad \text{and} \quad \nabla \cdot \mathbf{u}_h = \gamma_h \mathbf{E}^{3,2} \hat{\mathbf{u}}_h. \quad (13)$$

In a similar way, we may apply the differential operators to the basis functions ψ_h , α_h , β_h , and γ_h yielding

$$\nabla \psi_h = \alpha_h \mathbf{E}^{1,0}, \quad \nabla \times \alpha_h = \beta_h \mathbf{E}^{2,1}, \quad \text{and} \quad \nabla \cdot \beta_h = \gamma_h \mathbf{E}^{3,2}. \quad (14)$$

These incidence matrices satisfy the identities

$$\mathbf{E}^{2,1} \mathbf{E}^{1,0} = \mathbf{0}^{2,0}, \quad \text{and} \quad \mathbf{E}^{3,2} \mathbf{E}^{2,1} = \mathbf{0}^{3,1}, \quad (15)$$

where $\mathbf{0}^{2,0}$ is the $d_{\mathcal{U}} \times d_{\mathcal{P}}$ zero matrix and $\mathbf{0}^{3,1}$ is the $d_{\mathcal{Q}} \times d_{\mathcal{W}}$ zero matrix. Note that these two identities are directly related to the well known vector calculus identities $\nabla \times \nabla = 0$ and $\nabla \cdot \nabla \times = 0$. For a more detailed discussion of the incidence matrices and their properties the reader is directed to [14].

Additionally, these compatible function spaces satisfy an integration-by-parts property, such that the reverse mappings $(\tilde{\nabla} \cdot, \tilde{\nabla} \times, \tilde{\nabla})$, satisfy

$$0 \longleftarrow \mathcal{P}_h \xleftarrow{\tilde{\nabla} \cdot} \mathcal{W}_h \xleftarrow{\tilde{\nabla} \times} \mathcal{U}_h \xleftarrow{\tilde{\nabla}} \mathcal{Q}_h \longleftarrow \mathbb{R}. \quad (16)$$

and are represented by weak form adjoint relations. Assuming periodic or homogeneous boundary conditions we have that for $\mathbf{w}_h \in \mathcal{W}_h$, $\mathbf{u}_h \in \mathcal{U}_h$, and $c_h \in \mathcal{Q}_h$

$$\begin{aligned} \langle \beta_h, \tilde{\nabla} c_h \rangle &= -\langle \nabla \cdot \beta_h, c_h \rangle, \\ \langle \alpha_h, \tilde{\nabla} \times \mathbf{u}_h \rangle &= \langle \nabla \times \alpha_h, \mathbf{u}_h \rangle, \\ \langle \psi_h, \tilde{\nabla} \cdot \mathbf{w}_h \rangle &= -\langle \nabla \psi_h, \mathbf{w}_h \rangle. \end{aligned} \quad (17)$$

Remark 2. In (17) we have used an implicit matrix notation. As an example, consider the expression $\langle \beta_h, \beta_h \rangle$. This expression corresponds to

$$\langle \beta_h, \beta_h \rangle := \int_{\Omega} \beta_h^{\top} \beta_h \, d\Omega, \quad (18)$$

which is a matrix of dimensions $d_{\mathcal{U}} \times d_{\mathcal{U}}$ with element (i, j) , $\langle \beta_h, \beta_h \rangle_{i,j}$, given by

$$\langle \beta_h, \beta_h \rangle_{i,j} = \langle \beta_i, \beta_j \rangle.$$

As a more complex example, consider the expression $\langle \nabla \psi_h, \mathbf{w}_h \rangle$, with $\mathbf{w}_h \in \mathcal{W}_h$. In a similar way, this expression corresponds to

$$\langle \nabla \psi_h, \mathbf{w}_h \rangle = \int_{\Omega} \nabla \psi_h^{\top} \mathbf{w}_h \, d\Omega,$$

which is a column vector of $d_{\mathcal{W}}$ elements with element j , $\langle \nabla \psi_h, \mathbf{w}_h \rangle_j$, given by [17]

$$\langle \nabla \psi_h, \mathbf{w}_h \rangle_j \stackrel{(18)}{=} \left(\int_{\Omega} \nabla \psi_h^{\top} \mathbf{w}_h \right)_j \stackrel{(14)}{=} \left(\int_{\Omega} (\mathbf{E}^{1,0})^{\top} \alpha_h^{\top} \mathbf{w}_h \right)_j = \left((\mathbf{E}^{1,0})^{\top} \langle \alpha_h, \mathbf{w}_h \rangle \right)_j = \sum_{k=1}^{d_{\mathcal{W}}} \mathbf{E}_{k,j}^{1,0} \langle \alpha_k, \mathbf{w}_h \rangle.$$

This leads to the following discrete representations of the dual operators applied to $\mathbf{w}_h \in \mathcal{P}_h$, $\mathbf{u}_h \in \mathcal{U}_h$, and $c_h \in \mathcal{Q}_h$

$$\begin{aligned} \tilde{\nabla} c_h &= -\beta_h \langle \beta_h, \beta_h \rangle^{-1} (\mathbf{E}^{3,2})^{\top} \langle \gamma_h, \gamma_h \rangle \hat{c}_h, \\ \tilde{\nabla} \times \mathbf{u}_h &= \alpha_h \langle \alpha_h, \alpha_h \rangle^{-1} (\mathbf{E}^{2,1})^{\top} \langle \beta_h, \beta_h \rangle \hat{\mathbf{u}}_h, \\ \tilde{\nabla} \cdot \mathbf{w}_h &= -\psi_h \langle \psi_h, \psi_h \rangle^{-1} (\mathbf{E}^{1,0})^{\top} \langle \alpha_h, \alpha_h \rangle \hat{\mathbf{w}}_h. \end{aligned} \quad (19)$$

The discrete form of the variational 3D compressible Euler equations (4) may then be expressed as: Given $\mathbf{u}_h|_{t=0} \in \mathcal{U}$, $\rho|_{t=0}, \Theta|_{t=0} \in \mathcal{Q}$, find $\mathbf{u}_h \in \mathcal{U}$, $\rho, \Theta \in \mathcal{Q}$ such that

$$\left\langle \boldsymbol{\beta}_h, \frac{\partial \mathbf{u}_h}{\partial t} \right\rangle = -\langle \boldsymbol{\beta}_h, \mathbf{q}_h \times \mathbf{U}_h \rangle + \langle \nabla \cdot \boldsymbol{\beta}_h, \Phi_h \rangle + \langle \nabla \cdot \mathbb{P}^{\text{div}} [\theta_h \boldsymbol{\beta}_h], \Pi_h \rangle, \quad (20a)$$

$$\left\langle \gamma_h, \frac{\partial \rho_h}{\partial t} \right\rangle = -\langle \gamma_h, \nabla \cdot \mathbf{U}_h \rangle, \quad (20b)$$

$$\left\langle \gamma_h, \frac{\partial \Theta_h}{\partial t} \right\rangle = -\langle \gamma_h, \nabla \cdot \mathbb{P}^{\text{div}} [\theta_h \mathbf{U}_h] \rangle, \quad (20c)$$

and $\mathbf{U}_h \in \mathcal{U}$, $\Phi_h, \Pi_h \in \mathcal{Q}$ such that

$$\left\langle \boldsymbol{\beta}_h, \frac{\delta H_h}{\delta \mathbf{u}_h} \right\rangle = \langle \boldsymbol{\beta}_h, \mathbf{U}_h \rangle = \langle \boldsymbol{\beta}_h, \rho_h \mathbf{u}_h \rangle, \quad (21a)$$

$$\left\langle \gamma_h, \frac{\delta H_h}{\delta \rho_h} \right\rangle = \langle \gamma_h, \Phi_h \rangle = \langle \gamma_h, \frac{1}{2} \mathbf{u}_h \cdot \mathbf{u}_h + gz \rangle, \quad (21b)$$

$$\left\langle \gamma_h, \frac{\delta H_h}{\delta \Theta_h} \right\rangle = \langle \gamma_h, \Pi_h \rangle = \langle \gamma_h, c_p \left(\frac{R \Theta_h}{p_0} \right)^{\frac{R}{c_p}} \rangle. \quad (21c)$$

The system of equations (20), obtained after spatial discretisation, constitutes a system of ordinary differential equations. We now wish to devise a discrete time integration formulation that preserves the energy conservation properties of the spatial discretisation in (20). To introduce the energy conserving time integration scheme employed in this work, we recall the total energy H , (2),

$$H_h = \int_{\Omega} \underbrace{\frac{1}{2} \rho_h \mathbf{u}_h \cdot \mathbf{u}_h}_{K_h} + \underbrace{\rho_h gz}_{P_h} + \underbrace{\frac{c_v}{c_p} \Theta_h \Pi_h}_{I_h} d\Omega, \quad (22)$$

and its time derivative, (7),

$$\left\langle \frac{\delta H_h}{\delta \mathbf{u}_h}, \frac{\partial \mathbf{u}_h}{\partial t} \right\rangle + \left\langle \frac{\delta H_h}{\delta \rho_h}, \frac{\partial \rho_h}{\partial t} \right\rangle + \left\langle \frac{\delta H_h}{\delta \Theta_h}, \frac{\partial \Theta_h}{\partial t} \right\rangle = \left\langle \mathbf{U}_h, \frac{\partial \mathbf{u}_h}{\partial t} \right\rangle + \left\langle \Phi_h, \frac{\partial \rho_h}{\partial t} \right\rangle + \left\langle \Pi_h, \frac{\partial \Theta_h}{\partial t} \right\rangle = \frac{dH_h}{dt} = 0. \quad (23)$$

We now note that (20) can be written in matrix form as

$$\begin{bmatrix} \langle \boldsymbol{\beta}_h, \boldsymbol{\beta}_h \rangle \frac{d\hat{\mathbf{u}}_h}{dt} \\ \langle \gamma_h, \gamma_h \rangle \frac{d\hat{\rho}_h}{dt} \\ \langle \gamma_h, \gamma_h \rangle \frac{d\hat{\Theta}_h}{dt} \end{bmatrix} = \begin{bmatrix} -\langle \boldsymbol{\beta}_h, \mathbf{q}_h \times \boldsymbol{\beta}_h \rangle & \langle \nabla \cdot \boldsymbol{\beta}_h, \gamma_h \rangle & \langle \nabla \cdot \mathbb{P}^{\text{div}} [\theta_h \boldsymbol{\beta}_h], \gamma_h \rangle \\ -\langle \gamma_h, \nabla \cdot \boldsymbol{\beta}_h \rangle & 0 & 0 \\ -\langle \gamma_h, \nabla \cdot \mathbb{P}^{\text{div}} [\theta_h \boldsymbol{\beta}_h] \rangle & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{U}}_h \\ \hat{\Phi}_h \\ \hat{\Pi}_h \end{bmatrix}. \quad (24)$$

2.2. Temporal discretisation

Equation (24) is essentially a system of ordinary differential equations (ODEs) with the general form

$$\mathbf{M} \frac{d\hat{\mathbf{y}}(t)}{dt} = \mathbf{S}(\hat{\mathbf{y}}(t)) \hat{\mathbf{x}}(t), \quad (25)$$

where

$$\hat{\mathbf{y}}(t) := \begin{bmatrix} \hat{\mathbf{u}}_h \\ \hat{\rho}_h \\ \hat{\Theta}_h \end{bmatrix}, \quad \hat{\mathbf{x}}(t) := \begin{bmatrix} \hat{\mathbf{U}}_h \\ \hat{\Phi}_h \\ \hat{\Pi}_h \end{bmatrix}, \quad (26)$$

$\mathbf{S}(\hat{\mathbf{y}}(t))$ is a time dependent skew-symmetric matrix and \mathbf{M} is a constant in time symmetric positive definite matrix. The variable $\hat{\mathbf{y}}$ (26) is a column vector containing the coefficients of the discrete representation of the fields \mathbf{u}_h , ρ_h , and Θ_h , as introduced in (9) and (10). In this way we have that

$$\hat{y}_i = \hat{u}_i, \quad \text{for } i = 1, \dots, d_u, \quad \hat{y}_{i+d_u} = \hat{\rho}_i, \quad \text{for } i = 1, \dots, d_Q, \quad \text{and } \hat{y}_{i+d_u+d_Q} = \hat{\Theta}_i, \quad \text{for } i = 1, \dots, d_Q. \quad (27)$$

These coefficients $\hat{\mathbf{y}}$ are time dependent and their evolution is governed by the system of ODEs (25).

