Entropy-stable discontinuous spectral-element methods for the spherical shallow water equations in covariant form

Tristan Montoya[a,b,*], Andrés M. Rueda-Ramírez[c], Gregor J. Gassner[a,d]

[a]*Department of Mathematics and Computer Science, University of Cologne, Germany*
[b]*Department of Computer Science, University of Saskatchewan, Canada*
[c]*School of Aeronautics (ETSIAE), Universidad Politécnica de Madrid, Spain*
[d]*Center for Data and Simulation Science, University of Cologne, Germany*

Abstract

We introduce discontinuous spectral-element methods of arbitrary order that are well balanced, conservative of mass, and conservative or dissipative of total energy (i.e., a mathematical entropy function) for a covariant flux formulation of the rotating shallow water equations with variable bottom topography on curved manifolds such as the sphere. The proposed methods are based on a skew-symmetric splitting of the tensor divergence in covariant form, which we implement and analyze within a general flux-differencing framework using tensor-product summation-by-parts operators. Such schemes are proven to satisfy semi-discrete mass and energy conservation on general unstructured quadrilateral grids in addition to well balancing for arbitrary continuous bottom topographies, with energy dissipation resulting from a suitable choice of numerical interface flux. Furthermore, the proposed covariant formulation permits an analytical representation of the geometry and associated metric terms while satisfying the aforementioned entropy stability, conservation, and well-balancing properties without the need to approximate the metric terms so as to enforce discrete metric identities. Numerical experiments on cubed-sphere grids are presented in order to verify the schemes' structure-preservation properties as well as to assess their accuracy and robustness within the context of several standard test cases characteristic of idealized atmospheric flows. Our theoretical and numerical results support the further development of the proposed methodology towards a full dynamical core for numerical weather prediction and climate modelling, as well as broader applications to other hyperbolic and advection-dominated systems of partial differential equations on curved manifolds.

Keywords: Entropy stability, summation-by-parts, discontinuous Galerkin, geophysical fluid dynamics, hyperbolic partial differential equations, manifolds
2020 MSC: 65M12, 65M60, 65M70, 58J45, 76U60

1. Introduction

The development of efficient, robust, and flexible numerical methods for solving the shallow water equations in spherical geometry constitutes an important early step in the development of a new dynamical core for the atmosphere or ocean component of a global weather or climate model,

---

[*]Corresponding author, ORCiD: 0000-0002-4259-1449.
Email address: tristan.montoya@usask.ca (Tristan Montoya)

with verification through shallow water test cases (for example, those proposed by Williamson *et al.* [1], Galewsky *et al.* [2], and Läuter *et al.* [3]) serving as an important precursor to the solution of more complex geophysical flow problems. Consisting of a scalar equation for mass conservation and a vector equation for momentum conservation, the spherical shallow water equations also represent a prototypical example of a nonlinear hyperbolic system of balance laws on a curved manifold, thereby facilitating the development of methods and algorithms applicable to a broad class of similar partial differential equations arising in various scientific and engineering disciplines. Offering excellent parallel scalability and geometric flexibility as well as the ability to achieve low dissipative and dispersive errors on relatively coarse grids, high-order continuous and discontinuous spectral-element methods represent attractive choices for solving such systems of equations and have therefore been a major focus of recent advances in the development of next-generation numerical methods for global weather and climate models.

Foundational contributions to the development of continuous Galerkin (CG) methods (see, for example, Taylor *et al.* [4], Giraldo [5], and Thomas and Loft [6]) and discontinuous Galerkin (DG) methods (see, for example, Giraldo *et al.* [7] and Nair *et al.* [8]) for the spherical shallow water equations have since been extended to three-dimensional nonhydrostatic atmospheric models based on spectral-element dynamical cores[1] at major modelling centres worldwide (see, for example, the review by Marras *et al.* [9] and references therein). A major barrier to the widespread adoption of spectral-element approaches, and high-order methods more broadly, is that their efficiency often comes at the price of robustness for nonlinear problems as a result of their inherently low numerical dissipation, which renders such methods susceptible to aliasing-driven instabilities and nonphysical oscillations introduced by under-resolved high-frequency components of the numerical solution. As discussed in [9, Section 4], continuous as well as discontinuous spectral-element methods typically achieve robustness in practice through the use of *ad hoc* stabilization techniques such as modal filtering or artificial viscosity, which must be tuned carefully in order to avoid introducing excessive dissipative error into the numerical solution. Alternatively, overintegration techniques involving the use of a larger number of quadrature points than required for a linear problem (see, for example, Kirby and Karniadakis [10] and Mengaldo *et al.* [11]) can be used to reduce aliasing for nonlinear problems. However, overintegration significantly increases a scheme's computational expense, does not generally provide any theoretical assurance of stability, and has been shown to be insufficient in certain cases when applied to under-resolved discretizations of compressible flows (see, for example, Gassner *et al.* [12] and Winters *et al.* [13]).

Tracing their origins to the early work of Kreiss and Scherer [14] and Tadmor [15] as well as more recent developments by LeFloch *et al.* [16], Fisher [17], and Gassner [18], modern structure-preserving numerical methods based on summation-by-parts (SBP) operators have evolved into a powerful algebraic framework for the construction of schemes which overcome the robustness issues traditionally afflicting high-order discretizations. These developments are reviewed within the context of discontinuous spectral-element methods applied to computational fluid dynamics by Gassner and Winters [19], with the two key components being a discrete derivative operator which mimics integration by parts (i.e., satisfies the SBP property) and a modified approximation of the flux divergence based on a consistent skew-symmetric splitting or a *flux-differencing* formulation based

---

[1]In this paper, we use the term *spectral-element method* in a general sense to denote any high-order element-based discretization, whether or not it is derived from a standard Galerkin procedure, and whether or not a collocated nodal basis is used. This family of methods includes standard nodal and modal DG and CG schemes as special cases.

on specially chosen finite-volume-like fluxes coupling pairs of quadrature points. Such two-point flux functions can be tailored so as to reduce aliasing as well as preserve certain properties of the continuous problem, including conservation or dissipation of mathematical entropy (a strictly convex function generalizing the notion of thermodynamic entropy, corresponding, for example, to the total energy for a shallow water system).

Although entropy-stable discretizations have recently been developed for two-dimensional geophysical flow models based on the vector-invariant form of the spherical shallow water equations (see, for example, Ricardo *et al.* [20]), as well as for three-dimensional atmospheric models based on a Cartesian flux formulation (see, for example, Waruszewski *et al.* [21]), to the authors' knowledge, the use of a *covariant* flux formulation has not yet been explored within a split-form or flux-differencing framework. In the context of fluid dynamics in curved geometry, the covariant formulation involves the evolution of the contravariant momentum components as prognostic variables and the formulation of the governing equations as a system of balance laws in which divergences of tensor-valued flux functions are formulated in terms of the covariant derivative. This allows for discretizations to be constructed directly on curved manifolds, rather than within a higher-dimensional ambient space, for example, solving systems on two-dimensional surfaces embedded in three-dimensional space using two, rather than three, momentum equations. When extending such formulations to three-dimensional atmospheric models, the covariant form then allows for separation of the horizontal dynamics from the vertical dynamics, where the latter require special numerical treatment due to the finer mesh spacing and fast wave speeds (see, for example, Baldauf [22]). Unlike the vector-invariant form, the covariant form allows for a locally conservative treatment of the momentum balance, up to the terms introduced by the manifold's curvature, and does not require a specialized discretization of the vorticity. These considerations, as well as the fact that discontinuous spectral-element methods for nonhydrostatic atmospheric models in covariant form are the focus of several dynamical cores currently under development for large-scale weather and climate models [22–25], motivate the development of entropy-stable covariant formulations as a key contribution towards improving the robustness of atmospheric simulations.

In this paper, we address the above objective of extending the SBP framework for split-form and entropy-stable discretizations to nonlinear systems of balance laws in covariant form on curved manifolds. Specifically, we introduce discontinuous spectral-element methods for the covariant shallow water equations with variable bottom topography that achieve arbitrary-order accuracy on general unstructured quadrilateral meshes and satisfy semi-discrete conservation of mass, semi-discrete conservation or dissipation of total energy, and well balancing for a constant surface height and zero velocity. This is achieved through a novel flux-differencing discretization of the covariant derivative that generalizes the skew-symmetric split formulation developed by Gassner *et al.* [26] and Wintermeyer *et al.* [27] to curved manifolds, combined with tensor-product spectral-element operators with the SBP property, a suitable treatment of the nonconservative bottom-topography terms, and an entropy-stable interface flux between adjacent elements.

We now outline the contents and contributions of the remainder of this paper. In Section 2, we focus on the continuous problem, introducing the necessary notation and definitions for working with vectors and tensors on manifolds, the formulation and entropy analysis of the shallow water equations in covariant form, and the derivation of a novel skew-symmetric split form for the covariant derivative which will serve as the starting point for constructing our discretizations. In Section 3, we present the major components of the schemes, including the mapping from reference to physical space, the tensor-product SBP operators, and the flux-differencing discontinuous spectral-element

formulation. In Section 4, we introduce entropy-conservative and entropy-stable two-point flux functions based on our proposed skew-symmetric splitting and show that the resulting schemes are consistent and satisfy the aforementioned conservation, entropy stability, and well-balancing properties. Numerical experiments for an unsteady analytical solution [3, Example 3], the standard isolated mountain test described in [1, Case 5], the barotropic instability problem from [2], and the Rossby–Haurwitz wave problem described in [1, Case 6] are presented in Section 5, demonstrating the accuracy, structure-preserving properties, and robustness of the proposed schemes for idealized atmospheric test cases on cubed-sphere grids.

2. Shallow water equations on manifolds

In this section, we will discuss the formulation and analysis of balance laws on manifolds, with a specific focus on the covariant shallow water equations on general two-dimensional surfaces.

2.1. Vector and tensor notation

We begin by introducing some standard notation which will be used in the formulation of balance laws on manifolds. In particular, we will consider the case of a two-dimensional surface $S \subset \mathbb{R}^3$ embedded within a three-dimensional ambient space, which we take to be spanned by the Cartesian unit basis vectors $\vec{e}_x$, $\vec{e}_y$, and $\vec{e}_z$, and equipped with the standard Euclidean dot product and norm. Let us now define $\vec{X} \colon D \to E$ to be a smooth, bijective, and time-invariant mapping from the standard reference domain $D \subset \mathbb{R}^2$ to the physical domain $E \subset S$, which corresponds to an element of a non-overlapping tessellation (i.e., a mesh) of S, as illustrated in Figure 1. The space of tangent vectors at a point on E is spanned by the *covariant basis vectors*, which are given by

$$\vec{a}_1 := \partial_1 \vec{X}, \quad \vec{a}_2 := \partial_2 \vec{X}, \tag{1}$$

where we use ∂_1 and ∂_2 to denote partial derivatives with respect to the coordinates ξ^1 and ξ^2 on the reference domain D, which correspond to chart coordinates on S induced by the mapping $\vec{X}$. The covariant basis vectors in (1) can be used to obtain the *covariant components* of the metric tensor as $G_{ij} := \vec{a}_i \cdot \vec{a}_j$, which constitute the entries of a symmetric positive-definite (SPD) matrix

$$\boldsymbol{G} := \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix} \tag{2}$$

characterizing the geometry of the manifold at a given point. The entries of the inverse matrix $\boldsymbol{G}^{-1}$ are likewise known as the *contravariant components* of the metric tensor, which we denote by G^{ij}.