To solve this system of ODEs in time, we follow a similar approach to that for the spatial discretisation. We approximate the time dependent coefficients $\hat{\mathbf{y}}(t)$ on the time step interval $[t^n, t^{n+1}]$ with a temporal discrete polynomial function space of degree s , $\mathcal{T}_h^0([t^n, t^{n+1}]) := \text{span}\{l_0, \dots, l_s\}$ where $l_i(t)$, $i = 0, \dots, s$, are the $(s+1)$ nodal polynomials of degree s , as mentioned before for the spatial discretisation. In this way the polynomial temporal expansion of the spatial coefficients $\hat{\mathbf{y}}_i^h$ is

$$\hat{\mathbf{y}}_i^h = l_h \hat{\mathbf{y}}_i^h := \sum_{j=0}^s l_j(t) \hat{\mathbf{y}}_i^j, \quad t \in ([t^n, t^{n+1}]), \quad (28)$$

where, as in (9) and (10), $\hat{\mathbf{y}}_i^h := [\hat{\mathbf{y}}_i^1, \dots, \hat{\mathbf{y}}_i^{s+1}]$. The double hat notation $\hat{\mathbf{y}}_i^h$ in (28) highlights the double discretization taking place (spatial and temporal). The terms $\hat{\mathbf{y}}_i^h$ with $i = 1, \dots, d_u + 2d_Q$ correspond to the time dependent polynomial expansion of the i -th coefficient of the spatial polynomial expansion $\hat{\mathbf{y}} := [\hat{\mathbf{u}}_h, \hat{\rho}_h, \hat{\Theta}_h]^T$. The terms $\hat{\mathbf{y}}_i^j$ correspond to the j -th coefficient of the temporal polynomial expansion associated to the i -th coefficient of the spatial polynomial expansion.

The time dependent coefficients $\hat{\mathbf{x}}(t)$ are approximated with a temporal discrete polynomial function space of degree $(s-1)$, $\mathcal{T}_h^1([t^n, t^{n+1}]) := \text{span}\{e_1, \dots, e_s\}$ where $e_i(t)$, $i = 1, \dots, s$, are the s edge polynomials of degree $(s-1)$, as mentioned before for the spatial discretisation. In this way

$$\hat{\mathbf{x}}_i^h = e_h \hat{\mathbf{x}}_i^h := \sum_{j=1}^s e_j(t) \hat{\mathbf{x}}_i^j, \quad t \in ([t^n, t^{n+1}]). \quad (29)$$

As before, these two function spaces constitute a discrete de Rham complex

$$\mathbb{R} \longrightarrow \mathcal{T}_h^0([t^n, t^{n+1}]) \xrightarrow{\frac{d}{dt}} \mathcal{T}_h^1([t^n, t^{n+1}]) \longrightarrow 0, \quad (30)$$

and we have that

$$\frac{d\hat{\mathbf{y}}_i^h}{dt} = e_h \mathbf{E}_t^{1,0} \hat{\mathbf{y}}_i^h, \quad \text{and} \quad \frac{d\hat{\mathbf{x}}_i^h}{dt} = e_h \mathbf{E}_t^{1,0} \hat{\mathbf{x}}_i^h. \quad (31)$$

Remark 3. Note that here we added the subscript t to the incidence matrix, $\mathbf{E}_t^{1,0}$, to distinguish it from the spatial ones. Also, as mentioned before, the double hat highlights the fact that we have performed both discretisation in space and time and these coefficients are associated to this double discretisation. When an explicit reference to the temporal indices is required they are presented as superscript, as in (28).

Remark 4. The discrete de Rham complex (30) is associated to a single time step (for simplicity of exposition and computational efficiency). Naturally, as for the spatial discretisation, considering multiple time steps is a straightforward extension. In this work we perform the discretisation in a per step fashion.

The temporal discretisation then follows a similar mixed Galerkin approach as presented for the spatial discretisation: Given $\hat{\mathbf{x}}_i^h \in \mathcal{T}_h^1([t^n, t^{n+1}])$, with $i = 1, \dots, (d_u + 2d_Q)$, find $\hat{\mathbf{y}}_i^h \in \mathcal{T}_h^0([t^n, t^{n+1}])$, with $i = 1, \dots, (d_u + 2d_Q)$, such that

$$\mathbf{M} \langle e_h, \frac{d\hat{\mathbf{y}}_h(t)}{dt} \rangle_{[t^n, t^{n+1}]} = \langle e_h, \mathbf{S}(\hat{\mathbf{y}}_h(t)) \hat{\mathbf{x}}_h(t) \rangle_{[t^n, t^{n+1}]} . \quad (32)$$

Remark 5. The time-related inner products follow the same rules outlined before, namely

$$\langle e_h, \frac{d\hat{\mathbf{y}}_h(t)}{dt} \rangle_{[t^n, t^{n+1}]} := \int_{t^n}^{t^{n+1}} e_h^T \frac{d\hat{\mathbf{y}}_h(t)}{dt} dt. \quad (33)$$

Therefore, (32) is

$$\mathbf{M} \int_{t^n}^{t^{n+1}} e_h^\top \frac{d\hat{\mathbf{y}}_h(t)}{dt} dt = \int_{t^n}^{t^{n+1}} e_h^\top \mathbf{S}(\hat{\mathbf{y}}_h(t)) \hat{\mathbf{x}}_h(t) dt. \quad (34)$$

Note that we implicitly separate the spatial from the temporal coefficients, for simplicity of the notation, i.e., the constant in time matrix \mathbf{M} and the time dependent matrix $\mathbf{S}(\hat{\mathbf{y}}_h(t))$ apply at the spatial level also.

This choice of time integration is notable because it is energy conserving. To prove this, we recall the expression for the time variation of the energy of this system of equations (the Hamiltonian, H_h) (23)

$$\int_{t^n}^{t^{n+1}} \frac{dH_h}{dt} dt = H_h(\mathbf{y}_h(t^{n+1})) - H_h(\mathbf{y}_h(t^n)) = \int_{t^n}^{t^{n+1}} \left\langle \frac{\delta H_h}{\delta \mathbf{y}_h}, \frac{\partial \mathbf{y}_h}{\partial t} \right\rangle dt. \quad (35)$$

We may now use (21) to rewrite (35) as

$$\int_{t^n}^{t^{n+1}} \frac{dH_h}{dt} dt = H_h(\mathbf{y}_h(t^{n+1})) - H_h(\mathbf{y}_h(t^n)) = \int_{t^n}^{t^{n+1}} \left\langle \mathbf{x}_h, \frac{\partial \mathbf{y}_h}{\partial t} \right\rangle dt = \int_{t^n}^{t^{n+1}} \hat{\mathbf{x}}_h^\top \mathbf{M} \frac{d\hat{\mathbf{y}}_h}{dt} dt = \hat{\mathbf{x}}_h^\top \mathbf{M} \int_{t^n}^{t^{n+1}} e_h^\top \frac{d\hat{\mathbf{y}}_h}{dt} dt. \quad (36)$$

Finally, using the proposed time integration scheme, (34), we obtain energy conservation

$$H_h(\mathbf{y}_h(t^{n+1})) - H_h(\mathbf{y}_h(t^n)) = \hat{\mathbf{x}}_h^\top \mathbf{M} \int_{t^n}^{t^{n+1}} e_h^\top e_h dt (\hat{\mathbf{y}}_h^{n+1} - \hat{\mathbf{y}}_h^n) = \hat{\mathbf{x}}_h^\top \int_{t^n}^{t^{n+1}} e_h^\top \mathbf{S}(\hat{\mathbf{y}}_h(t)) \hat{\mathbf{x}}_h(t) dt = \hat{\mathbf{x}}_h^\top \left(\int_{t^n}^{t^{n+1}} e_h^\top \mathbf{S}(\hat{\mathbf{y}}_h(t)) e_h dt \right) \hat{\mathbf{x}}_h. \quad (37)$$

Since the term inside the parenthesis on the right hand side is skew-symmetric, the quadratic term cancels and we obtain energy conservation at the discrete level

$$H_h(\mathbf{y}_h(t^{n+1})) - H_h(\mathbf{y}_h(t^n)) = 0. \quad (38)$$

Remark 6. In equation (37), the term $\int_{t^n}^{t^{n+1}} e_h^\top e_h dt$ is a temporal mass matrix.

Remark 7. Two key ingredients need to be highlighted. The first is the construction of a system of equations of the form (24) and (25) where \mathbf{S} is skew-symmetric. This was only possible to obtain by a judicious choice of discrete function spaces and an equally careful selection of physical field quantities to employ in the construction of the system of equations. The second key ingredient is the construction of the time integration scheme also based on a sequence of polynomial spaces together with a Galerkin projection employing exact integration. Exact temporal integration is fundamental to guarantee the equality needed to perform the last step, equation (37).

This temporal discretisation is very closely related to the works by Hairer et al. [18, 19]. For the lowest order case (discussed below) this approach results in time stepping scheme identical to [19].

In this formulation any polynomial degree s may be employed to approximate the solution in time. In this work we choose $s = 1$ since this choice will greatly simplify the temporal discretisation. If this choice is made we will have

$$l_0(t) = \frac{t^{n+1} - t}{\Delta t}, \quad l_1(t) = \frac{t - t^n}{\Delta t}, \quad \text{and} \quad e_1(t) = \frac{1}{\Delta t}. \quad (39)$$

where $\Delta t := t^{n+1} - t^n$. In this case, since $e_h = e_1(t) = \frac{1}{\Delta t}$, the coefficients of the temporal polynomial expansion of $\hat{\mathbf{x}}_h(t)$ become

$$\hat{\mathbf{x}}_h(t) = [\hat{U}_1^1, \dots, \hat{U}_{d_u}^1, \hat{\Phi}_1^1, \dots, \hat{\Phi}_{d_Q}^1, \hat{\Pi}_1^1, \dots, \hat{\Pi}_{d_Q}^1]^\top e_1(t). \quad (40)$$

We recall now the properties of the edge polynomials, [12, 16], and its associated coefficients. Just as the nodal polynomial expansion interpolates a function, the edge polynomial expansion histopolates (interpolates the integral of) a function. While the coefficients of the nodal polynomial expansion correspond to pointwise sampling of a function, the coefficients of the edge polynomial expansion correspond to sampling of the integral of the function over intervals [12, 16]. Note that in the temporal discretization considered here there is only one polynomial basis in the edge polynomial expansion and therefore the temporal expansion of each spatial coefficient has only one coefficient.