Employing the summation convention throughout this work for Latin indices repeated across upper and lower positions, any tangent vector to S can be represented as $\vec{v} = v^i \vec{a}_i = v^1 \vec{a}_1 + v^2 \vec{a}_2$ in terms of its contravariant components v^i with respect to the covariant basis vectors $\vec{a}_i$, or equivalently as $\vec{v} = v_i \vec{a}^i = v_1 \vec{a}^1 + v_2 \vec{a}^2$ in terms of its covariant components v_i with respect to the contravariant basis vectors $\vec{a}^i$. Such components and bases are related by raising and lowering indices as

$$v_i = G_{ij} v^j \iff v^i = G^{ij} v_j, \quad \vec{a}_i = G_{ij} \vec{a}^j \iff \vec{a}^i = G^{ij} \vec{a}_j, \tag{3}$$

such that the Euclidean dot product and norm can be computed as $\vec{v} \cdot \vec{v} = v_i v^i$ and $\|\vec{v}\| = \sqrt{v_i v^i}$, respectively. Tensors of second order or higher can be expanded in terms of tensor products of basis
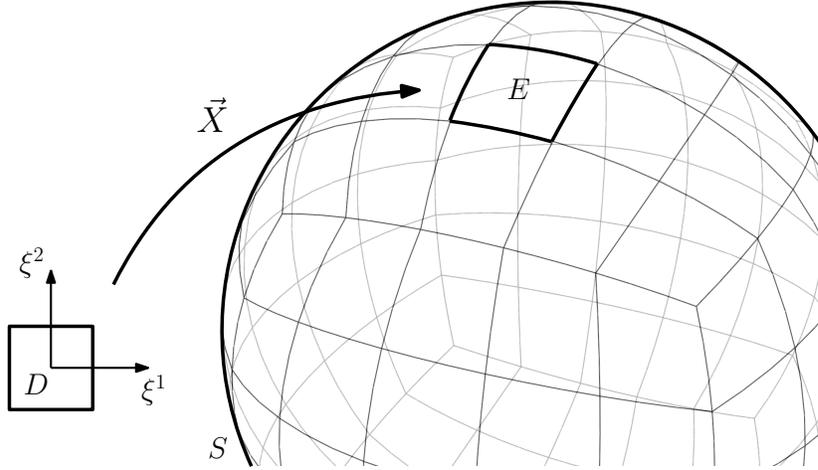
4

Figure 1: Illustration of the surface S, reference domain D, and physical element E

vectors, for example, as $\overset{\leftrightarrow}{\tau} = \tau^{ij}\vec{a}_i \otimes \vec{a}_j$, where the rules in (3) for raising and lowering indices can be used similarly to convert between covariant, contravariant, and mixed tensor representations.

In order to differentiate a vector field $\vec{v}$ or second-order tensor field $\overset{\leftrightarrow}{\tau}$ along a spatially varying basis vector $\vec{a}_k$ given as in (1), we require the *covariant derivative*, which is a tensor field with components given by

$$\nabla_k v^i := \partial_k v^i + \Gamma^i_{jk} v^j, \quad \nabla_k \tau^{ij} := \partial_k \tau^{ij} + \Gamma^i_{kl} \tau^{lj} + \Gamma^j_{kl} \tau^{il}, \tag{4}$$

where the *Christoffel symbols of the second kind* are given by

$$\Gamma^i_{jk} := \frac{1}{2} G^{il} \left(\partial_j G_{kl} + \partial_k G_{jl} - \partial_l G_{jk} \right), \tag{5}$$

and, following a standard abuse of notation, we emphasize that the operator ∇_k should be taken as applying to the entire vector field $\vec{v}$ or tensor field $\overset{\leftrightarrow}{\tau}$ rather than to a single component v^i or τ^{ij}. By construction, the covariant derivative satisfies the product rule as well as the *metric compatibility* conditions

$$\nabla_k G^{ij} = 0, \tag{6}$$

which can be derived in a straightforward manner from (4) and (5). The integral of a function $h \colon E \to \mathbb{R}$ can be computed in reference coordinates under the change of variables

$$\int_E h \, \mathrm{d}S = \int_D (h \circ \vec{X}) \sqrt{\det \boldsymbol{G}} \, \mathrm{d}\xi^1 \mathrm{d}\xi^2, \tag{7}$$

where, for convenience of notation, we will define $J := \sqrt{\det \boldsymbol{G}}$ and suppress composition with $\vec{X}$ when it is unlikely to cause ambiguity. The reader is referred to fluid mechanics texts such as Aris [28] or general relativity texts such as Misner *et al.* [29] for a more complete introduction to the basic concepts of tensor calculus and differential geometry within the context of continuum physics.

2.2. Shallow water equations in covariant form

Denoting the depth of the fluid layer as h and the velocity vector field as $\vec{v} = v^i \vec{a}_i$, the shallow water equations can be formulated on a two-dimensional manifold $S \subset \mathbb{R}^3$ as

$$\partial_t h + \nabla_j (hv^j) = 0, \tag{8a}$$

$$\partial_t (hv^i) + \nabla_j \tau^{ij} = -fJG^{ij} \varepsilon_{jk} hv^k - ghG^{ij} \partial_j b, \tag{8b}$$

where $t \geq 0$ is the time coordinate, $g > 0$ is the constant gravitational acceleration, b is the spatially varying bottom topography, f is the spatially varying Coriolis parameter, and ε_{jk} is the Levi-Civita symbol, which takes values of $\varepsilon_{12} = 1$, $\varepsilon_{21} = -1$, and $\varepsilon_{11} = \varepsilon_{22} = 0$. The contravariant components of the momentum flux tensor are given by

$$\tau^{ij} := hv^i v^j + \frac{g}{2} h^2 G^{ij}. \tag{9}$$

Using the definitions in (4) as well as the product rule and the repeated-index identity $\Gamma^i_{ik} = (\partial_k J)/J$, we can express the divergence of the mass and momentum fluxes as

$$\nabla_j (hv^j) = \frac{1}{J} \partial_j (Jhv^j), \quad \nabla_j \tau^{ij} = \frac{1}{J} \partial_j (J\tau^{ij}) + \Gamma^i_{jk} \tau^{jk}, \tag{10}$$

such that the continuity and momentum equations in (8a) and (8b), respectively, can be rewritten using partial derivatives rather than covariant derivatives:

$$\partial_t h + \frac{1}{J} \partial_j (Jhv^j) = 0, \tag{11a}$$

$$\partial_t (hv^i) + \frac{1}{J} \partial_j (J\tau^{ij}) = -fJG^{ij} \varepsilon_{jk} hv^k - ghG^{ij} \partial_j b - \Gamma^i_{jk} \tau^{jk}. \tag{11b}$$

The resulting system of balance laws can then be expressed as

$$\partial_t \boldsymbol{u} + \frac{1}{J} \partial_1 (J\boldsymbol{f}^1) + \frac{1}{J} \partial_2 (J\boldsymbol{f}^2) = \boldsymbol{s}_{\text{cor}} + \boldsymbol{s}_{\text{bot}} + \boldsymbol{s}_{\text{geo}}, \tag{12}$$

where the state vector and contravariant flux components are given by

$$\boldsymbol{u} := \begin{pmatrix} h \\ hv^1 \\ hv^2 \end{pmatrix}, \quad \boldsymbol{f}^1 := \begin{pmatrix} hv^1 \\ hv^1 v^1 + \frac{g}{2} h^2 G^{11} \\ hv^2 v^1 + \frac{g}{2} h^2 G^{21} \end{pmatrix}, \quad \boldsymbol{f}^2 := \begin{pmatrix} hv^2 \\ hv^1 v^2 + \frac{g}{2} h^2 G^{12} \\ hv^2 v^2 + \frac{g}{2} h^2 G^{22} \end{pmatrix}, \tag{13}$$

and the Coriolis, bottom topography, and geometric source terms are given, respectively, by

$$\boldsymbol{s}_{\text{cor}} := \begin{pmatrix} 0 \\ -fJG^{1j} \varepsilon_{jk} hv^k \\ -fJG^{2j} \varepsilon_{jk} hv^k \end{pmatrix}, \quad \boldsymbol{s}_{\text{bot}} := \begin{pmatrix} 0 \\ -ghG^{1j} \partial_j b \\ -ghG^{2j} \partial_j b \end{pmatrix}, \quad \boldsymbol{s}_{\text{geo}} := \begin{pmatrix} 0 \\ -\Gamma^1_{jk} \tau^{jk} \\ -\Gamma^2_{jk} \tau^{jk} \end{pmatrix}. \tag{14}$$

This form provides a natural starting point for the construction of DG (see, for example, Läuter et al. [30], Bao et al. [31], Kuang et al. [32], and Baldauf [23]) or finite-volume methods (see, for example, Rossmanith et al. [33] and Ullrich et al. [34]).

2.3. Entropy analysis

It is well known that continuously differentiable solutions to the shallow water equations satisfy an additional conservation law for the total energy per unit area $\eta := h v_i v^i / 2 + g h (h + b) / 2$, which serves as a mathematical *entropy function* for the system (see, for example, Tadmor [35]). Taking the contravariant *entropy flux* components to be $F^j := h v_i v^i v^j / 2 + g h^2 v^j + g h b v^j$, such a conservation law is given in covariant form as

$$\partial_t \eta + \nabla_j F^j = 0. \tag{15}$$

The corresponding *entropy variables* are then obtained by differentiating the entropy function with respect to the state variables as

$$\boldsymbol{w} := \frac{\partial \eta}{\partial \boldsymbol{u}} = \begin{pmatrix} g(h+b) - \frac{1}{2}(v_1 v^1 + v_2 v^2) \\ v_1 \\ v_2 \end{pmatrix}. \tag{16}$$

While the development of a general theory for entropy analysis of hyperbolic systems on manifolds is beyond the scope of this paper (although we note that the scalar case has been treated, for example, by Ben-Artzi and LeFloch [36]), we will nevertheless verify that, under suitable assumptions, η is strictly convex with respect to the state vector consisting of the fluid layer depth and contravariant momentum components, as required for a mathematical entropy function.

Lemma 1. *The mapping from $\boldsymbol{u}$ to η is strictly convex when $h > 0$ and $\boldsymbol{G}$ is SPD.*

Proof. Differentiating the entropy variables with respect to the conservative variables, we obtain the Hessian of the entropy function as

$$\frac{\partial \boldsymbol{w}}{\partial \boldsymbol{u}} = \frac{1}{h} \begin{pmatrix} gh + v_1 v^1 + v_2 v^2 & -v_1 & -v_2 \\ -v_1 & G_{11} & G_{12} \\ -v_2 & G_{21} & G_{22} \end{pmatrix}. \tag{17}$$

The lower-right two-by-two block of the above matrix is given by $\boldsymbol{G}/h$, which is SPD under the present assumptions, and its Schur complement corresponds to the positive scalar g. The matrix in (17) is therefore also SPD, which is equivalent to the strict convexity of η as a function of $\boldsymbol{u}$. $\qquad\square$

Remark 1. Due to the role of the total energy as a mathematical entropy function, we will often refer to the quantity η as the "entropy" for consistency with the literature on entropy-stable schemes. We will also assume throughout our analysis that the conditions of Lemma 1 are satisfied regarding the positivity of the fluid layer depth and the positive-definiteness of the metric.

Whereas continuously differentiable solutions conserve the integral of η within the domain S, physically admissible weak solutions satisfy an entropy inequality of the form

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_S \eta \, \mathrm{d}S \leq 0. \tag{18}$$

Since establishing such integral balances for a system such as (11) requires the use of the product and chain rules in space, which are not guaranteed to hold discretely, we will begin our construction of an entropy-stable discretization with the introduction of a skew-symmetric splitting which allows for the product and chain rules to be circumvented.

2.4. Skew-symmetric formulation

Building upon the work of Gassner *et al.* [26] and Wintermeyer *et al.* [27] within the context of the shallow water equations in planar geometry, we prove the following lemma to obtain a skew-symmetric splitting of the covariant-form divergence of the momentum flux.

Lemma 2. *For smooth solutions, the divergence of the momentum flux in* (9) *can be expressed as*

$$
\nabla_j \tau^{ij} = \frac{1}{2J}\left(\partial_j(Jhv^iv^j) + v^i\partial_j(Jhv^j) + G^{ik}Jhv^j\partial_j v_k\right) + gG^{ij}h\partial_j h
$$
$$
+ \frac{1}{2}\left(\Gamma^i_{jk}hv^jv^k - G^{ik}\Gamma^l_{jk}hv^jv_l\right). \tag{19}
$$

Proof. Distributing the covariant derivative across the convective and pressure terms of (9) and noting that the covariant derivative of a scalar reduces to the partial derivative, we can use the product rule and chain rule to obtain

$$
\nabla_j \tau^{ij} = \nabla_j(hv^iv^j) + \frac{1}{2}gh^2\nabla_j G^{ij} + ghG^{ij}\partial_j h, \tag{20}
$$

The first term on the right-hand side can then be expressed in two different ways as

$$
\nabla_j(hv^iv^j) = \frac{1}{J}\partial_j(Jhv^iv^j) + \Gamma^i_{jk}hv^jv^k
$$
$$
= \frac{1}{J}\left(v^i\partial_j(Jhv^j) + G^{ik}Jhv^j\partial_j v_k - G^{ik}\Gamma^l_{jk}hv^jv_l\right), \tag{21}
$$

where the first line results from the use of the second identity in (10), while the second line results from the first identity in (10) and several applications of the product rule. Averaging the two expressions and substituting the result into (20), we then note that the second term on the right-hand side of (20) is zero due to the metric compatibility condition in (6). Grouping the differential and algebraic terms therefore results in (19). □

When contracting (19) with the covariant velocity components and manipulating the indices using (3), the geometric source terms cancel as

$$
v_i(\Gamma^i_{jk}hv^jv^k - G^{ik}\Gamma^l_{jk}hv^jv_l) = v_i\Gamma^i_{jk}hv^jv^k - v^k\Gamma^l_{jk}hv^jv_l = 0, \tag{22}
$$

resulting in an expression analogous to that obtained in flat (i.e., Euclidean) space:

$$
v_i\nabla_j \tau^{ij} = \frac{1}{2J}\left(v_i\partial_j(Jhv^iv^j) + v_iv^i\partial_j(Jhv^j) + Jhv^iv^j\partial_j v_i\right) + v^jgh\partial_j h, \tag{23}
$$

allowing for an integral entropy balance to be derived purely using integration by parts in reference space, a procedure which, as we will see in the following section, is straightforward to mimic at the semi-discrete level. We therefore rewrite the momentum equations in (11b) as

$$
\partial_t(hv^i) + \frac{1}{2J}\left(\partial_j(Jhv^iv^j) + v^i\partial_j(Jhv^j) + G^{ik}Jhv^j\partial_j v_k\right) + gG^{ij}h\partial_j(h + b)
$$
$$
= -fJG^{ij}\varepsilon_{jk}hv^k - \frac{1}{2}\left(\Gamma^i_{jk}hv^jv^k - G^{ik}\Gamma^l_{jk}hv^jv_l\right), \tag{24}
$$

which constitutes the *skew-symmetric form* of the momentum balance. Together with the continuity equation (8a), this form of the equations will be used in the construction of the entropy-stable discretizations proposed in this work.

Remark 2. To the authors' knowledge, skew-symmetric formulations for nonlinear balance laws in covariant form have not been proposed prior to this work.

3. Discontinuous spectral-element methods

In this section, we will introduce the main components of the spatial discretization framework adopted in this paper, which involves the application of flux-differencing discontinuous spectral-element methods using tensor-product SBP operators to the spatial discretization of balance laws in covariant form.

3.1. Spherical quadrilateral mapping

Let us assume that S is a spherical shell of radius a, and consider a global Cartesian coordinate system with its origin at the centre of the sphere. The sphere is tessellated to construct a conforming mesh of curved quadrilateral elements, which can be arbitrary in topology. For each mesh element $E \subset S$, we then let $\vec{x}_1, \vec{x}_2, \vec{x}_3, \vec{x}_4$ denote position vectors from the origin to four distinct points on S defining the corners of a curved element bounded by segments of four great circles. If $\xi^1, \xi^2 \in [-1, 1]$ denote coordinates on the reference quadrilateral $D \subset \mathbb{R}^2$, the position vector from the origin to the corresponding point on E is then given by

$$\vec{X}(\xi^1, \xi^2) := a \frac{\vec{x}_e(\xi^1, \xi^2)}{\|\vec{x}_e(\xi^1, \xi^2)\|}, \tag{25}$$

where, following Guba *et al.* [37, Appendix A], we define the bilinear mapping

$$\begin{aligned}
\vec{x}_e(\xi^1, \xi^2) := \frac{1}{4} \Big(& (1 - \xi^1)(1 - \xi^2)\vec{x}_1 + (1 + \xi^1)(1 - \xi^2)\vec{x}_2 \\
& + (1 + \xi^1)(1 + \xi^2)\vec{x}_3 + (1 - \xi^1)(1 + \xi^2)\vec{x}_4 \Big)
\end{aligned} \tag{26}$$

and note that a similar approach is taken in the case of triangular elements in [30] and [23]. Differentiating (25) as in (1), the covariant basis vectors can then be computed analytically as

$$\vec{a}_i(\xi^1, \xi^2) = \frac{a}{\|\vec{x}_e(\xi^1, \xi^2)\|} \left(\partial_i \vec{x}_e(\xi^1, \xi^2) - \frac{\vec{x}_e(\xi^1, \xi^2) \cdot \partial_i \vec{x}_e(\xi^1, \xi^2)}{\|\vec{x}_e(\xi^1, \xi^2)\|^2} \vec{x}_e(\xi^1, \xi^2) \right) \tag{27}$$

and used to analytically compute the covariant metric components, contravariant metric components, contravariant basis vectors, and Christoffel symbols. We will commonly use x, y, and z to denote the Cartesian coordinates of position vectors $\vec{x} = x\vec{e}_x + y\vec{e}_y + z\vec{e}_z$ relative to a right-handed coordinate system with its origin at the centre of the sphere. If the positive x axis points towards a longitude of $\lambda = 0$ and the positive z axis points towards a latitude of $\theta = \pi/2$, the spherical coordinates corresponding to a position vector $\vec{x}$ can be computed in radians as $\lambda = \arctan2(y, x)$ and $\theta = \arcsin(z/a)$, where we note that $a^2 = \|\vec{x}\|^2 = x^2 + y^2 + z^2$ and that the two-argument arctangent function is used such that $-\pi \le \lambda \le \pi$. In contrast with the common practice of using a global spherical coordinate system to represent the covariant and contravariant basis vectors, we represent such bases using the aforementioned global Cartesian coordinate system. Tangent vectors such as the velocity or momentum can therefore be transformed between local contravariant and global Cartesian component representations as

$$\begin{pmatrix} v^1 \\ v^2 \end{pmatrix} = \begin{pmatrix} \vec{e}_x \cdot \vec{a}^1 & \vec{e}_y \cdot \vec{a}^1 & \vec{e}_z \cdot \vec{a}^1 \\ \vec{e}_x \cdot \vec{a}^2 & \vec{e}_y \cdot \vec{a}^2 & \vec{e}_z \cdot \vec{a}^2 \end{pmatrix} \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix}, \quad \begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} \vec{e}_x \cdot \vec{a}_1 & \vec{e}_x \cdot \vec{a}_2 \\ \vec{e}_y \cdot \vec{a}_1 & \vec{e}_y \cdot \vec{a}_2 \\ \vec{e}_z \cdot \vec{a}_1 & \vec{e}_z \cdot \vec{a}_2 \end{pmatrix} \begin{pmatrix} v^1 \\ v^2 \end{pmatrix}, \tag{28}$$

where the above transformation matrices are well defined at all points on S. This avoids the issue of representing vectors which are nonzero at the poles with respect to a spherical basis.

3.2. Spectral-element approximation

To construct a spectral-element spatial discretization on the reference quadrilateral D, we will approximate the state vector $\boldsymbol{u}$ by a tensor-product polynomial expansion of the form

$$\boldsymbol{u}(\vec{X}(\xi^1, \xi^2), t) \approx \sum_{i=0}^{N} \sum_{j=0}^{N} \boldsymbol{u}_{ij}(t) \ell_i(\xi^1) \ell_j(\xi^2), \tag{29}$$

where the *Lagrange polynomials* corresponding to the nodes $\{\xi_i\}_{i=0}^{N} \subset [-1, 1]$ are given by

$$\ell_i(\xi) := \prod_{m=0, m \neq i}^{N} \frac{\xi - \xi_m}{\xi_i - \xi_m}, \quad i \in \{0, \ldots, N\}, \tag{30}$$

and $\boldsymbol{u}_{ij}(t) \approx \boldsymbol{u}(\vec{X}(\xi_i, \xi_j), t)$ are the vector-valued nodal coefficients. In this work, we employ the Legendre–Gauss–Lobatto (LGL) quadrature points, which include both endpoints of the reference interval, and correspond to positive quadrature weights, which we denote by $\{\omega_i\}_{i=0}^{N}$. We can then use such a quadrature rule to approximate integrals under the change of variables in (7) as

$$\int_E h \, \mathrm{d}S \approx I_E[h] := \sum_{i,j=0}^{N} \omega_i \omega_j J_{ij} h_{ij} \tag{31}$$

where the double subscript $(\cdot)_{ij}$ denotes a quantity evaluated at reference coordinates (ξ_i, ξ_j). By summing (31) over all elements, we can similarly approximate integrals over the global domain S as

$$\int_S h \, \mathrm{d}S \approx I_S[h] := \sum_{E \subset S} I_E[h]. \tag{32}$$

Discrete integrals such as those in (31) and (32) will be used extensively in our theoretical analysis and numerical verification of the schemes developed in this paper.

3.3. Summation-by-parts operators

Adopting a zero-based indexing convention for discretization matrices, the entries of the collocation derivative matrix $\boldsymbol{D}$ and diagonal mass matrix $\boldsymbol{M}$ corresponding to the Lagrange polynomials $\{\ell_i\}_{i=0}^{N}$ and quadrature weights $\{\omega_i\}_{i=0}^{N}$ are given by

$$D_{ij} := \left. \frac{\mathrm{d}\ell_j}{\mathrm{d}\xi} \right|_{\xi = \xi_i}, \quad M_{ij} := \omega_i \delta_{ij}, \quad i, j \in \{0, \ldots, N\}. \tag{33}$$

Defining the matrices $\boldsymbol{Q} := \boldsymbol{M}\boldsymbol{D}$ and $\boldsymbol{B} := \mathrm{diag}(-1, 0, \ldots, 0, 1)$, it was shown by Gassner [18] that the SBP property

$$\boldsymbol{Q} + \boldsymbol{Q}^{\mathrm{T}} = \boldsymbol{B} \tag{34}$$

is satisfied for collocated spectral-element operators based on LGL quadrature, mimicking the continuous integration-by-parts formula in one spatial dimension, which is given by

$$\int_{-1}^{1} u \frac{\mathrm{d}v}{\mathrm{d}\xi} \, \mathrm{d}\xi + \int_{-1}^{1} \frac{\mathrm{d}u}{\mathrm{d}\xi} v \, \mathrm{d}\xi = uv \Big|_{-1}^{1}. \tag{35}$$

We also note that the mass matrix $\boldsymbol{M}$ is guaranteed to be SPD due to the positivity of the LGL quadrature weights and that the following identity holds due to the exact differentiation of constant functions:

$$\sum_{m=0}^{N} D_{im} = 0, \quad i \in \{0, \ldots, N\}. \tag{36}$$

While we use collocated spectral-element operators in this work, the theory applies to any operators satisfying (34) and (36), including the finite-difference SBP schemes described, for example, in the review papers by Del Rey Fernández *et al.* [38] and Svärd and Nordström [39].