Given the integral interpolation properties of the edge basis functions, the single coefficient of this expansion is the temporal integral

$$\hat{\mathbf{x}}_i^1 = \hat{U}_i^1 = \int_{t^n}^{t^{n+1}} \hat{U}_i^h(t) dt, \quad \hat{\mathbf{x}}_{j+d_U}^1 = \hat{\Phi}_j^1 = \int_{t^n}^{t^{n+1}} \hat{\Phi}_j^h(t) dt, \quad \hat{\mathbf{x}}_{k+d_U+d_Q}^1 = \hat{\Pi}_k^1 = \int_{t^n}^{t^{n+1}} \hat{\Pi}_k^h(t) dt, \quad (41)$$

with $i = 1, \dots, d_U$, and $j, k = 1, \dots, d_Q$. We now introduce

$$\overline{\hat{\mathbf{x}}_i^h} := \frac{\hat{\mathbf{x}}_i^1}{\Delta t} = \frac{\int_{t^n}^{t^{n+1}} \hat{\mathbf{x}}_i^h dt}{\Delta t}, \quad (42)$$

the exact time integral of the temporal approximation between time levels n and $n + 1$. Following the notation introduced before, (9), we will use

$$\overline{\hat{\mathbf{x}}_h} := [\overline{\hat{\mathbf{x}}_1^h}, \dots, \overline{\hat{\mathbf{x}}_{d_U+2d_Q}^h}]^\top \stackrel{(40)+(41)+(42)}{=} \frac{1}{\Delta t} [\hat{U}_1^1, \dots, \hat{U}_{d_U}^1, \hat{\Phi}_1^1, \dots, \hat{\Phi}_{d_Q}^1, \hat{\Pi}_1^1, \dots, \hat{\Pi}_{d_Q}^1]^\top. \quad (43)$$

If we substitute this into (34) we obtain

$$\mathbf{M} (\hat{\mathbf{y}}_h^{n+1} - \hat{\mathbf{y}}_h^n) = \left(\int_{t^n}^{t^{n+1}} \mathbf{S}(\hat{\mathbf{y}}_h(t)) dt \right) \frac{\hat{\mathbf{x}}_h}{\Delta t} \stackrel{(42)}{=} \left(\int_{t^n}^{t^{n+1}} \mathbf{S}(\hat{\mathbf{y}}_h(t)) dt \right) \overline{\hat{\mathbf{x}}_h}. \quad (44)$$

Since $\mathbf{S}(\hat{\mathbf{y}}_h(t))$ depends only linearly on time, we may use lowest order Gauss integration (midpoint rule) to exactly integrate the term involving $\mathbf{S}(\hat{\mathbf{y}}_h(t))$, yielding

$$\mathbf{M} (\hat{\mathbf{y}}_h^{n+1} - \hat{\mathbf{y}}_h^n) = \Delta t \mathbf{S} \left(\frac{\hat{\mathbf{y}}_h^{n+1} + \hat{\mathbf{y}}_h^n}{2} \right) \overline{\hat{\mathbf{x}}_h}. \quad (45)$$

Therefore this system of equations may be cast in a fully discrete form as:

$$\begin{aligned} \begin{bmatrix} \langle \boldsymbol{\beta}_h, \boldsymbol{\beta}_h \rangle \hat{\mathbf{u}}_h^{n+1} \\ \langle \gamma_h, \gamma_h \rangle \hat{\rho}_h^{n+1} \\ \langle \gamma_h, \gamma_h \rangle \hat{\Theta}_h^{n+1} \end{bmatrix} &= \begin{bmatrix} \langle \boldsymbol{\beta}_h, \boldsymbol{\beta}_h \rangle \hat{\mathbf{u}}_h^n \\ \langle \gamma_h, \gamma_h \rangle \hat{\rho}_h^n \\ \langle \gamma_h, \gamma_h \rangle \hat{\Theta}_h^n \end{bmatrix} - \\ \Delta t \begin{bmatrix} \langle \boldsymbol{\beta}_h, \mathbf{q}_h^{n+\frac{1}{2}} \times \boldsymbol{\beta}_h \rangle & -(\mathbf{E}^{3,2})^\top \langle \gamma_h, \gamma_h \rangle & -\langle \boldsymbol{\beta}_h, \theta_h^{n+\frac{1}{2}} \boldsymbol{\beta}_h \rangle \langle \boldsymbol{\beta}_h, \boldsymbol{\beta}_h \rangle^{-1} (\mathbf{E}^{3,2}) \langle \gamma_h, \gamma_h \rangle \\ \langle \gamma_h, \gamma_h \rangle \mathbf{E}^{3,2} & 0 & 0 \\ \langle \gamma_h, \gamma_h \rangle \mathbf{E}^{3,2} \langle \boldsymbol{\beta}_h, \boldsymbol{\beta}_h \rangle^{-1} \langle \boldsymbol{\beta}_h, \theta_h^{n+\frac{1}{2}} \boldsymbol{\beta}_h \rangle & 0 & 0 \end{bmatrix} \begin{bmatrix} \overline{\hat{U}}_h \\ \overline{\hat{\Phi}}_h \\ \overline{\hat{\Pi}}_h \end{bmatrix}. \end{aligned} \quad (46)$$

As mentioned before, $\overline{\hat{U}}_h$, $\overline{\hat{\Phi}}_h$, and $\overline{\hat{\Pi}}_h$ are the exact time integrals of the temporal polynomial reconstruction of \hat{U}_h , $\hat{\Phi}_h$, and $\hat{\Pi}_h$. An important point to note is the following. Since \hat{U}_h , $\hat{\Phi}_h$, and $\hat{\Pi}_h$ are functions of $\hat{\mathbf{u}}$, $\hat{\rho}$, and $\hat{\Theta}$, (3), we employ discrete piecewise linear approximations between time levels n and $n + 1$ to be described in the following section, and $\theta_h^{n+\frac{1}{2}} = (\theta_h^n + \theta_h^{n+1})/2$, $\mathbf{q}_h^{n+\frac{1}{2}} = (\mathbf{q}_h^n + \mathbf{q}_h^{n+1})/2$ are the time centered potential temperature and potential vorticity.

3. HEVI splitting

The above system (46) may be dimensionally split into an implicit solve for the vertical dynamics (incorporating the horizontal divergence terms) and explicit momentum advection by defining the horizontal and vertical velocity components respectively as $\mathbf{v}_h \in \mathcal{U}_h^\parallel$, $\mathbf{w}_h \in \mathcal{U}_h^\perp$, such that $\mathbf{u}_h = \mathbf{v}_h + \mathbf{w}_h \in \mathcal{U}_h$ [8]. Similarly the discrete vorticity vector, $\mathbf{q}_h \in \mathcal{W}_h$ may be partitioned into its respective vertical and horizontal components (in global coordinates), $q_h^\perp = \mathbf{q}_h \cdot \mathbf{e}_z$, $q_h^\parallel = (-\mathbf{q}_h \cdot \mathbf{e}_\phi, \mathbf{q}_h \cdot \mathbf{e}_\lambda)$, where \mathbf{e}_λ , \mathbf{e}_ϕ , \mathbf{e}_z are unit vectors in the zonal, meridional and vertical global

coordinates. The discrete strong form divergence operator, as given by the incidence matrix $\mathbf{E}^{3,2}$, may itself be decomposed into its horizontal, $\mathbf{E}_{\parallel}^{3,2}$ and vertical $\mathbf{E}_{\perp}^{3,2}$ components, such that

$$\nabla \cdot \mathbf{u}_h \stackrel{(13)}{=} \gamma_h \mathbf{E}^{3,2} \hat{\mathbf{u}}_h =: \gamma_h \left(\mathbf{E}_{\parallel}^{3,2} \hat{\mathbf{v}}_h + \mathbf{E}_{\perp}^{3,2} \hat{\mathbf{w}}_h \right). \quad (47)$$

The second order HEVI splitting is then given as:

Step 1: Explicit horizontal momentum solve

$$\begin{aligned} \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle \hat{\mathbf{v}}_h' &= \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle \hat{\mathbf{v}}_h^{n-1} - 2\Delta t \langle \boldsymbol{\beta}_h^{\parallel}, \mathbf{q}_h^{\perp, n} \times \boldsymbol{\beta}_h^{\parallel} \rangle \hat{\mathbf{v}}_h^n + 2\Delta t \langle \boldsymbol{\beta}_h^{\parallel}, \mathbf{q}_h^{\parallel, n}, \boldsymbol{\beta}_h^{\perp} \rangle \hat{\mathbf{w}}_h^n + \\ & 2\Delta t \left(\mathbf{E}_{\parallel}^{3,2} \right)^{\top} \langle \gamma_h, \gamma_h \rangle \hat{\Phi}_h^n + 2\Delta t \langle \boldsymbol{\beta}_h^{\parallel}, \theta_h^n \boldsymbol{\beta}_h^{\parallel} \rangle \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle^{-1} \left(\mathbf{E}_{\parallel}^{3,2} \right)^{\top} \langle \gamma_h, \gamma_h \rangle \hat{\Pi}_h^n \end{aligned} \quad (48)$$

Step 2: Implicit vertical solve (including horizontal divergence)

$$\begin{aligned} \begin{bmatrix} \langle \boldsymbol{\beta}_h^{\perp}, \boldsymbol{\beta}_h^{\perp} \rangle \hat{\mathbf{w}}_h^{n+1} \\ \langle \gamma_h, \gamma_h \rangle \hat{\rho}_h^{n+1} \\ \langle \gamma_h, \gamma_h \rangle \hat{\Theta}_h^{n+1} \end{bmatrix} &= \begin{bmatrix} \langle \boldsymbol{\beta}_h^{\perp}, \boldsymbol{\beta}_h^{\perp} \rangle \hat{\mathbf{w}}_h^n \\ \langle \gamma_h, \gamma_h \rangle \hat{\rho}_h^n \\ \langle \gamma_h, \gamma_h \rangle \hat{\Theta}_h^n \end{bmatrix} - \\ \Delta t \begin{bmatrix} \mathbf{0} & -\left(\mathbf{E}_{\perp}^{3,2} \right)^{\top} \langle \gamma_h, \gamma_h \rangle & -\langle \boldsymbol{\beta}_h^{\perp}, \theta_h^{n+1/2} \boldsymbol{\beta}_h^{\perp} \rangle \langle \boldsymbol{\beta}_h^{\perp}, \boldsymbol{\beta}_h^{\perp} \rangle^{-1} \left(\mathbf{E}_{\perp}^{3,2} \right)^{\top} \langle \gamma_h, \gamma_h \rangle \\ \langle \gamma_h, \gamma_h \rangle \mathbf{E}_{\perp}^{3,2} & \mathbf{0} & \mathbf{0} \\ \langle \gamma_h, \gamma_h \rangle \mathbf{E}_{\perp}^{3,2} \langle \boldsymbol{\beta}_h^{\perp}, \boldsymbol{\beta}_h^{\perp} \rangle^{-1} \langle \boldsymbol{\beta}_h^{\perp}, \theta_h^{n+1/2} \boldsymbol{\beta}_h^{\perp} \rangle & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \bar{\hat{\mathbf{w}}}_h \\ \bar{\hat{\Phi}}_h \\ \bar{\hat{\Pi}}_h \end{bmatrix} - \\ \Delta t \begin{bmatrix} \langle \boldsymbol{\beta}_h^{\perp}, \mathbf{q}_h^{\parallel, n+1/2} \cdot \boldsymbol{\beta}_h^{\parallel} \rangle \bar{\hat{\mathbf{v}}}_h \\ \langle \gamma_h, \gamma_h \rangle \mathbf{E}_{\parallel}^{3,2} \bar{\hat{\mathbf{v}}}_h \\ \langle \gamma_h, \gamma_h \rangle \mathbf{E}_{\parallel}^{3,2} \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle^{-1} \langle \boldsymbol{\beta}_h^{\parallel}, \theta_h^{n+1/2} \boldsymbol{\beta}_h^{\parallel} \rangle \bar{\hat{\mathbf{v}}}_h \end{bmatrix} \end{aligned} \quad (49)$$

Step 3: Explicit horizontal momentum solve

$$\begin{aligned} \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle \hat{\mathbf{v}}_h^{n+1} &= \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle \hat{\mathbf{v}}_h^n - \Delta t \langle \boldsymbol{\beta}_h^{\parallel}, \mathbf{q}_h^{\perp, n+1/2} \times \boldsymbol{\beta}_h^{\parallel} \rangle \bar{\hat{\mathbf{v}}}_h + \Delta t \langle \boldsymbol{\beta}_h^{\parallel}, \mathbf{q}_h^{\parallel, n+1/2}, \boldsymbol{\beta}_h^{\perp} \rangle \bar{\hat{\mathbf{w}}}_h + \\ & \Delta t \left(\mathbf{E}_{\parallel}^{3,2} \right)^{\top} \langle \gamma_h, \gamma_h \rangle \bar{\hat{\Phi}}_h + \Delta t \langle \boldsymbol{\beta}_h^{\parallel}, \theta_h^{n+1/2} \boldsymbol{\beta}_h^{\parallel} \rangle \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle^{-1} \left(\mathbf{E}_{\parallel}^{3,2} \right)^{\top} \langle \gamma_h, \gamma_h \rangle \bar{\hat{\Pi}}_h \end{aligned} \quad (50)$$

Left multiplication of (49) by $\begin{bmatrix} \bar{\hat{\mathbf{w}}}_h^{\top} & \bar{\hat{\Phi}}_h^{\top} & \bar{\hat{\Pi}}_h^{\top} \end{bmatrix}$ and (50) by $\bar{\hat{\mathbf{v}}}_h^{\top}$ leads to the cancellation of all forcing terms in both equations, resulting in the expression

$$\bar{\hat{\mathbf{v}}}_h \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle (\hat{\mathbf{v}}_h^{n+1} - \hat{\mathbf{v}}_h^n) + \bar{\hat{\mathbf{w}}}_h \langle \boldsymbol{\beta}_h^{\perp}, \boldsymbol{\beta}_h^{\perp} \rangle (\hat{\mathbf{w}}_h^{n+1} - \hat{\mathbf{w}}_h^n) + \bar{\hat{\Phi}}_h \langle \gamma_h, \gamma_h \rangle (\hat{\rho}_h^{n+1} - \hat{\rho}_h^n) + \bar{\hat{\Pi}}_h \langle \gamma_h, \gamma_h \rangle (\hat{\Theta}_h^{n+1} - \hat{\Theta}_h^n) \neq 0 \quad (51)$$

Note that since \mathbf{v}_h' and \mathbf{v}_h^{n+1} differ, energy is not strictly conserved in the temporal discretisation of the horizontal momentum equation. There is a conservation error of magnitude $\bar{\hat{\mathbf{v}}}_h \langle \boldsymbol{\beta}_h^{\parallel}, \boldsymbol{\beta}_h^{\parallel} \rangle (\hat{\mathbf{v}}_h' - \hat{\mathbf{v}}_h^{n+1})$, such that (51) is not strictly a discrete analogue of the continuous expression

$$\left(\frac{\delta H}{\delta \mathbf{y}} \right)^{\top} \cdot \frac{\partial \mathbf{y}}{\partial t} = 0. \quad (52)$$

The above formulation is similar to the “u-forward, pressure-backward” formulation [6], by which the horizontal divergence term is treated implicitly (using the most recent horizontal velocity update), and the horizontal pressure gradient term is treated explicitly. However in the present case the horizontal pressure gradient term at the final step is evaluated at the same time level as the (implicit) horizontal divergence term.

The exact, second order time integrals of the variational derivatives are then given as:

$$\langle \beta_h^\parallel, \beta_h^\parallel \rangle \widehat{V}_h = \frac{1}{3} \langle \beta_h^\parallel, \rho_h^n \beta_h^\parallel \rangle \widehat{v}_h^n + \frac{1}{6} \langle \beta_h^\parallel, \rho_h^{n+1} \beta_h^\parallel \rangle \widehat{v}_h^n + \frac{1}{6} \langle \beta_h^\parallel, \rho_h^n \beta_h^\parallel \rangle \widehat{v}_h' + \frac{1}{3} \langle \beta_h^\parallel, \rho_h^{n+1} \beta_h^\parallel \rangle \widehat{v}_h' \quad (53a)$$

$$\langle \beta_h^\perp, \beta_h^\perp \rangle \widehat{W}_h = \frac{1}{3} \langle \beta_h^\perp, \rho_h^n \beta_h^\perp \rangle \widehat{w}_h^n + \frac{1}{6} \langle \beta_h^\perp, \rho_h^{n+1} \beta_h^\perp \rangle \widehat{w}_h^n + \frac{1}{6} \langle \beta_h^\perp, \rho_h^n \beta_h^\perp \rangle \widehat{w}_h^{n+1} + \frac{1}{3} \langle \beta_h^\perp, \rho_h^{n+1} \beta_h^\perp \rangle \widehat{w}_h^{n+1} \quad (53b)$$

$$\begin{aligned} \langle \gamma_h, \gamma_h \rangle \widehat{\Phi}_h &= \frac{1}{6} \langle \gamma_h, \mathbf{v}_h^n \cdot \beta_h^\parallel \rangle \widehat{v}_h^n + \frac{1}{6} \langle \gamma_h, \mathbf{v}_h' \cdot \beta_h^\parallel \rangle \widehat{v}_h^n + \frac{1}{6} \langle \gamma_h, \mathbf{v}_h' \cdot \beta_h^\parallel \rangle \widehat{v}_h' + \\ &\quad \frac{1}{6} \langle \gamma_h, w_h^n \beta_h^\perp \rangle \widehat{w}_h^n + \frac{1}{6} \langle \gamma_h, w_h^{n+1} \beta_h^\perp \rangle \widehat{w}_h^n + \frac{1}{6} \langle \gamma_h, w_h^{n+1} \beta_h^\perp \rangle \widehat{w}_h^{n+1} + g \langle \gamma_h, \gamma_h \rangle \widehat{z}_h \end{aligned} \quad (53c)$$

$$\langle \gamma_h, \gamma_h \rangle \widehat{\Pi}_h = \frac{1}{2} \langle \gamma_h, \gamma_h \rangle \widehat{\Pi}_h^n + \frac{1}{2} \langle \gamma_h, \gamma_h \rangle \widehat{\Pi}_h^{n+1} \quad (53d)$$

Note that while a second order “leap-frog” method has been used to determine the provisional velocity, \widehat{v}_h' in (48), a first order, forward Euler step is probably also acceptable, since in either case the resulting mass flux (53a) will be second order in time.

We have not as yet made mention of the spaces used to represent the quantities found within the operator in (49): \mathbf{q}_h, θ_h . Conservation of energy is satisfied for any representation these quantities, provided that the skew-symmetry of this operator is preserved. We choose to represent $\mathbf{q}_h \in \mathcal{W}_h$, such that this may be derived from \mathbf{u}_h and ρ_h via a weak curl operator [8], while θ_h is represented in the vertical component of the $H(\text{div}, \Omega)$ space, \mathcal{U}_h^\perp (such that it is C^0 continuous in the vertical direction only). Previous work has shown that this representation of θ_h yields an improved representation of the dispersion relation for linearised buoyancy modes [20].

The implicit vertical system in (49) is solved in serial using a direct LU method, while the projections onto the horizontal \mathcal{U}_h^\parallel space in (48), (50), as well as to derive the horizontal pressure gradients and temperature fluxes in (49) and the horizontal mass flux in (53a) are solved in parallel using the conjugate gradient method. These are both done using the PETSc library [21–23].

3.1. Stability analysis

In order to study the stability of the HEVI splitting scheme described above this is applied to the two dimensional linearised compressible Boussinesq equations [24], which are given as:

$$\frac{\partial u}{\partial t} + \frac{\partial p}{\partial x} = 0 \quad (54a)$$

$$\frac{\partial w}{\partial t} + \frac{\partial p}{\partial z} - b = 0 \quad (54b)$$

$$\frac{\partial p}{\partial t} + c^2 \left(\frac{\partial u}{\partial x} + \frac{\partial w}{\partial z} \right) = 0 \quad (54c)$$

$$\frac{\partial b}{\partial t} + N^2 w = 0 \quad (54d)$$

where u and w are the horizontal and vertical velocities, p is the pressure, b is the buoyancy, c is the speed of sound and N is the Brunt-Väisälä frequency. Using the splitting scheme detailed above (but with a first order forward Euler step for the initial provisional velocity, u') and assuming periodic solutions of the form $(u, w, p, b)(t)e^{ikx+ilz}$ these may be time-stepped as:

$$u' = u^n - \Delta t ikp^n \quad (55a)$$

$$w^{n+1} + \frac{\Delta t il}{2} p^{n+1} - \frac{\Delta t}{2} b^{n+1} = w^n - \frac{\Delta t il}{2} p^n + \frac{\Delta t}{2} b^n \quad (55b)$$

$$p^{n+1} + \frac{\Delta t ilc^2}{2} w^{n+1} = p^n - \frac{\Delta t ilc^2}{2} w^n - \frac{\Delta t ikc^2}{2} u^n - \frac{\Delta t ikc^2}{2} u' \quad (55c)$$

$$b^{n+1} + \frac{\Delta t N^2}{2} w^{n+1} = b^n - \frac{\Delta t N^2}{2} w^n \quad (55d)$$

$$u^{n+1} + \frac{\Delta t ik}{2} p^{n+1} = u^n - \frac{\Delta t ik}{2} p^n \quad (55e)$$

Note that a temporal description of the HEVI scheme that omits the spatial discretisation is provided in Appendix A. Substituting the expression for u' into that for p^{n+1} and expressing in matrix form gives

$$\begin{bmatrix} 1 & 0 & \frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & \frac{\Delta t}{2} \frac{il}{l} & -\frac{\Delta t}{2} \\ 0 & \frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & \frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \begin{bmatrix} u^{n+1} \\ w^{n+1} \\ p^{n+1} \\ b^{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & -\frac{\Delta t}{2} \frac{il}{l} & \frac{\Delta t}{2} \\ -\Delta t \frac{ikc^2}{N^2} & -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & -\frac{\Delta t^2 k^2 c^2}{2} \\ 0 & -\frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \begin{bmatrix} u^n \\ w^n \\ p^n \\ b^n \end{bmatrix}. \quad (56)$$

In order to analyse the stability of the splitting scheme we inspect the eigenvalues of the operator

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & \frac{\Delta t}{2} \frac{il}{l} & -\frac{\Delta t}{2} \\ 0 & \frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & \frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & -\frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & -\frac{\Delta t}{2} \frac{il}{l} & \frac{\Delta t}{2} \\ -\Delta t \frac{ikc^2}{N^2} & -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & -\frac{\Delta t^2 k^2 c^2}{2} \\ 0 & -\frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \quad (57)$$

For a given eigenvalue, λ_i , this will be unstable if $\lambda_i \lambda_i^* > 1$. The four eigenvalues correspond to two acoustic modes, which differ in the sign of their imaginary component only, and two gravity (buoyancy) modes which similarly differ only in the sign of the imaginary term.

The results for the new HEVI scheme are compared to the standard second order trapezoidal HEVI scheme, for which the vertical terms are all treated implicitly, and the horizontal terms explicitly. This scheme may be written as:

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{\Delta t}{2} \frac{il}{l} & -\frac{\Delta t}{2} \\ 0 & \frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & \frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \begin{bmatrix} u' \\ w' \\ p' \\ b' \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\Delta t \frac{ik}{l} & 0 \\ 0 & 1 & -\frac{\Delta t}{2} \frac{il}{l} & \frac{\Delta t}{2} \\ -\Delta t \frac{ikc^2}{N^2} & -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & \frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \begin{bmatrix} u^n \\ w^n \\ p^n \\ b^n \end{bmatrix} \quad (58)$$

$$\begin{bmatrix} u^{n+1} \\ w^{n+1} \\ p^{n+1} \\ b^{n+1} \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & -\frac{\Delta t}{2} \frac{il}{l} & \frac{\Delta t}{2} \\ -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & -\frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \begin{bmatrix} u^n \\ w^n \\ p^n \\ b^n \end{bmatrix} + \begin{bmatrix} 0 & 0 & -\frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 0 & -\frac{\Delta t}{2} \frac{il}{l} & \frac{\Delta t}{2} \\ -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & 0 & 0 \\ 0 & -\frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 0 \end{bmatrix} \begin{bmatrix} u' \\ w' \\ p' \\ b' \end{bmatrix} \quad (59)$$

Combining the above expressions gives a linear operator by which the solution at time level $n + 1$ may be determined from the solution at time level n , from which the corresponding eigenvalues may be derived.