3.4. Flux-differencing discretization

The discretizations developed in this work will be formulated within the general framework of a nonconservative, variable-coefficient flux-differencing formulation, building on prior work by Renac [40], Waruszewski *et al.* [21], and Rueda-Ramírez *et al.* [41]. Defining the skew-symmetric matrix $\boldsymbol{S} := 2\boldsymbol{Q} - \boldsymbol{B}$, a flux-differencing tensor-product spatial discretization is given by

$$
\begin{aligned}
\omega_i \omega_j J_{ij} \frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{u}_{ij} = {} & \omega_j \left(-\sum_{m=0}^{N} S_{im} (J\boldsymbol{f}^1)^{\#}_{ij,mj} + \delta_{i0} (J\boldsymbol{f}^1)^*_{0j,0j+} - \delta_{iN} (J\boldsymbol{f}^1)^*_{Nj,Nj+} \right) \\
& + \omega_i \left(-\sum_{m=0}^{N} S_{jm} (J\boldsymbol{f}^2)^{\#}_{ij,im} + \delta_{j0} (J\boldsymbol{f}^2)^*_{i0,i0+} - \delta_{jN} (J\boldsymbol{f}^2)^*_{iN,iN+} \right) + \omega_i \omega_j J_{ij} \boldsymbol{s}_{ij},
\end{aligned} \tag{37}
$$

where we have suppressed the explicit dependence on time. Here, *flux differencing* refers to the use of two-point volume flux functions of the form

$$(J\boldsymbol{f}^1)^{\#}_{L,R} := (J\boldsymbol{f}^1)^{\#}\big((\boldsymbol{u}, \boldsymbol{G}, b)_L, (\boldsymbol{u}, \boldsymbol{G}, b)_R\big), \quad (J\boldsymbol{f}^2)^{\#}_{L,R} := (J\boldsymbol{f}^2)^{\#}\big((\boldsymbol{u}, \boldsymbol{G}, b)_L, (\boldsymbol{u}, \boldsymbol{G}, b)_R\big), \tag{38}$$

as well as interface flux functions of the form

$$(J\boldsymbol{f}^1)^*_{L,R} := (J\boldsymbol{f}^1)^*\big((\boldsymbol{u}, \boldsymbol{G}, b)_L, (\boldsymbol{u}, \boldsymbol{G}, b)_R\big), \quad (J\boldsymbol{f}^2)^*_{L,R} := (J\boldsymbol{f}^2)^*\big((\boldsymbol{u}, \boldsymbol{G}, b)_L, (\boldsymbol{u}, \boldsymbol{G}, b)_R\big), \tag{39}$$

which depend on the solution $\boldsymbol{u}$ as well as the metric $\boldsymbol{G}$ and bottom topography b at arbitrary "left" and "right" states $(\cdot)_L$ and $(\cdot)_R$, and are not necessarily symmetric in their two arguments. For example, $(J\boldsymbol{f}^1)^{\#}_{ij,mj} = (J\boldsymbol{f}^1)^{\#}((\boldsymbol{u}, \boldsymbol{G}, b)_{ij}, (\boldsymbol{u}, \boldsymbol{G}, b)_{mj})$ denotes a contravariant volume flux in the ξ^1 direction evaluated between the nodes (ξ_i, ξ_j) and (ξ_m, ξ_j), and is not necessarily equal to $(J\boldsymbol{f}^1)^{\#}_{mj,ij}$. Any index with a "+" superscript denotes an external state transformed from an adjacent element, where the local-to-global and global-to-local transformations in (28) are combined to form the matrix

$$\boldsymbol{A}_{R\to L} := \begin{pmatrix} \vec{e}_x \cdot \vec{a}^{\,1} & \vec{e}_y \cdot \vec{a}^{\,1} & \vec{e}_z \cdot \vec{a}^{\,1} \\ \vec{e}_x \cdot \vec{a}^{\,2} & \vec{e}_y \cdot \vec{a}^{\,2} & \vec{e}_z \cdot \vec{a}^{\,2} \end{pmatrix}_L \begin{pmatrix} \vec{e}_x \cdot \vec{a}_1 & \vec{e}_x \cdot \vec{a}_2 \\ \vec{e}_y \cdot \vec{a}_1 & \vec{e}_y \cdot \vec{a}_2 \\ \vec{e}_z \cdot \vec{a}_1 & \vec{e}_z \cdot \vec{a}_2 \end{pmatrix}_R, \tag{40}$$

which transforms the contravariant representation of the momentum vector with respect to the local coordinate system at $(\cdot)_R$ into the corresponding local coordinate system at $(\cdot)_L$.

Remark 3. To simplify our notation and analysis, we will assume that for all interface fluxes, the state vectors on both sides have been transformed so as to include momentum components with respect to the coordinate system associated with the first subscript, corresponding to the interior state, and all geometric quantities used in the flux are evaluated on that side of the interface.

Using the SBP property in (34) as well as the row-sum identity in (36), the weak-form flux-differencing discretization in (37) can be rewritten in strong form as

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u}_{ij} = &-\frac{1}{J_{ij}}\sum_{m=0}^{N}2D_{im}(J\boldsymbol{f}^1)^{\#}_{ij,mj} - \frac{1}{J_{ij}}\sum_{m=0}^{N}2D_{jm}(J\boldsymbol{f}^2)^{\#}_{ij,im} + \boldsymbol{s}_{ij} \\
&+ \frac{\delta_{i0}}{J_{ij}\omega_i\omega_j}\left[(J\boldsymbol{f}^1)^*_{0j,0j+} - (J\boldsymbol{f}^1)_{0j}\right] - \frac{\delta_{iN}}{J_{ij}\omega_i\omega_j}\left[(J\boldsymbol{f}^1)^*_{Nj,Nj+} - (J\boldsymbol{f}^1)_{Nj}\right] \\
&+ \frac{\delta_{j0}}{J_{ij}\omega_i\omega_j}\left[(J\boldsymbol{f}^2)^*_{i0,i0+} - (J\boldsymbol{f}^2)_{i0}\right] - \frac{\delta_{jN}}{J_{ij}\omega_i\omega_j}\left[(J\boldsymbol{f}^2)^*_{iN,iN+} - (J\boldsymbol{f}^2)_{iN}\right],
\end{aligned}
\tag{41}
$$

where we have assumed that two-point flux functions are consistent with (12) such that $(J\boldsymbol{f})^{\#}_{ij,ij} = (J\boldsymbol{f})^*_{ij,ij} = (J\boldsymbol{f})_{ij}$. The first line of (41) can then be interpreted as a discretization of the governing equations on a given element, while the second and third lines consist of penalty terms (often referred to as *simultaneous approximation terms* in the SBP literature, following Carpenter *et al.* [42]) which weakly impose the coupling between adjacent elements. Although the weak form in (37) yields a simpler and more efficient implementation and is more easily amenable to entropy analysis, the strong form will be useful in demonstrating consistency and well balancing of the proposed discretizations. Finally, we note that in the specific case that the two-point fluxes are taken to be

$$
(J\boldsymbol{f}^1)^{\#}_{L,R} := \frac{1}{2}\left[(J\boldsymbol{f}^1)_L + (J\boldsymbol{f}^1)_R\right], \quad (J\boldsymbol{f}^2)^{\#}_{L,R} := \frac{1}{2}\left[(J\boldsymbol{f}^2)_L + (J\boldsymbol{f}^2)_R\right],
\tag{42}
$$

and the source term is taken as $\boldsymbol{s} := \boldsymbol{s}_{\mathrm{cor}} + \boldsymbol{s}_{\mathrm{bot}} + \boldsymbol{s}_{\mathrm{geo}}$, the resulting scheme is equivalent to a standard tensor-product DG method based on collocated LGL quadrature, which is given in weak form by

$$
\begin{aligned}
\omega_i\omega_j J_{ij}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u}_{ij} = &\,\omega_j\left(\sum_{m=0}^{N}Q_{mi}(J\boldsymbol{f}^1)_{mj} + \delta_{i0}(J\boldsymbol{f}^1)^*_{0j,0j+} - \delta_{iN}(J\boldsymbol{f}^1)^*_{Nj,Nj+}\right) \\
&+ \omega_i\left(\sum_{m=0}^{N}Q_{mj}(J\boldsymbol{f}^2)_{im} + \delta_{j0}(J\boldsymbol{f}^2)^*_{i0,i0+} - \delta_{jN}(J\boldsymbol{f}^2)^*_{iN,iN+}\right) + \omega_i\omega_j J_{ij}\boldsymbol{s}_{ij}.
\end{aligned}
\tag{43}
$$

In the following section, we will construct novel formulations of the volume flux, interface flux, and source term so as to ensure consistency, mass conservation, entropy stability, and well balancing of the resulting spatial discretization.

4. Entropy-stable covariant formulation

Considering the semi-discrete entropy balance on a single element, we can use (31) and apply the chain rule in time to obtain

$$
\frac{\mathrm{d}}{\mathrm{d}t}\int_E \eta\,\mathrm{d}S \approx \frac{\mathrm{d}}{\mathrm{d}t}I_E[\eta] = \sum_{i,j=0}^{N}\omega_i\omega_j J_{ij}\boldsymbol{w}_{ij}^{\mathrm{T}}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{u}_{ij},
\tag{44}
$$

where $\boldsymbol{w}_{ij}$ denotes the vector of entropy variables in (16) evaluated in terms of the discrete nodal coefficients $\boldsymbol{u}_{ij}$. In this section, we will develop an approach for constructing flux-differencing

12

discretizations which satisfy a semi-discrete form of (18) by ensuring that the time derivative of $I_E[\eta]$ is bounded from above by telescoping interface terms, leading to a semi-discrete bound on the entropy integrated numerically over the global domain S. We will also show that the resulting discretizations are conservative of mass and well balanced for arbitrary continuous bottom topographies, preserving a constant surface height and zero velocity.

4.1. Entropy-conservative and entropy-stable two-point flux functions

Our next step is to discretize the skew-symmetric formulation proposed in Section 2.4 within the flux-differencing framework introduced in Section 3.4. The following lemma presents a two-point flux, which, when used with an appropriate formulation of the source term, accomplishes this task.

Lemma 3. *Taking the volume flux and interface flux terms in* (37) *to be*

$$
(J\boldsymbol{f}^j)_{L,R}^{\#} := \begin{pmatrix} \frac{1}{2}\big[(Jhv^j)_L + (Jhv^j)_R\big] \\ \frac{1}{4}\big[(Jhv^jv^1)_L + (Jhv^jv^1)_R + (Jhv^j)_R(v^1)_L + (G^{1k}Jhv^j)_L(v_k)_R\big] \\ \frac{1}{4}\big[(Jhv^jv^2)_L + (Jhv^jv^2)_R + (Jhv^j)_R(v^2)_L + (G^{2k}Jhv^j)_L(v_k)_R\big] \end{pmatrix}
$$
$$
+ \begin{pmatrix} 0 \\ \frac{g}{2}(G^{1j}Jh)_L h_R \\ \frac{g}{2}(G^{2j}Jh)_L h_R \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{g}{2}(G^{1j}Jh)_L(b_R - b_L) \\ \frac{g}{2}(G^{2j}Jh)_L(b_R - b_L) \end{pmatrix}
\tag{45}
$$

results in a consistent discretization of (11) *when the source term is given by*

$$
\boldsymbol{s} := \begin{pmatrix} 0 \\ -\frac{1}{2}\big(\Gamma^1_{jk}hv^jv^k - G^{1k}\Gamma^l_{jk}hv^jv_l\big) - fJG^{1j}\varepsilon_{jk}hv^k \\ -\frac{1}{2}\big(\Gamma^2_{jk}hv^jv^k - G^{2k}\Gamma^l_{jk}hv^jv_l\big) - fJG^{2j}\varepsilon_{jk}hv^k \end{pmatrix}.
\tag{46}
$$

Proof. Considering the equivalent strong form in (41), we first note that when all interface states are expressed consistently within a given local coordinate system, the flux differences which constitute the surface terms vanish for globally continuous solutions. Similarly to [12, Lemma 1], the flux-differencing volume terms can be rewritten as

$$
\sum_{m=0}^{N} 2D_{im}(J\boldsymbol{f}^1)_{ij,mj}^{\#} = \sum_{m=0}^{N} \begin{pmatrix} D_{im}(Jhv^1)_{mj} \\ \frac{1}{2}\big[D_{im}(Jhv^1v^1)_{mj} + (v^1)_{ij}D_{im}(Jhv^1)_{mj} + (G^{1k}Jhv^1)_{ij}D_{im}(v_k)_{mj}\big] \\ \frac{1}{2}\big[D_{im}(Jhv^1v^2)_{mj} + (v^2)_{ij}D_{im}(Jhv^1)_{mj} + (G^{2k}Jhv^1)_{ij}D_{im}(v_k)_{mj}\big] \end{pmatrix}
$$
$$
+ \sum_{m=0}^{N} \begin{pmatrix} 0 \\ g(JG^{11}h)_{ij}D_{im}(h+b)_{mj} \\ g(JG^{21}h)_{ij}D_{im}(h+b)_{mj} \end{pmatrix}
\tag{47}
$$

and

$$
\sum_{m=0}^{N} 2D_{jm}(J\boldsymbol{f}^2)_{ij,im}^{\#} = \sum_{m=0}^{N} \begin{pmatrix} D_{jm}(Jhv^2)_{im} \\ \frac{1}{2}\big[D_{jm}(Jhv^2v^1)_{im} + (v^1)_{ij}D_{jm}(Jhv^2)_{im} + (G^{1k}Jhv^2)_{ij}D_{jm}(v_k)_{im}\big] \\ \frac{1}{2}\big[D_{jm}(Jhv^2v^2)_{im} + (v^2)_{ij}D_{jm}(Jhv^2)_{im} + (G^{2k}Jhv^2)_{ij}D_{jm}(v_k)_{im}\big] \end{pmatrix}
$$
$$
+ \sum_{m=0}^{N} \begin{pmatrix} 0 \\ g(JG^{12}h)_{ij}D_{jm}(h+b)_{im} \\ g(JG^{22}h)_{ij}D_{jm}(h+b)_{im} \end{pmatrix},
\tag{48}
$$

where we have used the fact that terms of the form $D_{im}(\cdot)_{ij}$ and $D_{jm}(\cdot)_{ij}$ do not contribute to sums over m as a consequence of (36). After dividing by J_{ij}, we then see that the right-hand sides of (47) and (48) correspond directly to the differential operators appearing in the skew-symmetric formulation, with the partial derivatives in (11a) and (24) replaced by nodal approximations using the SBP derivative matrix in (33). Noting that the momentum source terms in (46) are identical to those appearing on the right-hand side of (24), the result therefore follows from Lemma 2 and the consistency of the spectral-element derivative approximation. $\qquad\square$

Remark 4. In the special case of the Euclidean metric $G_{ij} = \delta_{ij}$, the two-point flux in (45) is the same as that proposed in [27, Eq. (4.3)] for the shallow water equations in planar geometry. We also note that in addition to making use of split-form approximation of the covariant derivative, the bottom topography term $gG^{ij}\partial_j b/2$ is treated as a nonconservative contribution to the two-point flux in (45) rather than as a pointwise source term.

Next, we present the following lemma establishing that the proposed flux is *entropy conservative* in the sense that it satisfies a particular *shuffle condition*, also known as a *non-symmetric Tadmor condition* (borrowing terminology from Chan *et al.* [43] with reference to [15]), which we will employ in the following subsection to establish entropy stability of the resulting scheme.

Lemma 4. *The two-point flux in* (45) *is entropy conservative, satisfying the shuffle condition*

$$\boldsymbol{w}_R^{\mathrm{T}}(J\boldsymbol{f}^j)_{R,L}^{\#} - \boldsymbol{w}_L^{\mathrm{T}}(J\boldsymbol{f}^j)_{L,R}^{\#} = (J\Psi^j)_R - (J\Psi^j)_L, \tag{49}$$

where the contravariant flux potential components are given by $\Psi^j := \boldsymbol{w}^{\mathrm{T}}\boldsymbol{f}^j - F^j = gh^2 v^j/2$.

Proof. If we first consider the case of $b = 0$ and recall the definition of the entropy variables in (16), the left-hand side of (49) can be expanded term-by-term as

$$
\begin{aligned}
\boldsymbol{w}_R^{\mathrm{T}}(J\boldsymbol{f}^j)_{R,L}^{\#} - \boldsymbol{w}_L^{\mathrm{T}}(J\boldsymbol{f}^j)_{L,R}^{\#} = {} & \underbrace{\frac{g}{2}h_R(Jhv^j)_L}_{(A)} - \underbrace{\frac{g}{2}h_L(Jhv^j)_L + \frac{g}{2}h_R(Jhv^j)_R - \frac{g}{2}h_L(Jhv^j)_R}_{(B)} \\
& - \underbrace{\frac{1}{4}(v_i v^i)_R(Jhv^j)_L}_{(C)} + \underbrace{\frac{1}{4}(v_i v^i)_L(Jhv^j)_L}_{(D)} - \underbrace{\frac{1}{4}(v_i v^i)_R(Jhv^j)_R}_{(E)} + \underbrace{\frac{1}{4}(v_i v^i)_L(Jhv^j)_R}_{(F)} \\
& + \underbrace{\frac{1}{4}(v_i)_R(Jhv^i v^j)_R}_{(E)} - \underbrace{\frac{1}{4}(v_i)_L(Jhv^i v^j)_L}_{(D)} + \underbrace{\frac{1}{4}(v_i)_R(Jhv^i v^j)_L}_{(G)} - \underbrace{\frac{1}{4}(v_i)_L(Jhv^i v^j)_R}_{(H)} \\
& + \underbrace{\frac{1}{4}(v_i)_R(Jhv^j)_L(v^i)_R}_{(C)} - \underbrace{\frac{1}{4}(v_i)_L(Jhv^j)_R(v^i)_L}_{(F)} + \underbrace{\frac{1}{4}(v_i)_R(G^{ik}Jhv^j)_R(v_k)_L}_{(H)} \\
& - \underbrace{\frac{1}{4}(v_i)_L(G^{ik}Jhv^j)_L(v_k)_R}_{(G)} + \underbrace{\frac{g}{2}(v_i)_R(G^{ij}Jh)_R h_L}_{(B)} - \underbrace{\frac{g}{2}(v_i)_L(G^{ij}Jh)_L h_R}_{(A)}.
\end{aligned}
\tag{50}
$$

Noting the cancellation of pairs labelled (A) through (H), we see that the remaining terms constitute the right-hand side of (49). Next, incorporating the influence of the bottom topography b within the

entropy variables as well as the two-point fluxes, we obtain

$$
\boldsymbol{w}_R^{\mathrm{T}}(J\boldsymbol{f}^j)_{R,L}^{\#} - \boldsymbol{w}_L^{\mathrm{T}}(J\boldsymbol{f}^j)_{L,R}^{\#} = (J\Psi^j)_R - (J\Psi^j)_L + \underbrace{\frac{g}{2}b_R(Jhv^j)_L}_{(A)} - \underbrace{\frac{g}{2}b_L(Jhv^j)_L}_{(B)} + \underbrace{\frac{g}{2}b_R(Jhv^j)_R}_{(C)}
$$

$$
- \underbrace{\frac{g}{2}b_L(Jhv^j)_R}_{(D)} + \underbrace{\frac{g}{2}(v_i)_R(G^{ij}Jh)_R b_L}_{(D)} - \underbrace{\frac{g}{2}(v_i)_L(G^{ij}Jh)_L b_R}_{(A)} \tag{51}
$$

$$
- \underbrace{\frac{g}{2}(v_i)_R(G^{ij}Jh)_R b_R}_{(C)} + \underbrace{\frac{g}{2}(v_i)_L(G^{ij}Jh)_L b_L}_{(B)},
$$

where, similarly to (50), the pairs labelled (A) through (D) cancel, again resulting in (49) when the effect of variable bottom topography is included. $\qquad\square$

At element interfaces, we augment the entropy-conservative flux in (45) with local Lax–Friedrichs dissipation applied to the jump in the conservative variables in order to obtain

$$
(J\boldsymbol{f}^j)_{L,R}^* := (J\boldsymbol{f}^j)_{L,R}^{\#} - \frac{J_L}{2}\max(\lambda_L^j, \lambda_R^j)(\boldsymbol{u}_R - \boldsymbol{u}_L), \tag{52}
$$

where the contravariant wave speed estimate is computed by evaluating the spectral radius of the flux Jacobian $\partial\boldsymbol{f}^j/\partial\boldsymbol{u}$ as

$$
\lambda^j := \left|v^j\right| + \sqrt{ghG^{jj}} \tag{53}
$$

for which derivations can be found, for example, in [23, Appendix D] and [24, Appendix E]. Following Ranocha's analysis in [44, Section 6.1], one can show that the resulting numerical interface flux in (52) is *entropy stable*, satisfying

$$
\boldsymbol{w}_R^{\mathrm{T}}(J\boldsymbol{f}^j)_{R,L}^* - \boldsymbol{w}_L^{\mathrm{T}}(J\boldsymbol{f}^j)_{L,R}^* \leq (J\Psi^j)_R - (J\Psi^j)_L \tag{54}
$$

when $h > 0$ and the bottom topography b is continuous across element interfaces.

Remark 5. To show that an entropy-conservative two-point flux augmented with a standard local Lax–Friedrichs dissipation term is entropy stable, Ranocha relies on the Fundamental Theorem of Calculus in [44, Eq. (68)] to express the jump in the conservative variables across an element interface as

$$
\boldsymbol{u}_R - \boldsymbol{u}_L = \left(\int_0^1 \frac{\partial\boldsymbol{u}}{\partial\boldsymbol{w}}\big((1-\sigma)\boldsymbol{w}_L + \sigma\boldsymbol{w}_R\big)\,\mathrm{d}\sigma\right)(\boldsymbol{w}_R - \boldsymbol{w}_L), \tag{55}
$$

so as to ensure that the dissipation term does not contribute positively to the entropy balance under the assumption that $\partial\boldsymbol{u}/\partial\boldsymbol{w}$ is SPD along the entire path from $\boldsymbol{w}_L$ to $\boldsymbol{w}_R$. When the entropy variables $\boldsymbol{w}$ depend on the bottom topography as in (16), however, one cannot invoke (55) across an interface with $b_L \neq b_R$ and thus the arguments in [44] do not imply entropy stability for discontinuous bottom topographies. While we do not consider discontinuous bottom topographies in this work, one entropy-stable option in that setting is to instead penalize the jump in the entropy variables as

$$
(J\boldsymbol{f}^j)_{L,R}^* := (J\boldsymbol{f}^j)_{L,R}^{\#} - \frac{J_L}{2}\max(\lambda_L^j, \lambda_R^j)\boldsymbol{H}_{L,R}(\boldsymbol{w}_R - \boldsymbol{w}_L), \tag{56}
$$

where $\boldsymbol{H}_{L,R}$ is SPD (or, at least, symmetric positive semi-definite) and typically corresponds to the evaluation of $\partial\boldsymbol{u}/\partial\boldsymbol{w}$ at an average between the two states. The construction of such a dissipation term was explored in detail by Ranocha [45] in the context of the planar shallow water equations.

4.2. Mass conservation

Noting that the two-point mass flux $(Jhv^j)^{\#}_{L,R}$ corresponding to the first entry of (45) is simply an arithmetic mean, it follows that the resulting flux-differencing discretization of the continuity equation is identical to that used by the standard DG method recovered in (43). While we therefore expect Baldauf's proof of conservation in [23, Appendix E] to extend to the present formulation, we will proceed more generally with the following theorem, which applies to any symmetric two-point flux for a scalar variable discretized in flux form.

Theorem 1. *If mass flux components $(Jhv^1)^{\#}_{L,R}$ and $(Jhv^2)^{\#}_{L,R}$ are symmetric with respect to $(\cdot)_L$ and $(\cdot)_R$, and the corresponding entry in the source term vector $\boldsymbol{s}$ is zero, the semi-discrete mass balance on a single element is then given for the scheme in (37) by*

$$\frac{\mathrm{d}}{\mathrm{d}t} I_E[h] = \sum_{j=0}^{N} \omega_j \left[(Jhv^1)^*_{0j,0j^+} - (Jhv^1)^*_{Nj,Nj^+} \right] + \sum_{i=0}^{N} \omega_i \left[(Jhv^2)^*_{i0,i0^+} - (Jhv^2)^*_{iN,iN^+} \right], \quad (57)$$

where $(Jhv^1)^*_{L,R}$ and $(Jhv^2)^*_{L,R}$ are the corresponding interface flux components for the continuity equation.