For completeness we also compare against the Crank-Nicolson scheme, which is semi-implicit in both horizontal and vertical dimensions, and neutrally stable for all vertical and horizontal wave numbers for both buoyancy and acoustic modes. The amplification factors for the Crank-Nicolson scheme are given as:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & \frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & \frac{\Delta t}{2} \frac{il}{l} & -\frac{\Delta t}{2} \\ \frac{\Delta t}{2} \frac{ilc^2}{N^2} & \frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & \frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & -\frac{\Delta t}{2} \frac{ik}{l} & 0 \\ 0 & 1 & -\frac{\Delta t}{2} \frac{il}{l} & \frac{\Delta t}{2} \\ -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & -\frac{\Delta t}{2} \frac{ilc^2}{N^2} & 1 & 0 \\ 0 & -\frac{\Delta t}{2} \frac{N^2}{N^2} & 0 & 1 \end{bmatrix} \quad (60)$$

The amplification factors for the acoustic and gravity modes for the three different time splitting schemes are given in figs. 1 and 2 respectively for a periodic domain of size 1000.0m in both dimensions, with $c = 340.0\text{ms}^{-1}$, $N = 0.01\text{s}^{-1}$. While the Crank-Nicolson scheme is neutrally stable, $|\lambda| = 1$ for both modes, and the trapezoidal HEVI scheme is unstable, $|\lambda| > 1$ for both modes (albeit with a very small amplification factor for the gravity modes), the energetically balanced splitting scheme is neutrally stable for all acoustic modes below some horizontal wave number, and unconditionally stable for all gravity modes. This represents a qualitative improvement on the standard second order HEVI splitting.

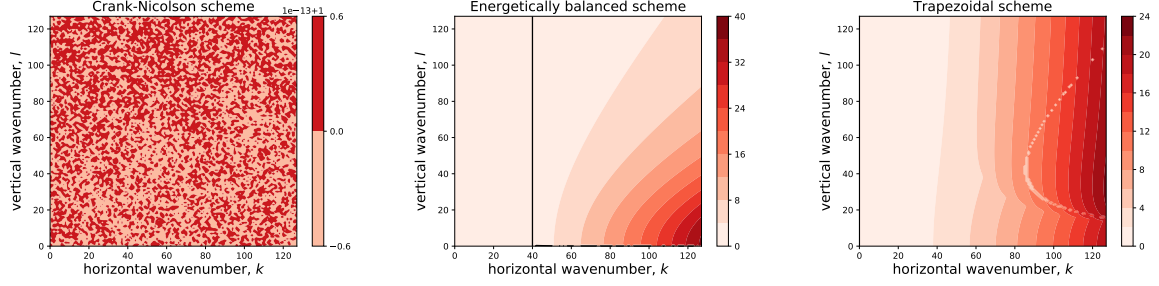


Figure 1: Amplification factor for the acoustic waves using the Crank-Nicolson (left), new (center) and trapezoidal HEVI (right) splitting. The black line on the center plot indicates an amplification factor of $1.0 + 10^{-12}$, indicating the onset of unstable modes. The Crank-Nicolson scheme is neutrally stable for all modes, the new scheme is neutrally stable for all modes below a finite horizontal wave number and the trapezoidal scheme is unstable for all modes.

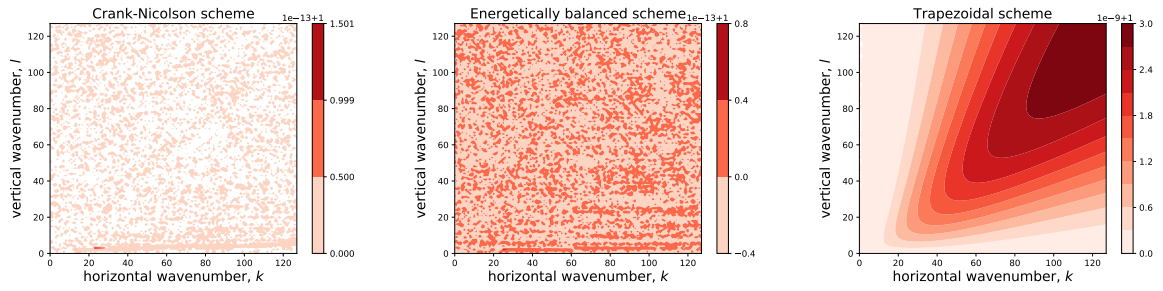


Figure 2: Amplification factor for the gravity waves using the Crank-Nicolson (left), new (center) and trapezoidal HEVI (right) splitting. The Crank-Nicolson scheme and the new scheme are both neutrally stable for all modes, while the trapezoidal scheme is unstable for all modes.

4. Potential temperature upwinding

This section describes the suppression of spurious oscillations in potential temperature via upwinding of the potential temperature diagnostic equation in an energetically consistent manner. Since the potential vorticity and potential temperature are represented within the skew-symmetric operator for the discrete variational form of the compressible Euler equations (46), these quantities may be upwinded so as to suppress oscillations without breaking energy conservation (so long as skew-symmetry is preserved). This strategy has been used previously to suppress oscillations for the shallow water equations by upwinding potential vorticity so as to dissipate potential enstrophy [25–27], and for the compressible Euler equations by upwinding potential temperature so as to generate entropy [28]. In the present context the upwinding of potential temperature is applied not to the material form advection equation, but rather to the diagnostic equation by which the potential temperature is derived from the prognostic variables of density (1b) and density-weighted potential temperature (1c).

While this is typically achieved for variational methods by augmenting the test space with a linear correction which allows for dissipation in the direction of the flow [29], in the present case this is achieved by integrating the test functions backwards along velocity characteristics [27]. Here the upwinding is applied in the vertical direction only. Since the discrete potential temperature is represented as $\theta_h \in \mathcal{U}_h^\perp$ [8], for which the vertical component of the tensor product basis is represented by piecewise linear polynomials, for this case the upwinded bases are effectively equivalent to the SUPG method [28, 29]. However in the present case this upwinding is applied to the diagnostic equation (the discrete form of the relation $\theta := \Theta/\rho$, (62)), and not the material form of the advection equation, $\partial\theta/\partial t + \mathbf{u} \cdot \nabla\theta = 0$, as is customary in the SUPG method. Since C^0 continuity for the \mathcal{U}_h^\perp space is enforced only in the vertical direction, upwinding of θ_h in the horizontal directions will have no effect.

The upwinded trial functions, $\beta_h^{\perp,u} \in \mathcal{U}_h^\perp$ are computed by evaluating the quadrature points at vertically *downwind* locations, ζ^d , which are computed to first order as [27]:

$$\zeta^d = \zeta + 0.5\Delta t \sum_i \hat{w}_i \beta_i^\perp(\xi, \eta, \zeta), \quad (61)$$

where ξ , η and ζ are the local element coordinates. Note that in the above expression the velocity is interpolated to its local element value, and not its global physical value. Note also that the choice to integrate the vertical quadrature points downwind for a time of $0.5\Delta t$ is arbitrary, and may be tuned to optimise the balance between reduced oscillation and less diffusive solutions. The upwinded trial functions are then evaluated as $\beta_h^{\perp,u} = \beta_h^\perp(\xi, \eta, \zeta^d)$. The upwinded potential temperature is then diagnosed as:

$$\langle \beta_h^{\perp,u}, \rho_h \beta_h^\perp \rangle \hat{\theta}_h = \langle \beta_h^{\perp,u}, \gamma_h \rangle \hat{\Theta}_h. \quad (62)$$

Note that in the above expression only the test functions for the potential temperature diagnostic equation are evaluated at upstream locations, while the trial functions remain static. Moreover since continuity is enforced only in the vertical direction for the trial space of the potential temperature, \mathcal{U}_h^\perp , it is only in this direction that upwinding has any material effect.

While the variational form of the compressible Euler equations (4) conserve entropy, $s = c_p \log(\theta) + s_0$ (see Appendix B), this is not true in the discrete form, since transcendental functions such as logarithms cannot be integrated exactly in the discrete form. As such, while the upwinding of the test functions as described above will generate additional entropy, there are already additional entropy sources at the discrete level.

5. Results

5.1. Baroclinic instability

In order to verify the energetically balanced HEVI scheme, this is applied to the solution of a standard test case for baroclinic instability on the sphere [30]. This test involves a small perturbation to an otherwise hydrostatically and geostrophically balanced atmosphere, which over several days generates a baroclinic wave. The model is stabilised using a biharmonic viscosity term on both the horizontal potential temperature equation and momentum equation (which does not conserve energy due to the explicit time stepping), and a Rayleigh friction term in the top layers of the atmosphere in the vertical (which is not strictly necessary however helps to accelerate the initial hydrostatic

adjustment process) [8]. Figures 3, 4 and 5 show the surface level Exner pressure, as well as the vertical vorticity component and potential temperature at a height of $z = 1.57\text{km}$ at days 8 and 10 respectively. As an initial validation exercise these compare well against previously published results [9, 30], however they do not allow us to draw any particular conclusion as to the superiority of any particular scheme. The signature of the baroclinic instability is also observed in the perturbation to the steady hydrostatic pressure profile, for which a vertical cross section at 50°N is shown at days 8 and 10 in fig 6.

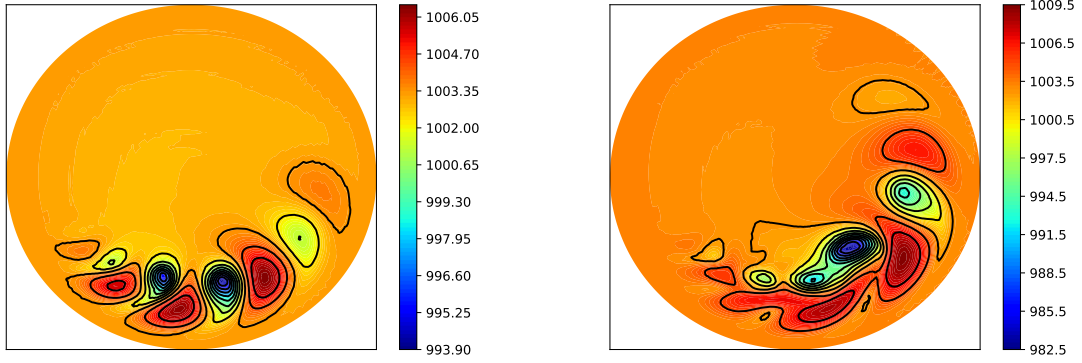


Figure 3: Baroclinic instability test case: Bottom level Exner pressure, days 8 and 10.

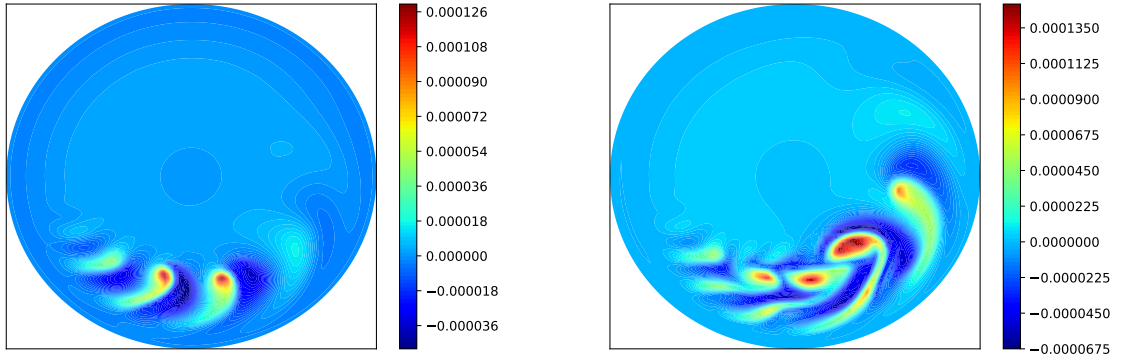


Figure 4: Baroclinic instability test case: Vertical vorticity component at $z = 1.57\text{km}$, days 8 and 10.