Proof. Considering the component of (37) corresponding to the continuity equation and summing over all quadrature nodes, we obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} I_E[h] = & \sum_{j=0}^{N} \omega_j \left(- \sum_{i,m=0}^{N} S_{im}(Jhv^1)^{\#}_{ij,mj} + \sum_{i=0}^{N} \left[\delta_{i0}(Jhv^1)^*_{0j,0j^+} - \delta_{iN}(Jhv^1)^*_{Nj,Nj^+} \right] \right) \\
& + \sum_{i=0}^{N} \omega_i \left(- \sum_{j,m=0}^{N} S_{jm}(Jhv^2)^{\#}_{ij,im} + \sum_{j=0}^{N} \left[\delta_{j0}(Jhv^2)^*_{i0,i0^+} - \delta_{jN}(Jhv^2)^*_{iN,iN^+} \right] \right),
\end{aligned} \quad (58)$$

due to the absence of source terms in the mass balance. Noting that the right-hand side of (57) results from summing over the Kronecker delta, we must show that the contributions from the flux-differencing volume terms vanish. Using the SBP property in (34) to obtain

$$\boldsymbol{S} = \boldsymbol{Q} - \boldsymbol{Q}^{\mathrm{T}}, \quad (59)$$

we can rewrite the contribution on the first line of (58) as

$$\sum_{i,m=0}^{N} S_{im}(Jhv^1)^{\#}_{ij,mj} = \sum_{i,m=0}^{N} Q_{im}(Jhv^1)^{\#}_{ij,mj} - \sum_{i,m=0}^{N} Q_{mi}(Jhv^1)^{\#}_{ij,mj}, \quad (60)$$

where the first and second terms on the right-hand side cancel when the two-point mass flux is symmetric with respect to its two subscripts. Applying an analogous procedure to the volume term on the second line of (58), we therefore obtain (57). $\qquad\square$

The right-hand side of (57) is a discretization of the line integral of the inward normal component of the numerical mass flux over the boundary of the quadrilateral element. To show that the scheme is globally conservative of mass, we proceed similarly to [23, Appendix E] under the assumption that every node on the boundary of the element E, associated with the state $(\cdot)_L$, coincides in physical

16

space with a node on the boundary of an adjacent element, associated with the state $(\cdot)_R$, for which the corresponding quadrature weight is equal to that associated with $(\cdot)_L$. Summing (57) over all elements to obtain the global integral in (32) and considering the contribution to the right-hand side from each pair of coincident interface nodes, we see that the discretization is globally conservative when the interface mass fluxes cancel as

$$(Jhv^j)^*_{L,R}(n_j)_L = -(Jhv^j)^*_{R,L}(n_j)_R \tag{61}$$

for all coincident pairs of interface nodes, where n_j denotes the jth component of the outward unit normal vector in reference coordinates. Recalling from Remark 3 that evaluating the left-hand side of (61) requires transforming the momentum vector at $(\cdot)_R$ into the $(\cdot)_L$ coordinate system, while the right-hand side requires transforming the momentum vector at $(\cdot)_L$ into the $(\cdot)_R$ coordinate system, it follows that the resulting discretization is conservative across element interfaces if the transformation matrices in (40) satisfy

$$\boldsymbol{A}_{L \to R} = \boldsymbol{A}_{R \to L}^{-1}, \tag{62}$$

which is the case when the manifold is smooth and the covariant and contravariant basis vectors are computed analytically, as described in Section 3.1.

Remark 6. In addition to the element-wise conservation property established above, the flux-differencing discretization in (37) satisfies conservation of mass at the *nodal* level for any symmetric two-point mass flux. This property follows from the results of Fisher *et al.* [46, 47] and Carpenter *et al.* [48], who showed that flux-differencing formulations using diagonal-norm SBP operators including boundary nodes, such as those used in our LGL-based discontinuous spectral-element method, can be expressed in an equivalent telescoping flux form compatible with the Lax–Wendroff theorem when combined with symmetric two-point fluxes within a one-dimensional or tensor-product discretization. The existence of such an equivalence is important for the use of subcell limiting techniques (see, for example, Rueda-Ramírez *et al.* [49]).

4.3. Entropy conservation and entropy stability

We now introduce the following definitions for the numerical entropy flux and entropy production terms, which will be used to facilitate the semi-discrete entropy analysis of the covariant-form discretization through a similar approach to that employed in [41] for discretizations of standard Cartesian formulations of nonconservative systems.

Definition 1. The *numerical entropy flux* between two states $(\cdot)_L$ and $(\cdot)_R$ on each side of an element interface is given by

$$(JF^j)^*_{L,R} := \frac{1}{2}\Big[\boldsymbol{w}_L^{\text{T}}(J\boldsymbol{f}^j)^*_{L,R} + \boldsymbol{w}_R^{\text{T}}(J\boldsymbol{f}^j)^*_{R,L}\Big] - \frac{1}{2}\Big[(J\Psi^j)_L + (J\Psi^j)_R\Big]. \tag{63}$$

The *numerical entropy production* terms associated with pairs of volume or interface nodes are given, respectively, by

$$r^{j\#}_{L,R} := \Big[\boldsymbol{w}_R^{\text{T}}(J\boldsymbol{f}^j)^\#_{R,L} - \boldsymbol{w}_L^{\text{T}}(J\boldsymbol{f}^j)^\#_{L,R}\Big] - \Big[(J\Psi^j)_R - (J\Psi^j)_L\Big], \tag{64a}$$

$$r^{j*}_{L,R} := \Big[\boldsymbol{w}_R^{\text{T}}(J\boldsymbol{f}^j)^*_{R,L} - \boldsymbol{w}_L^{\text{T}}(J\boldsymbol{f}^j)^*_{L,R}\Big] - \Big[(J\Psi^j)_R - (J\Psi^j)_L\Big]. \tag{64b}$$

17

Remark 7. The entropy conservation and entropy stability conditions in (49) and (54) correspond to $r_{L,R}^{j\#} = 0$ and $r_{L,R}^{j*} \leq 0$, respectively.

Besides the entropy conservation and entropy stability properties of the two-point flux functions, our proofs of entropy conservation and entropy stability will make use of the fact that the source term yields zero net contribution to the entropy balance, which is established as follows.

Lemma 5. *The source term in* (46) *satisfies* $\boldsymbol{w}^{\mathrm{T}}\boldsymbol{s} = 0$.

Proof. Contracting the source term in (46) with the vector of entropy variables results in

$$\boldsymbol{w}^{\mathrm{T}}\boldsymbol{s} = -v_i \left(\frac{1}{2} \Gamma_{jk}^i h v^j v^k - \frac{1}{2} G^{ik} \Gamma_{jk}^l h v^j v_l + f J G^{ij} \varepsilon_{jk} h v^k \right). \tag{65}$$

Noting that the contravariant metric tensor components in the last two terms in parentheses raise the covariant index in v_i to obtain the contravariant components v^k and v^j, respectively, the contribution from the geometric source term then vanishes as in (22), while the contribution from the Coriolis term vanishes due to the skew-symmetry of the Levi-Civita symbol. $\qquad\square$

We now present the main result of this section, establishing the semi-discrete entropy balance.

Theorem 2. *Using the numerical entropy flux and production terms introduced in Definition 1, the time rate of change in the mathematical entropy integrated over a given element is given for the spatial discretization in* (37) *by*

$$\frac{\mathrm{d}}{\mathrm{d}t} I_E[\eta] = \sum_{j=0}^{N} \omega_j \left[(JF^1)_{0j,0j+}^* - (JF^1)_{Nj,Nj+}^* \right] + \sum_{i=0}^{N} \omega_i \left[(JF^2)_{i0,i0+}^* - (JF^2)_{iN,iN+}^* \right]$$

$$+ \frac{1}{2} \sum_{j=0}^{N} \omega_j \left(r_{0j+,0j}^{1*} + r_{Nj,Nj+}^{1*} \right) + \frac{1}{2} \sum_{i=0}^{N} \omega_i \left(r_{i0+,i0}^{2*} + r_{iN,iN+}^{2*} \right) \tag{66}$$

$$+ \sum_{i,j,m=0}^{N} \omega_i \omega_j \left(D_{im} r_{ij,mj}^{1\#} + D_{jm} r_{ij,im}^{2\#} \right).$$

Proof. Contracting both sides of (37) with the vector of entropy variables $\boldsymbol{w}_{ij}$ and summing over all quadrature nodes, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} I_E[\eta] = \sum_{j=0}^{N} \omega_j \left(-\sum_{i,m=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} S_{im} (J\boldsymbol{f}^1)_{ij,mj}^{\#} + \sum_{i=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} \left[\delta_{i0} (J\boldsymbol{f}^1)_{0j,0j+}^* - \delta_{iN} (J\boldsymbol{f}^1)_{Nj,Nj+}^* \right] \right)$$

$$+ \sum_{i=0}^{N} \omega_i \left(-\sum_{j,m=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} S_{jm} (J\boldsymbol{f}^2)_{ij,im}^{\#} + \sum_{j=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} \left[\delta_{j0} (J\boldsymbol{f}^2)_{i0,i0+}^* - \delta_{jN} (J\boldsymbol{f}^2)_{iN,iN+}^* \right] \right) \tag{67}$$

$$+ \sum_{i,j=0}^{N} \omega_i \omega_j J_{ij} \boldsymbol{w}_{ij}^{\mathrm{T}} \boldsymbol{s}_{ij}.$$

Considering the contribution from the flux-differencing volume term on the first line of (67) and using (59), we can use the definition of the entropy production term in (64a) to obtain

$$
\begin{aligned}
\sum_{i,m=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} S_{im} (J\boldsymbol{f}^1)_{ij,mj}^{\#} &= \sum_{i,m=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} (Q_{im} - Q_{mi})(J\boldsymbol{f}^1)_{ij,mj}^{\#} \\
&= \sum_{i,m=0}^{N} Q_{im} \left[\boldsymbol{w}_{ij}^{\mathrm{T}} (J\boldsymbol{f}^1)_{ij,mj}^{\#} - \boldsymbol{w}_{mj}^{\mathrm{T}} (J\boldsymbol{f}^1)_{mj,ij}^{\#} \right] \\
&= \sum_{i,m=0}^{N} Q_{im} \left[(J\Psi^1)_{mj} - (J\Psi^1)_{ij} - r_{ij,mj}^{1\#} \right] \\
&= \sum_{i,m=0}^{N} Q_{im} (J\Psi^1)_{mj} - \sum_{i=0}^{N} (J\Psi^1)_{ij} \sum_{m=0}^{N} Q_{im} - \sum_{i,m=0}^{N} Q_{im} r_{ij,mj}^{1\#} \\
&= (J\Psi^1)_{Nj} - (J\Psi^1)_{0j} - \sum_{i,m=0}^{N} \omega_i D_{im} r_{ij,mj}^{1\#},
\end{aligned}
\tag{68}
$$

where the final equality results from the properties in (34) and (36) as well as the definition of $\boldsymbol{Q}$. Following the same procedure for the flux-differencing term on the second line of (67), we likewise obtain

$$
\sum_{j,m=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} S_{jm} (J\boldsymbol{f}^2)_{ij,im}^{\#} = (J\Psi^2)_{iN} - (J\Psi^2)_{i0} - \sum_{j,m=0}^{N} \omega_j D_{jm} r_{ij,im}^{2\#}.
\tag{69}
$$

We now turn our attention to the interface terms, which contract with the entropy variables as

$$
\begin{aligned}
\sum_{i=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} \left[\delta_{i0} (J\boldsymbol{f}^1)_{0j,0j+}^{*} - \delta_{iN} (J\boldsymbol{f}^1)_{Nj,Nj+}^{*} \right] &= (J\Psi^1)_{0j} - (J\Psi^1)_{Nj} \\
&+ (JF^1)_{0j,0j+}^{*} - (JF^1)_{Nj,Nj+}^{*} + \frac{1}{2} \left(r_{0j+,0j}^{1*} + r_{Nj,Nj+}^{1*} \right)
\end{aligned}
\tag{70}
$$

and

$$
\begin{aligned}
\sum_{j=0}^{N} \boldsymbol{w}_{ij}^{\mathrm{T}} \left[\delta_{j0} (J\boldsymbol{f}^2)_{i0,i0+}^{*} - \delta_{jN} (J\boldsymbol{f}^2)_{iN,iN+}^{*} \right] &= (J\Psi^2)_{i0} - (J\Psi^2)_{iN} \\
&+ (JF^2)_{i0,i0+}^{*} - (JF^2)_{iN,iN+}^{*} + \frac{1}{2} \left(r_{i0+,i0}^{2*} + r_{iN,iN+}^{2*} \right),
\end{aligned}
\tag{71}
$$

where we have used Lemma 4 as well as (63) and (64b). Substituting (68), (69), (70), and (71) into (67), noting the cancellation of the terms $(J\Psi^1)_{0j} - (J\Psi^1)_{Nj}$ and $(J\Psi^2)_{i0} - (J\Psi^2)_{iN}$, and invoking Lemma 5 to show that the source term in (46) does not contribute to the entropy balance, we therefore obtain (66). $\qquad\square$

Remark 8. Unlike existing proofs of entropy stability in curvilinear coordinates (see, for example, [17, Appendix B.1.1]), the proof of Theorem 2 does not rely on any discrete metric identities. This allows us to use an analytical representation of the geometry and tangent basis vectors, which ensures that (62) is satisfied.

The first line on the right-hand side of (66) represents a discrete integral of the numerical entropy flux directed into the element, the second line contains interface production terms which are zero for an entropy-conservative interface flux and non-positive for an entropy-stable interface flux, and the third line contains volume production terms which are zero for an entropy-conservative volume flux. Under such conditions, the resulting entropy balance becomes

$$\frac{\mathrm{d}}{\mathrm{d}t} I_E[\eta] \le \sum_{j=0}^{N} \omega_j \left[(JF^1)^*_{0j,0j+} - (JF^1)^*_{Nj,Nj+} \right] + \sum_{i=0}^{N} \omega_i \left[(JF^2)^*_{i0,i0+} - (JF^2)^*_{iN,iN+} \right], \quad (72)$$

which holds as an equality when the volume and interface fluxes are both entropy conservative. A semi-discrete bound on the integrated entropy analogous to (18) can therefore be derived under the same assumptions invoked for global mass conservation in Section 4.2 through the cancellation of numerical entropy flux contributions at element interfaces.

4.4. Well balancing

It is often desired for a numerical method to exactly preserve certain equilibrium states in order to minimize the error in discretizations of problems corresponding to small perturbations of such equilibria. For the shallow water equations, the "lake at rest" or "atmosphere at rest" state is an important example of such an equilibrium, corresponding to a steady state given everywhere on S by

$$H = \text{constant}, \quad \vec{v} = \vec{0}, \tag{73}$$

where $H := h + b$ denotes the total surface height, corresponding to a balance between the pressure gradient and the bottom topography term. The following theorem establishes that the proposed scheme preserves such an equilibrium.

Theorem 3. *Assuming that the bottom topography b is continuous, the discretization in (37) with the volume flux and source term introduced in Lemma 3 and the surface flux in (52) is well balanced in the sense that any solution satisfying (73) results in $\mathrm{d}\boldsymbol{u}_{ij}/\mathrm{d}t = 0$.*

Proof. Proceeding similarly to the proof of Lemma 3, we begin with the strong form in (41) and again note that the interface terms vanish for continuous bottom topographies. Rewriting the volume terms as in (47) and (48), we then set the contravariant velocity components v^1 and v^2 both to zero and combine the contributions from h and b to obtain

$$\frac{\mathrm{d}}{\mathrm{d}t} \boldsymbol{u}_{ij} = - \sum_{m=0}^{N} \begin{pmatrix} 0 \\ g(G^{11}h)_{ij} D_{im} H_{mj} \\ g(G^{21}h)_{ij} D_{im} H_{mj} \end{pmatrix} - \sum_{m=0}^{N} \begin{pmatrix} 0 \\ g(G^{12}h)_{ij} D_{jm} H_{im} \\ g(G^{22}h)_{ij} D_{jm} H_{im} \end{pmatrix}. \tag{74}$$

Finally, we can use (36) to show that both terms on the right-hand side of (74) vanish when the surface height $h + b$ is constant. □

Remark 9. The assumption of continuous bottom topography is used in the proof of Theorem 3 to show that the interface terms in (41) vanish for constant $h + b$. For discontinuous bottom topographies (i.e., $b_L \ne b_R$ across an interface), these cancellations generally do not hold for the present discretization. In such contexts, well-balanced schemes can be obtained, for example, using hydrostatic reconstruction procedures based on the work of Audusse *et al.* [50] and related well-balancing approaches for discontinuous topography (see, for example, Fjordholm *et al.* [51]).

5. Numerical experiments

The numerical experiments presented in this section are performed using `TrixiAtmo.jl`,[2] an open-source package written in Julia [52] extending the `Trixi.jl` simulation framework [53] to atmospheric flow problems. The scripts required to produce and post-process the results in this paper are provided within the article's reproducibility repository [54]. All initial conditions and problem data are prescribed in SI units (kg-m-s), and, following [1], we make use of the following values for the Earth's radius, rotation rate, and gravitational acceleration:

$$a = 6.37122 \times 10^6 \text{ m}, \quad \Omega = 7.292 \times 10^{-5} \text{ s}^{-1}, \quad g = 9.80616 \text{ m s}^{-2}. \tag{75}$$

Taking $\vec{e}_z$ to be the Earth's rotation axis, the Coriolis parameter is given by $f = 2\Omega y/a = 2\Omega \sin\theta$ for all cases considered here. The mesh is constructed as described in Section 3.1, with the corner vertices placed according to a uniform subdivision of the cube mapped onto the sphere using an equiangular gnomonic projection (see, for example, Ronchi *et al.* [55] as well as earlier work by Sadourny [56]), and all metric terms are computed exactly. Similarly to [31], we consider the nominal resolution for such cubed-sphere grids to be the average equatorial distance between quadrature points, which can be computed as

$$\text{Nominal resolution} := \frac{\pi}{2}\frac{a}{MN}, \tag{76}$$

where M denotes the number of elements per direction on each face of the cube and N denotes the polynomial degree of the tensor-product spectral-element approximation. Time integration is performed using the fourth-order, low-storage explicit Runge–Kutta method of Carpenter and Kennedy [57] implemented in `DifferentialEquations.jl` [58]. We make use of an adaptive time step based on the Courant–Friedrichs–Lewy (CFL) condition, as given by

$$\Delta t = C \min_{E \subset S} \min_{i,j=0}^{N} \left(\frac{2}{N+1}\frac{1}{|\lambda_{ij}^1| + |\lambda_{ij}^2|} \right), \tag{77}$$

where $C > 0$ is the Courant number, λ_{ij}^1 and λ_{ij}^2 correspond to the nodal values of the contravariant wave speeds in (53), and we note that the quantity in parentheses in (77) is an approximation of the ratio of the local mesh spacing to the maximum wave speed in physical space, which we have adapted from Ranocha *et al.* [59, 60]. In order to ensure that the temporal discretization error is negligible relative to the spatial discretization error and that the robustness of the fully discrete algorithm is not impacted by the CFL condition for the explicit temporal discretization, we take $C = 0.1$ for all numerical experiments in this section. The remainder of this section will consider a sequence of four test cases, which will be used to assess the accuracy and robustness of the proposed schemes as well as to support the theoretical analysis presented in Section 4. We will present results for an entropy-conservative formulation (abbreviated in this section as "EC") in which both the volume and interface fluxes are computed using (45), as well as an entropy-stable formulation (abbreviated as "ES") in which the volume and interface fluxes are given by (45) and (52), respectively.

---

[2]`https://github.com/trixi-framework/TrixiAtmo.jl`

5.1. Unsteady solid-body rotation

The first case we consider is an unsteady analytical solution to the spherical shallow water equations proposed in [3, Example 3]. To obtain such a solution, we introduce the rotating frame

$$\vec{b}_x(t) := \cos(\Omega t)\vec{e}_x + \sin(\Omega t)\vec{e}_y, \quad \vec{b}_y(t) := -\sin(\Omega t)\vec{e}_x + \cos(\Omega t)\vec{e}_y, \quad \vec{b}_z(t) := \vec{e}_z, \quad (78)$$

as well as the vector-valued function $\vec{\varphi}(\vec{x}, t) := (\vec{x} \cdot \vec{b}_x(t))\vec{e}_x + (\vec{x} \cdot \vec{b}_y(t))\vec{e}_y + (\vec{x} \cdot \vec{b}_z(t))\vec{e}_z$. Defining the Earth's rotation vector as $\vec{\Omega} := \Omega \vec{e}_z$, prescribing the bottom topography as

$$b(\vec{x}) := \frac{1}{2g}(\vec{\Omega} \cdot \vec{x})^2, \quad (79)$$

and defining the fixed vector $\vec{c} := -\sin\alpha \vec{e}_x + \cos\alpha \vec{e}_y$, it can be shown that the following surface height and velocity fields solve the spherical shallow water equations:

$$H(\vec{x}, t) = \frac{1}{g}\left(-\frac{(\vec{\Omega} \cdot \vec{x} + V\,\vec{\varphi}(\vec{c}, t) \cdot \vec{x}/a)^2}{2} + \frac{(\vec{\Omega} \cdot \vec{x})^2}{2} + K\right), \quad \vec{v}(\vec{x}, t) = V\,\vec{\varphi}(\vec{c}, t) \times \frac{\vec{x}}{a}. \quad (80)$$

In our numerical experiments, we take values of $\alpha = \pi/4$, $V = 2\pi a/(12 \cdot 86400\text{ s})$, and $K = 133681$ m^2/s^2 for consistency with the results in [3]. The numerical solution is initialized on each element by evaluating the state vector $\boldsymbol{u}$ in terms of $H(\vec{x}, 0)$, $\vec{v}(\vec{x}, 0)$, and $b(\vec{x})$ at each node position $\vec{x} = \vec{X}(\xi_i, \xi_j)$ in physical space, where (28) is used to transform the velocity vector from global Cartesian to local contravariant components.

Using H_{exact} to denote the analytical height field given in (80), we compute the L^2 error in such a quantity at $t = 5$ days following the conventions suggested in [1], wherein we employ the quadrature formula in (32) and normalize to obtain

$$\text{Normalized } L^2 \text{ height error} := \frac{I_S[(H - H_{\text{exact}})^2]^{1/2}}{I_S[(H_{\text{exact}})^2]^{1/2}}. \quad (81)$$

Holding the polynomial degree N fixed and successively doubling the number of elements per dimension on each face of the cubed sphere, we observe that the ES scheme converges at an optimal rate of $N + 1$, while the EC scheme converges at a rate of N for odd polynomial degrees and $N + 1$ for even polynomial degrees, as shown for $N = 3$ and $N = 4$ in Figure 2a and Figure 2b, respectively. Holding the number of elements fixed and varying the polynomial degree, we observe exponential convergence in Figure 2c for both the EC and ES formulations. In all cases, the error values are smaller for the ES scheme than for the EC scheme. Such results are consistent with the expected convergence behaviour for discontinuous spectral-element discretizations in Euclidean space.

5.2. Flow over an isolated mountain

Our second test case involves an initially zonal flow over an isolated mountain. Following [1, Case 5] we consider a mountain of height $b_{\text{ref}} = 2000$ m centred at a longitude of $\lambda_0 = -\pi/2$ and a latitude of $\theta_0 = \pi/6$, as given by

$$b(\lambda, \theta) := b_{\text{ref}}\left(1 - \sqrt{\min(R^2, (\lambda - \lambda_0)^2 + (\theta - \theta_0)^2)}/R\right), \quad (82)$$

where we take $R = \pi/9$. Representing the velocity field in terms of a spherical basis as $\vec{v} = u\vec{i} + v\vec{j}$, where u and v denote the zonal and meridional velocity components and $\vec{i}$ and $\vec{j}$ denote the

(a) Convergence with mesh refinement for $N = 3$

(b) Convergence with mesh refinement for $N = 4$

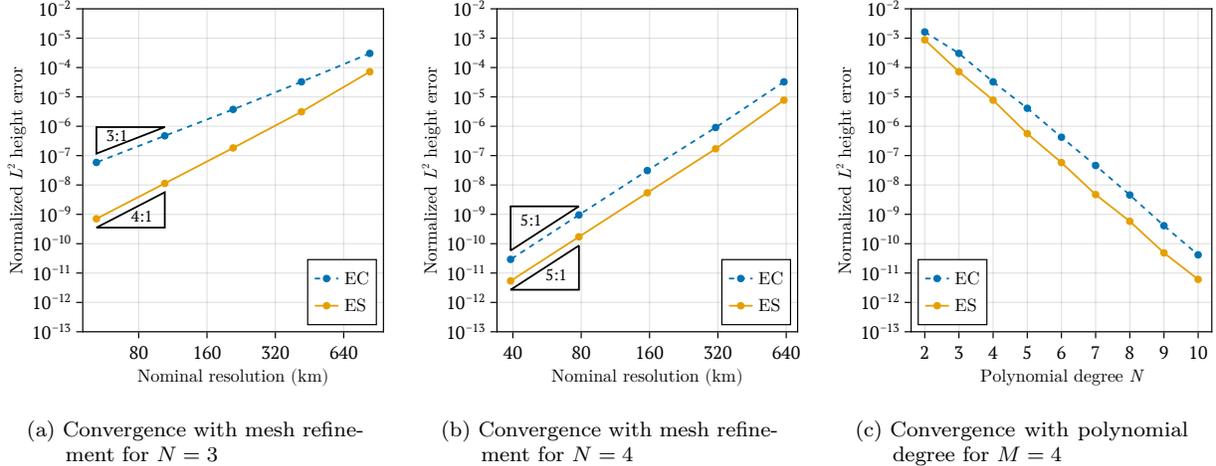(c) Convergence with polynomial degree for $M = 4$