5.1.1. Comparison to TRAP(2,3,2)

The energetically balanced HEVI scheme is compared to the horizontally third order, vertically second order trapezoidal scheme, TRAP(2,3,2) [6, 7]. In most respects the energetics of the two schemes are similar, despite the fact that the TRAP(2,3,2) scheme is not energetically balanced, but is higher order in the horizontal (also a time step of $\Delta t = 60\text{s}$ is used for the energetically balanced scheme, while the TRAP(2,3,2) simulation was run with a time step of $\Delta t = 120\text{s}$). While the TRAP(2,3,2) scheme is stable for longer time steps than the energetically balanced scheme, it also involves an additional substep in both the vertical and horizontal directions. Both schemes are stabilised using biharmonic viscosity in the horizontal and Rayleigh friction in the top layers in the vertical [8, 9].

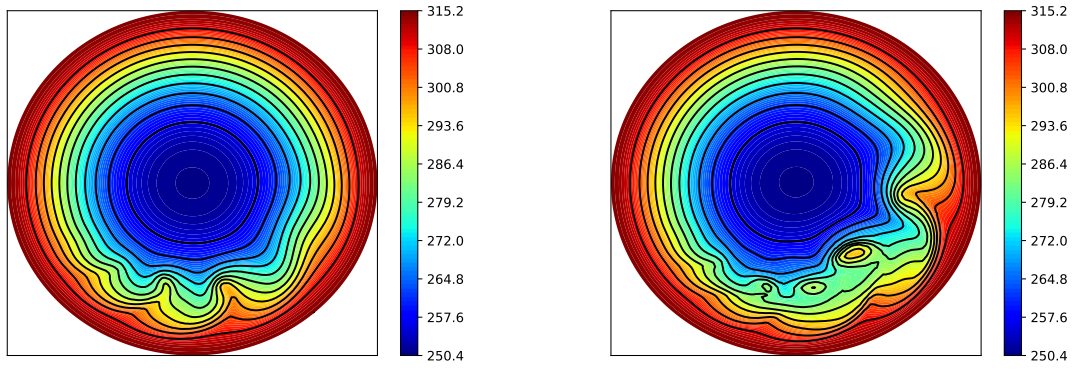


Figure 5: Baroclinic instability test case: Potential temperature at $z = 1.57 \text{ km}$, days 8 and 10.

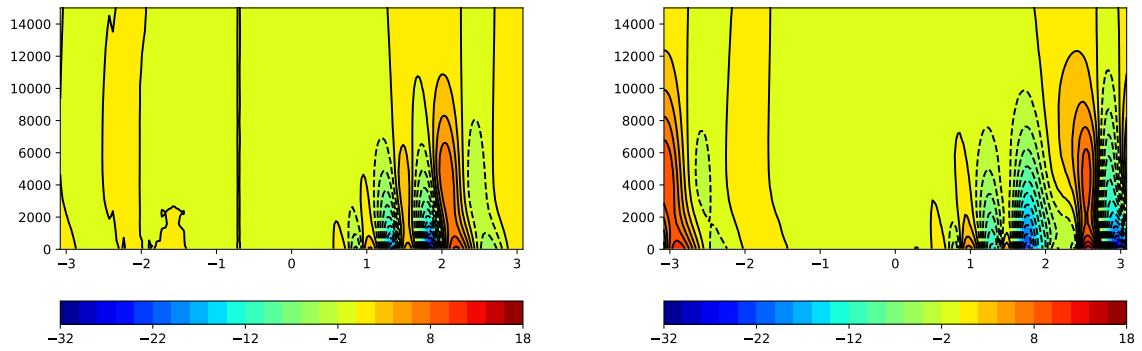


Figure 6: Baroclinic instability test case: Zonal-vertical pressure slice at 50°N , days 8 and 10.

The one respect in which the two schemes differ markedly is the potential to kinetic power exchanges, the transition of potential to kinetic energy over time. This is computed (assuming homogeneous Dirichlet boundary conditions for the vertical velocity) as the discrete equivalent of the term $-\partial P/\partial t = g \int z \partial(\rho w)/\partial z d\Omega$ as $-\partial P_h/\partial t = g \hat{z}_h \langle \gamma_h, \gamma_h \rangle \mathbf{E}_\perp^{3,2} \bar{\mathbf{W}}_h$ [8]. For the potential to kinetic power exchange the energetically balanced scheme exhibits a higher frequency oscillation not present in the TRAP(2,3,2) scheme, as shown in fig. 7. Fourier analysis of these power exchanges, taken between days 6 and 10, shows that while the amplitudes of the low-frequency modes ($k < 10$) match each other closely, both solutions exhibit a secondary peak, which for the TRAP(2,3,2) scheme occurs at $k = 33$ (2.91 hours), while for the energetically balanced scheme this peak occurs at $k = 131$ (0.73 hours). Curiously, the secondary peak for the energetically balanced scheme is almost exactly four times the frequency observed for the TRAP(2,3,2) scheme. Since the secondary peak for the TRAP(2,3,2) scheme is closer in frequency to the low frequency modes, which exhibit very similar behaviour for both schemes, there is a possibility that this may alias onto the low frequency time scales of the baroclinic motions.

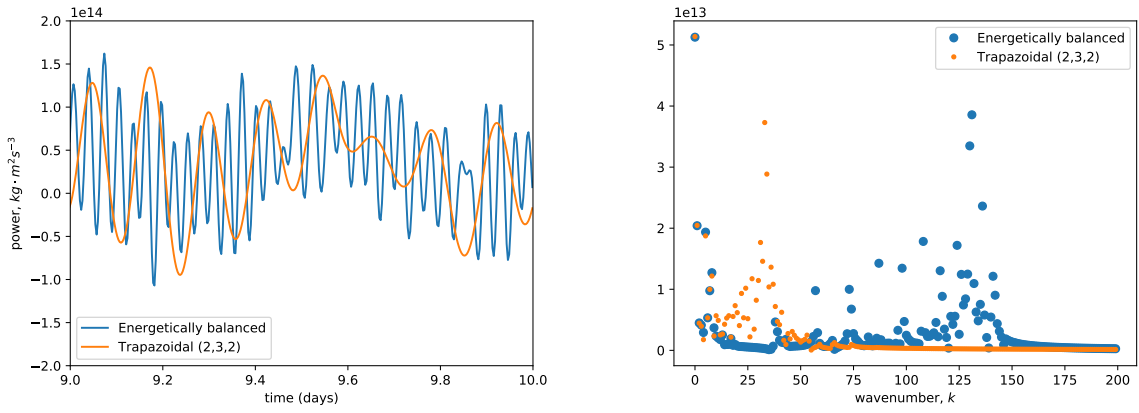


Figure 7: Baroclinic instability test case: Comparison of potential to kinetic energy power exchanges for the energetically balanced and TRAP(2,3,2) HEVI schemes; left: real space representation from day 9 to 10, right: amplitude of the corresponding Fourier modes from days 6 to 10.

5.2. 3D Rising bubble

In order to verify the scheme for non-hydrostatic dynamics this is applied to a standard test case for a 3D rising bubble, which is initialised as a small perturbation in potential temperature against a constant background value, with the Exner pressure specified so as to satisfy hydrostatic balance [2, 31, 32]. The domain is configured as a horizontally periodic box of size $1000\text{m} \times 1000\text{m} \times 1500\text{m}$ with a flat bottom and top with homogeneous boundary conditions, and discretised using 24×24 elements of degree $p = 3$ in the horizontal and 150 vertical levels, with a time step of $\Delta t = 0.01\text{s}$. No dissipation of any kind is applied in the vertical, or to the potential temperature advection equation. However the energy conserving upwinding is applied to the potential temperature diagnostic equation (62) in order to suppress the development of oscillations associated with non-hydrostatic motions. This is not applied to the baroclinic instability test case, which evolves in a predominantly hydrostatic regime.

Since the implicit vertical time stepping scheme conserves energy [9], and the energetic exchanges are exactly balanced in the horizontal, no stabilisation is required of the energetically balanced HEVI scheme, provided that the horizontal CFL condition is satisfied. This is in contrast to the TRAP(2,3,2) scheme, which requires a horizontal biharmonic viscosity in order to suppress the onset of numerical instability. In order to directly compare to the TRAP(2,3,2) scheme, simulations of the energetically balanced HEVI scheme have been run both with and without horizontal biharmonic viscosity on the momentum equation [11], with a coefficient of $624.78\text{m}^4\text{s}^{-1}$.

While the results qualitatively match those previously published in terms of both the position and deformation of the bubble [31, 32], in the absence of potential temperature upwinding a secondary oscillation develops in the wake of

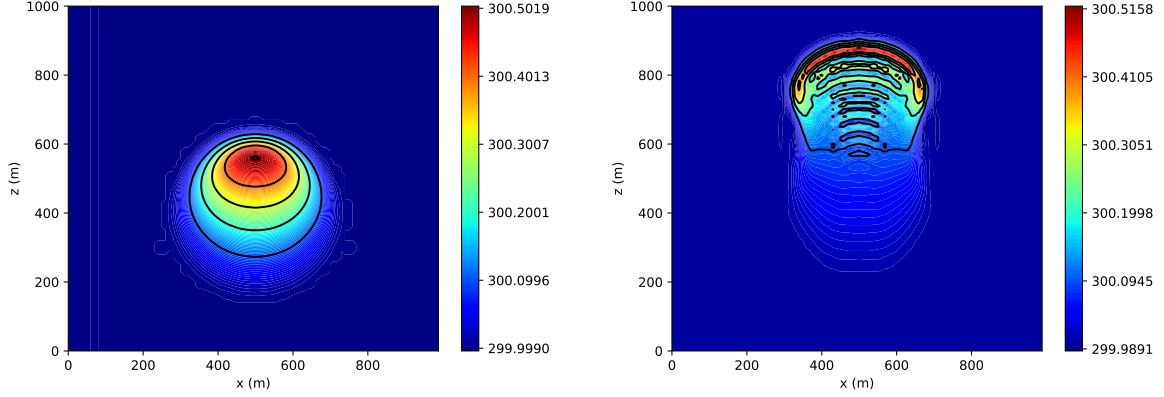


Figure 8: 3D rising bubble: Potential temperature cross section at $y = 0$ at times 200s (left) and 400s (right).

the bubble as it steepens, as shown in fig. 8. This oscillation is also observed for the TRAP (2,3,2) scheme [9]. The evolution of the energy and power exchanges (computed as in [8]) are shown in fig 9, and clearly show the signature of the high frequency internal gravity wave, as well as the increase in vertical kinetic energy and decrease in potential energy which represent the ascension of the bubble.

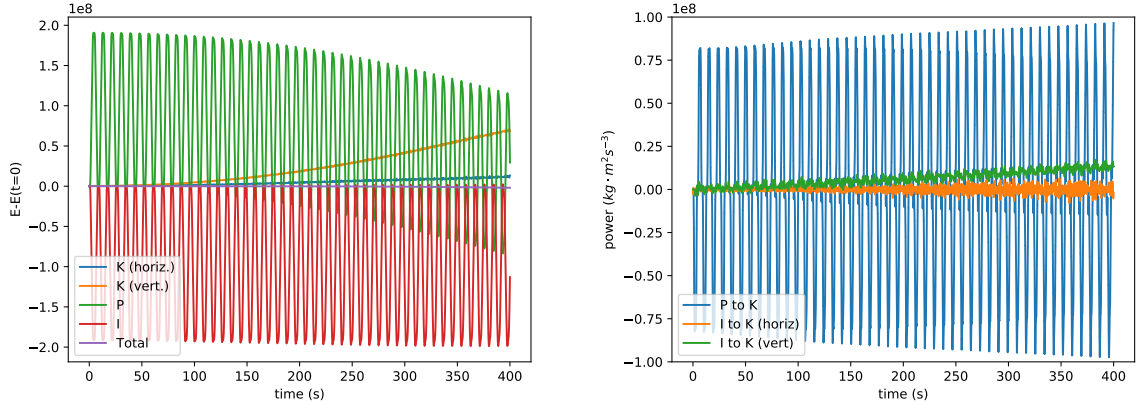


Figure 9: 3D rising bubble: Difference in energy from initial value (left), and power exchanges associated with the changes in potential to kinetic and internal to kinetic energy with time (right). The power exchanges are computed as the discrete analogues of the relations $g \int z \partial(\rho w) / \partial z d\Omega$ and $\int \nabla \cdot (\rho u \theta) \Pi d\Omega$ respectively.

5.2.1. Upwinding of potential temperature

In order to suppress the spurious potential temperature oscillations observed in fig. 8 in an energetically consistent manner, the upwinding formulation presented in section 4 is applied to the rising bubble test case. As observed in fig. 10, this upwinding formulation effectively clears up the spurious overshoots and oscillations. Since the upwinding is performed within the skew-symmetric operator, as detailed in (46), the energy conservation errors are almost identical to those for the original formulation, as shown in fig. 11.

Fig. 11 also shows the energy conservation errors for the TRAP(2,3,2) scheme (with biharmonic viscosity) and the energetically balanced scheme in the absence of viscosity. While horizontal biharmonic viscosity is necessary

to stabilise the TRAP(2,3,2) scheme, the energetically balanced scheme may be run entirely without dissipation. As observed, the difference in the energy conservation error of the energetically balanced scheme with and without biharmonic viscosity is small compared to the overall energy conservation error, indicating that most of the energy conservation error arises from the explicit time stepping of the horizontal velocity, and not the dissipation term. In all cases the total energy of the system decreases with time.

This figure also shows the power exchanges for the TRAP(2,3,2) scheme. These are almost identical to those for the energetically balanced scheme shown in fig. 9, such that the difference in the time scale of the internal gravity mode observed for the baroclinic test case is not observed for the high resolution rising bubble test case, for which the ratio of the horizontal and vertical resolutions is closer to unity, and the vertical time scales are effectively resolved by the time step.

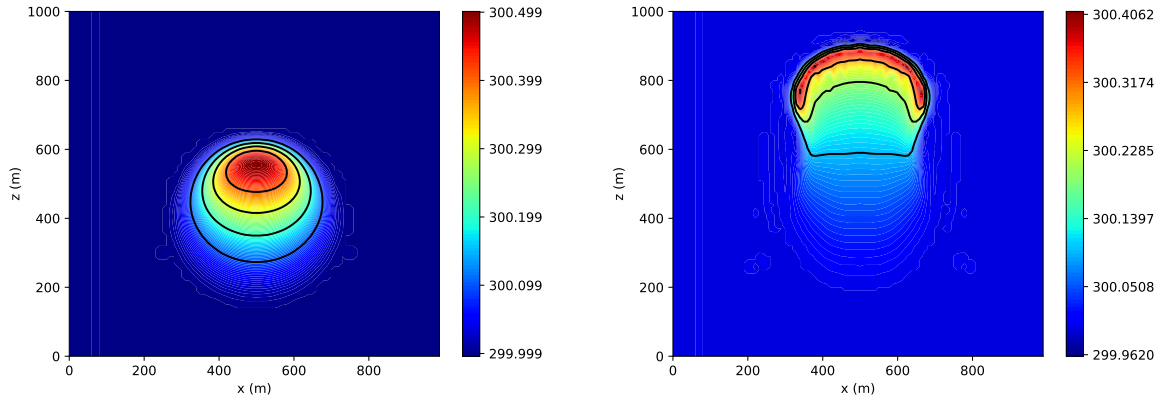


Figure 10: 3D rising bubble: Upwinded potential temperature cross section at $y = 0$ at times 200s (left) and 400s (right).

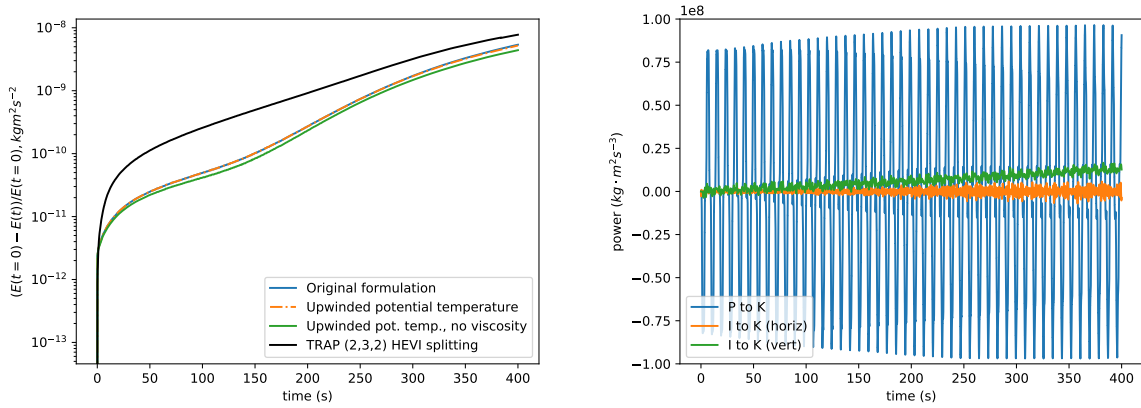


Figure 11: 3D rising bubble: Comparison of energy conservation errors for the original and upwinded potential temperature formulations, as well as the upwinded inviscid solution and the TRAP(2,3,2) scheme (left). Power exchanges associated with the changes in potential to kinetic and internal to kinetic energy with time for the TRAP (2,3,2) scheme (right).

6. Conclusions

A new horizontally explicit/vertically implicit time splitting scheme for non-hydrostatic atmospheric dynamics is introduced. The scheme allows for the exact balance of energetic exchanges when coupled with a spatial and vertical time integration scheme that preserve the skew-symmetric property of the Hamiltonian formulation of the equations of motion. However since the choice of horizontal velocity for which exchanges balance differs from the horizontal velocity at the end of the time step, the scheme still permits an energy conservation error. Linear eigenvalue analysis shows that the second order scheme improves upon the existing second order trapezoidal HEVI splitting by ensuring the neutral stability of all buoyancy modes, and all acoustic modes below a certain horizontal CFL number. Comparison to a horizontally third order, vertically second order trapezoidal splitting for the full 3D compressible Euler equations on the sphere shows that both schemes exhibit a secondary oscillation in the potential to kinetic power exchanges, associated with an internal gravity mode, however this oscillation has a temporal frequency approximately four times faster for the new energetically balanced scheme, such that the time scales of the oscillation for the trapezoidal scheme are closer to those of the baroclinic motions.

An energetically consistent formulation for potential temperature upwinding is also presented, by which the test functions are evaluated at vertically downstream locations. This ensures that spurious oscillations in potential temperature can be removed for vertical motions at non-hydrostatic regimes. This upwinding is applied within the skew-symmetric formulation, so as to suppress oscillations associated with vertical motions without altering the energetic properties of the model, for which the exchanges remain balanced.

Notably, for the rising bubble test case, which is run on an affine geometry, the energetically balanced HEVI scheme may be run stably without dissipation of any kind. For the baroclinic instability test case on the sphere however, horizontal biharmonic viscosity is still required in order to stabilise the simulation. Since the model uses inexact integration, and the sphere is represented as a non-affine geometry, it is possible that aliasing errors may provide an additional source of instability. As part of future work exact integration will be employed in order to determine if the scheme can be run entirely without dissipation for the baroclinic instability test case on the sphere also.

The code used in this article is publicly available at: <https://github.com/davelee2804/MiMSEM>.

Appendix A: Temporal discretisation of the energetically balanced HEVI scheme

For the benefit of readers who are not familiar with compatible finite element methods, in this appendix we present a formulation of the energetically balanced HEVI scheme that is discretised in the temporal dimension only. Note that in order to preserve energetic balance a spatial discretisation that respects the skew-symmetric structure of the equations of motion must also be employed. The temporal scheme is expressed as follows:

Step 1: Explicit horizontal momentum solve

$$\mathbf{v}' = \mathbf{v}^n - \Delta t \mathbf{q}_\perp^n \times \mathbf{V}^n + \Delta t \mathbf{q}_\parallel^n W^n - \Delta t \nabla_h \Phi^n - \Delta t \theta^n \nabla_h \Pi^n \quad (63)$$

Step 2: Implicit vertical solve (including horizontal divergence)

$$w^{n+1} = w^n - \Delta t \partial_z \bar{\Phi} - \Delta t \theta^{n+1/2} \partial_z \bar{\Pi} - \Delta t \mathbf{q}_\parallel^{n+1/2} \cdot \bar{\mathbf{V}} \quad (64a)$$

$$\rho^{n+1} = \rho^n - \Delta t \partial_z \bar{W} - \Delta t \nabla_h \cdot \bar{\mathbf{V}} \quad (64b)$$

$$\Theta^{n+1} = \rho^n - \Delta t \partial_z (\theta^{n+1/2} \bar{W}) - \Delta t \nabla_h \cdot (\theta^{n+1/2} \bar{\mathbf{V}}) \quad (64c)$$

with

$$\bar{V} = \frac{1}{6}(2\rho^n \mathbf{v}^n + \rho^{n+1} \mathbf{v}^n + \rho^n \mathbf{v}' + 2\rho^{n+1} \mathbf{v}') \quad (65a)$$

$$\bar{W} = \frac{1}{6}(2\rho^n w^n + \rho^{n+1} w^n + \rho^n w^{n+1} + 2\rho^{n+1} w^{n+1}) \quad (65b)$$

$$\bar{\Phi} = \frac{1}{6}(\mathbf{v}^n \mathbf{v}^n + \mathbf{v}^n \mathbf{v}' + \mathbf{v}' \mathbf{v}' + w^n w^n + w^n w^{n+1} + w^{n+1} w^{n+1}) + gz \quad (65c)$$

$$\bar{\Pi} = \frac{1}{2}(\Pi^n + \Pi^{n+1}) \quad (65d)$$

Step 3: Explicit horizontal momentum solve

$$\mathbf{v}^{n+1} = \mathbf{v}^n - \Delta t q_{\perp}^{n+1/2} \times \bar{\mathbf{V}} + \Delta t \mathbf{q}_{\parallel}^n \bar{W} - \Delta t \nabla_h \bar{\Phi} - \Delta t \theta^{n+1/2} \nabla_h \bar{\Pi} \quad (66)$$

where ∇_h is the horizontal gradient operator.

Note that if energetic balance is not of concern, a standard time centered construction of the mass flux terms, V , W and Bernoulli function, Φ at $n + 1/2$ should also preserve second order accuracy. Also note that a simple first order Euler integration is used here in step 1, which is sufficient to construct a second order mass flux and Bernoulli function.

Appendix B: Entropy conservation for the variational form of the compressible Euler equations

The time derivative of the internal energy is derived by multiplying the density-weighted potential temperature advection equation,

$$\frac{\partial \rho \theta}{\partial t} + \nabla \cdot (\rho \mathbf{u} \theta) = 0, \quad (67)$$

by the Exner pressure,

$$\frac{\delta I}{\delta \Theta} = \Pi = c_p \left(\frac{R \rho \theta}{p_0} \right)^{R/c_v}, \quad (68)$$

such that the internal energy, I evolves as

$$\frac{\partial I}{\partial t} = \left\langle \frac{\delta I}{\delta \Theta}, \frac{\partial \Theta}{\partial t} \right\rangle = \left\langle \Pi, \frac{\partial \rho \theta}{\partial t} \right\rangle = \frac{c_v}{c_p} \frac{\partial}{\partial t} \int \rho \theta \Pi d\Omega = -\langle \Pi, \nabla \cdot (\mathbf{u} \rho \theta) \rangle, \quad (69)$$

since $R = c_p - c_v$ and $\Theta = \rho \theta$. This expression may then be expanded as:

$$\frac{c_v}{c_p} \frac{\partial}{\partial t} \int \rho \theta \Pi d\Omega = -\langle \Pi, \theta \nabla \cdot (\rho \mathbf{u}) + \rho \mathbf{u} \cdot \nabla \theta \rangle \quad (70a)$$

$$= -\left\langle \rho \theta \Pi, \frac{1}{\rho} \nabla \cdot (\rho \mathbf{u}) + \frac{\mathbf{u}}{\theta} \cdot \nabla \theta \right\rangle \quad (70b)$$

$$= \left\langle \rho \theta \Pi, \frac{\partial \log(\rho)}{\partial t} - \mathbf{u} \cdot \nabla \log(\theta) \right\rangle \quad (70c)$$

The left hand side of this expression may be re-arranged as:

$$\frac{c_v}{c_p} \int \frac{\partial \rho \theta \Pi}{\partial t} d\Omega = \frac{c_v}{c_p} \left\langle \frac{\rho \theta \Pi}{\rho \theta \Pi}, \frac{\partial \rho \theta \Pi}{\partial t} \right\rangle = \frac{c_v}{c_p} \left\langle \rho \theta \Pi, \frac{\partial \log(\rho \theta \Pi)}{\partial t} \right\rangle = \left\langle \rho \theta \Pi, \frac{\partial \log(\rho)}{\partial t} - \mathbf{u} \cdot \nabla \log(\theta) \right\rangle \quad (71)$$

For $\rho \theta \Pi \in L^2(\Omega)$, cancellation of this factor from both sides yields the variational expression

$$\frac{c_v}{c_p} \left\langle \gamma, \frac{\partial \log(\rho \theta \Pi)}{\partial t} \right\rangle = \left\langle \gamma, \frac{\partial \log(\rho)}{\partial t} - \mathbf{u} \cdot \nabla \log(\theta) \right\rangle \quad \forall \gamma \in L^2(\Omega) \quad (72)$$

Recalling the equation of state gives

$$\rho\theta\Pi = c_p \left(\frac{R}{p_0} \right)^{R/c_v} (\rho\theta)^{c_p/c_v}, \quad (73)$$

$$\left\langle \gamma, \frac{\partial \log(\rho\theta)}{\partial t} \right\rangle = \left\langle \gamma, \frac{\partial \log(\rho)}{\partial t} - \mathbf{u} \cdot \nabla \log(\theta) \right\rangle \quad \forall \gamma \in L^2(\Omega) \quad (74)$$

which is further simplified as:

$$\left\langle \gamma, \frac{\partial \log(\theta)}{\partial t} + \mathbf{u} \cdot \nabla \log(\theta) \right\rangle = 0 \quad \forall \gamma \in L^2(\Omega) \quad (75)$$

Recalling the original definition of the entropy then gives

$$\left\langle \gamma, \frac{\partial s}{\partial t} + \mathbf{u} \cdot \nabla s \right\rangle = 0 \quad \forall \gamma \in L^2(\Omega) \quad (76)$$

such that entropy is materially conserved. Setting the test function as the density $\gamma = \rho$ in the above expression, and multiplying the continuity equation by the entropy, we have

$$\left\langle \rho, \frac{\partial s}{\partial t} + \mathbf{u} \cdot \nabla s \right\rangle = 0 \quad (77a)$$

$$\left\langle s, \frac{\partial \rho}{\partial t} + \nabla \cdot \mathbf{u}\rho \right\rangle = 0. \quad (77b)$$

Adding this to the continuity equation, and recalling that $\mathbf{U} := \rho\mathbf{u}$ gives

$$\int \frac{\partial \rho s}{\partial t} + \nabla \cdot (\mathbf{u}\rho s) d\Omega = 0, \quad (78)$$

which yields an additional conservation law for the entropy function-entropy flux pair $\rho s, \mathbf{u}\rho s$ [33].

7. Acknowledgments

David Lee would like to thank Dr Justin Freeman for his continued encouragement, and Drs. Marcus Thatcher and John McGregor for their support and access to computing resources. This project was supported by resources and expertise provided by CSIRO IMT Scientific Computing. We are also grateful to the two anonymous reviewers, whose comments and insights helped to improve the quality of this article.

References

- [1] P. Ullrich, C. Jablonowski, Operator-split Runge–Kutta–Rosenbrock methods for nonhydrostatic atmospheric models, *Mon. Wea. Rev.* 140 (2012) 1257–1284.
- [2] F. X. Giraldo, J. F. Kelly, E. M. Constantinescu, Implicit-explicit formulations of a three-dimensional nonhydrostatic unified model of the atmosphere (NUMA), *SIAM J. Sci. Comput.* 35 (2013) B1162–B1194.
- [3] L. Bao, R. Klöforn, R. D. Nair, Horizontally explicit and vertically implicit (HEVI) time discretization scheme for a discontinuous Galerkin nonhydrostatic model, *Mon. Wea. Rev.* 143 (2015) 972–990.
- [4] D. J. Gardner, J. E. Guerra, F. P. Hamon, D. R. Reynolds, P. A. Ullrich, C. S. Woodward, Implicit-explicit (IMEX) Runge–Kutta methods for non-hydrostatic atmospheric models, *Geosci. Model Dev.* 11 (2018) 1497–1515.
- [5] A. Steyer, C. J. Vogl, M. Taylor, O. Guba, Efficient IMEX Runge–Kutta Methods for Nonhydrostatic dynamics, *arXiv:1906.07219* (2019)
- [6] H. Weller, S. J. Lock, N. Wood, Runge–Kutta IMEX schemes for the Horizontally Explicit/Vertically Implicit (HEVI) solution of the wave equations, *J. Comp. Phys.* 252 (2013) 365–381.
- [7] S.-J. Lock, N. Wood, H. Weller, Numerical analyses of Runge–Kutta implicit–explicit schemes for horizontally explicit, vertically implicit solutions of atmospheric models, *Q. J. R. Meteorol. Soc.* 140 (2014) 1654–1669.
- [8] D. Lee, A. Palha, A mixed mimetic spectral element model of the 3D compressible Euler equations on the cubed sphere, *J. Comp. Phys.* 401 (2020) 108993.

- [9] D. Lee, An energetically balanced, quasi-Newton integrator for non-hydrostatic vertical atmospheric dynamics, *J. Comp. Phys.* (2021) 109988.
- [10] D. Lee, A. Palha, M. Gerritsma, Discrete conservation properties for shallow water flows using mixed mimetic spectral elements, *J. Comp. Phys.* 357 (2018) 282–304.
- [11] D. Lee, A. Palha, A mixed mimetic spectral element model of the rotating shallow water equations on the cubed sphere, *J. Comp. Phys.* 375 (2018) 240–262.
- [12] M. Gerritsma, Edge functions for spectral element methods, in: *Spectral and high order methods for partial differential equations*, Lecture Notes in Computational Science and Engineering, Springer 76 (2011) 199–207.
- [13] J. Kreeft, M. Gerritsma, Mixed mimetic spectral element method for Stokes flow: A pointwise divergence-free solution, *J. Comp. Phys.* 240 (2013) 284–309.
- [14] A. Palha, P. P. Rebelo, R. Hiemstra, J. Kreeft, M. Gerritsma, Physics-compatible discretization techniques on single and dual grids, with application to the Poisson equation of volume forms, *J. Comp. Phys.* 257 (2014) 1394–1422.
- [15] R. Hiemstra, D. Toshniwal, R. H. M. Huijsmans, M. I. Gerritsma, High order geometric methods with exact conservation properties, *J. Comp. Phys.* 257 (2014) 1444–1471.
- [16] V. Jain, Y. Zhang, A. Palha, M. Gerritsma, Construction and application of algebraic dual polynomial representations for finite element methods on quadrilateral and hexahedral meshes, *Comput. Math. Appl.* in press (2020).
- [17] M. Gerritsma, P. Bochev, A spectral mimetic least-squares method for the Stokes equations with no-slip boundary conditions, *Comput. Math. Appl.* 71 (2016) 2285–2300.
- [18] E. Hairer, Energy-Preserving Variant of Collocation Methods, *JNAIAM J. Numer. Anal. Ind. Appl. Math.* 5 (2010) 7–84.
- [19] D. Cohen, E. Hairer, Linear energy-preserving integrators for Poisson systems, *BIT Numer. Math.* 51 (1) (2011) 91–101.
- [20] T. Melvin, T. Benacchio, J. Thuburn, C. Cotter, Choice of function spaces for thermodynamic variables in mixed finite-element methods, *Q. J. Royal. Meteorol. Soc.* 144 (2018) 900–916.
- [21] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, K. Rupp, B. F. Smith, S. Zampini, H. Zhang, PETSc Web page, <http://www.mcs.anl.gov/petsc> (2017).
- [22] S. Balay, S. Abhyankar, M. F. Adams, J. Brown, P. Brune, K. Buschelman, L. Dalcin, V. Eijkhout, W. D. Gropp, D. Kaushik, M. G. Knepley, D. A. May, L. C. McInnes, K. Rupp, P. Sanan, B. F. Smith, S. Zampini, H. Zhang, PETSc users manual, Tech. Rep. ANL-95/11 - Revision 3.8, Argonne National Laboratory (2017).
- [23] S. Balay, W. D. Gropp, L. C. McInnes, B. F. Smith, Efficient management of parallelism in object oriented numerical software libraries, in: E. Arge, A. M. Bruaset, H. P. Langtangen (Eds.), *Modern Software Tools in Scientific Computing*, Birkhäuser Press, 1997, pp. 163–202.
- [24] D. Durran, *Numerical Methods for Fluid Dynamics: With Applications in Geophysics*, Springer (2010) 516pp.
- [25] R. Sadourny, C. Basdevant, Parameterization of Subgrid Scale Barotropic and Baroclinic Eddies in Quasi-geostrophic Models: Anticipated Potential Vorticity Method, *J. Atmos. Sci.* 42 (1985) 1353–1363.
- [26] A. T. T. McRae, C. J. Cotter, Energy- and enstrophy conserving schemes for the shallow water equations, based on mimetic finite elements, *Q. J. R. Meteorol. Soc.* 140 (2014) 2223–2234.
- [27] D. Lee, Petrov-Galerkin flux upwinding for mixed mimetic spectral elements, and its application to geophysical flow problems, *Comput. Math. Appl.* 89 (2021) 68–77.
- [28] G. A. Wimmer, C.J. Cotter, W. Bauer, Energy conserving SUPG methods for compatible finite element schemes in numerical weather prediction, *arXiv:2001.09590* (2020).
- [29] A. N. Brooks, T. J. R. Hughes, Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations, *Comput. Meth. Appl. Mech. Engrg.* 32 (1982) 199–259.
- [30] P. A. Ullrich, T. Melvin, C. Jablonowski, A. Staniforth, A proposed baroclinic wave test case for deep- and shallow-atmosphere dynamical cores, *Q. J. R. Meteorol. Soc.* 140 (2014) 1590–1602.
- [31] D. S. Abdi, F. X. Giraldo, Efficient construction of unified continuous and discontinuous Galerkin formulations for the 3D Euler equations, *J. Comp. Phys.* 320 (2016) 46–68.
- [32] T. Melvin, T. Benacchio, B. Shipway, N. Wood, J. Thuburn, C. Cotter, A mixed finite-element, finite-volume, semi-implicit discretisation for atmospheric dynamics: Cartesian geometry, *Q. J. R. Meteorol. Soc.* (2019) 1–19.
- [33] E. Tadmor, Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems, *Acta Numerica* 12 (2003) 451–512.