Figure 2: Convergence results for the unsteady solid-body rotation problem at $t = 5$ days

longitudinal and latitudinal basis vectors, respectively, we prescribe the initial values of H, u, and v as functions of longitude and latitude given by

$$H_0(\lambda, \theta) := h_{\text{ref}} - \frac{1}{g}\left(a\Omega V + \frac{1}{2}V^2\right)\sin^2\theta, \quad u_0(\lambda, \theta) := V\cos\theta, \quad v_0(\lambda, \theta) := 0 \text{ m/s}, \quad (83)$$

where we take $h_{\text{ref}} = 5960$ m. In our numerical experiments, we consider both the case of $V = 0$ m/s, for which a well-balanced discretization is expected to maintain the initial condition up to roundoff error, as well as the case of $V = 20$ m/s described in [1]. For initial velocity fields specified in zonal and meridional components as in (83), we first transform to a Cartesian representation as

$$\begin{pmatrix} v_x \\ v_y \\ v_z \end{pmatrix} = \begin{pmatrix} -\sin\lambda & -\cos\lambda\sin\theta \\ \cos\lambda & -\sin\lambda\sin\theta \\ 0 & \cos\theta \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \quad (84)$$

before using (28) to obtain the contravariant components. We do not apply any smoothing to the sharp orography in (82); however, we note that the use of the collocation derivative operator to compute the bottom topography gradient within a flux-differencing formulation effectively approximates b through polynomial interpolation on each element prior to differentiation.

Using $(N, M) = (3, 20)$, as was chosen in several previous studies, including [31, Section 4.3], and beginning with the case of $V = 0$ m/s, we first note that the initial time derivative of the state vector is on the order of machine precision at all nodes for both the EC and ES variants of the scheme, consistent with Theorem 3. Plotting the temporal growth of the normalized L^2 error over a period of 15 days in Figure 3a, we observe a slow deviation from the initial balanced state due to roundoff error accumulation; such error growth occurs more slowly for the ES variant than for the EC variant due to the dissipative effect of the local Lax–Friedrichs numerical flux. In order to assess the conservation properties of the schemes with respect to mass and total energy (i.e., mathematical entropy), we consider the case of $V = 20$ m/s and evaluate the metrics

$$\text{Normalized mass change} := \frac{I_S[h] - I_S[h_0]}{I_S[h_0]}, \quad \text{Normalized entropy change} := \frac{I_S[\eta] - I_S[\eta_0]}{I_S[\eta_0]}, \quad (85)$$

23

(a) Normalized L^2 height error with $V = 0$ m/s

(b) Normalized mass change with $V = 20$ m/s

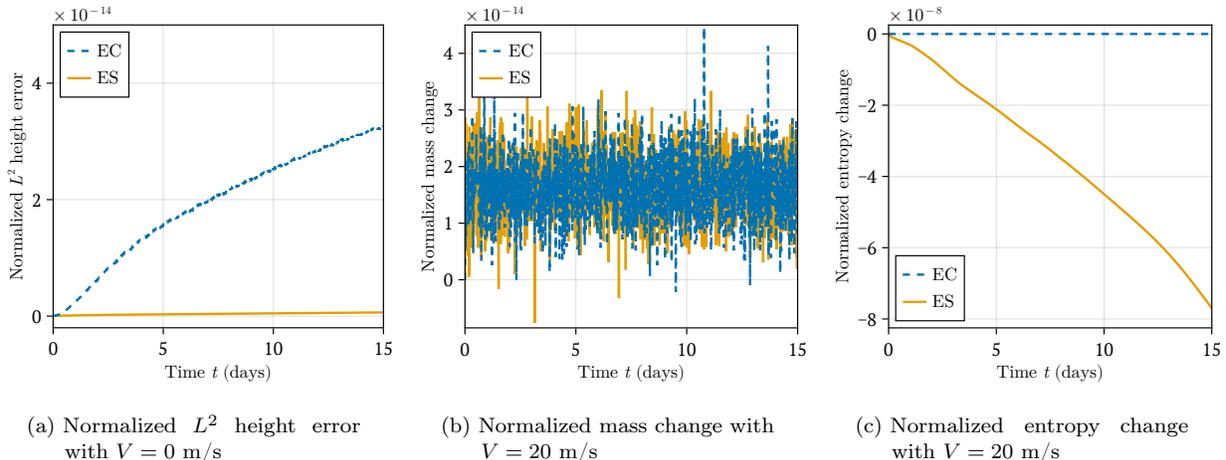(c) Normalized entropy change with $V = 20$ m/s

Figure 3: Error growth, mass conservation, and total energy (i.e., mathematical entropy) conservation or dissipation for the isolated mountain problem with $(N, M) = (3, 20)$

where h_0 and η_0 are the layer depth and mathematical entropy fields evaluated at the initial solution state. Plotting the evolution of such quantities in Figure 3b and Figure 3c, we see that $I_S[h]$ and $I_S[\eta]$ remain constant up to roundoff error for the EC scheme, reflecting the fact that the spatial discretization is conservative of mass and entropy as a consequence of Theorems 1 and 2, while the small time step renders the effect of the temporal discretization on such properties negligible. For the ES scheme, $I_S[h]$ remains constant up to roundoff error, but $I_S[\eta]$ decays monotonically, reflecting discrete mass conservation and entropy stability. These results are consistent with the theoretical analysis in Section 4.

In Figure 4, we plot the relative vorticity field, which is given by

$$\zeta := \varepsilon^{ij} \nabla_i v_j = \frac{1}{J} \big(\partial_1 v_2 - \partial_2 v_1 \big), \tag{86}$$

at $t = 7$ days, considering the case of $V = 20$ m/s. The resulting contours are qualitatively similar to those presented in [24, Fig. 9] and [31, Fig. 11]. While some oscillatory behaviour is observed in the EC case, likely due to the absence of numerical dissipation, such oscillations are nevertheless smaller than those observed in [31, Fig. 11b] for a standard flux-form DG method at the same polynomial degree and resolution, despite the fact that their method made use of a dissipative interface flux. For our ES formulation using the interface flux in (52), which employs the same dissipation term and wave speed estimate used in [31], such oscillations are not present, suggesting that they may be the result of spurious numerical behaviour which is suppressed through the use of a flux-differencing ES formulation rather than a standard DG scheme. Further comparisons to results obtained using standard DG methods will be presented in Sections 5.3 and 5.4.

5.3. Barotropic instability

The third case we consider is a barotropic instability initiated by a perturbation applied to a mid-latitude jet, as proposed in [2]. The initial velocity field is prescribed as a purely zonal flow
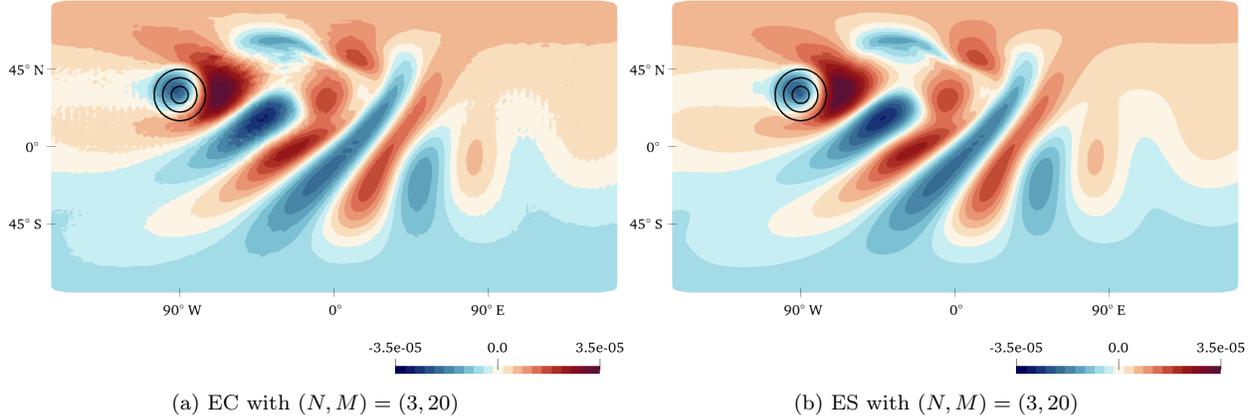
(a) EC with $(N, M) = (3, 20)$ (b) ES with $(N, M) = (3, 20)$

Figure 4: Relative vorticity field (s^{-1}) at $t = 7$ days for the isolated mountain problem with $V = 20$ m/s, where topography contours at 500 m, 1000 m, and 1500 m are shown in black

given by

$$
u_0(\lambda, \theta) := \begin{cases} \dfrac{V \exp((\theta - \theta_0)^{-1}(\theta - \theta_1)^{-1})}{\exp(-4(\theta_1 - \theta_0)^{-2})}, & \theta_0 < \theta < \theta_1, \\ 0, & \text{otherwise,} \end{cases} \tag{87}
$$

and $v_0(\lambda, \theta) = 0$ m/s, where $V = 80$ m/s, $\theta_0 = \pi/7$, and $\theta_1 = \pi/2 - \theta_0$. Taking $h_{\text{ref}} = 10158$ m as in [23] so as to achieve a global mean layer depth of approximately 10000 m as specified in [2], we compute the unperturbed balanced height field as

$$
h_{\text{bal}}(\lambda, \theta) := h_{\text{ref}} - \frac{a}{g} \int_{-\pi/2}^{\theta} u_0(\lambda, \theta') \left(2\Omega \sin(\theta') + u_0(\lambda, \theta') \frac{\tan(\theta')}{a} \right) \, \mathrm{d}\theta', \tag{88}
$$

where the integral is evaluated numerically using `QuadGK.jl` [61], employing an adaptive scheme based on a composite 15-point Gauss–Kronrod quadrature rule (which itself embeds a seven-point Legendre–Gauss quadrature rule) with a relative tolerance of $\sqrt{\varepsilon}$, where $\varepsilon \approx 2.22 \times 10^{-16}$ denotes machine epsilon in double-precision arithmetic. Adding a small perturbation to the depth field to trigger the instability, we obtain

$$
h_0(\lambda, \theta) = \begin{cases} h_{\text{bal}}(\lambda, \theta) + \delta h \cos \theta \exp(-(\lambda/\alpha)^2) \exp(-((\theta_2 - \theta)/\beta)^2), & -\pi < \lambda < \pi, \\ h_{\text{bal}}(\lambda, \theta), & \text{otherwise,} \end{cases} \tag{89}
$$

where $\alpha = 1/3$, $\beta = 1/15$, $\theta_2 = \pi/4$, and we consider the unperturbed case of $\delta h = 0$ m as well as the perturbed case of $\delta h = 120$ m, as recommended in [2].

Considering a polynomial degree of $N = 3$ and successively refined grids with $M = 16$, $M = 32$, and $M = 64$, which, according to (76), correspond to nominal resolutions of approximately 208 km, 104 km, and 52 km, respectively, we evolve the unperturbed solution forward in time and plot the evolution of the normalized L^2 error (based on the deviation from the balanced initial state) for the EC and ES schemes over a period of 12 days in Figure 5. As the well balancing property established in Theorem 3 applies only to the rest-state equilibrium in (73), the discretization is not expected to preserve the initial condition for all time, and we indeed observe significant deviation in all cases due to numerical perturbations which are introduced through discretization error, roundoff error, as
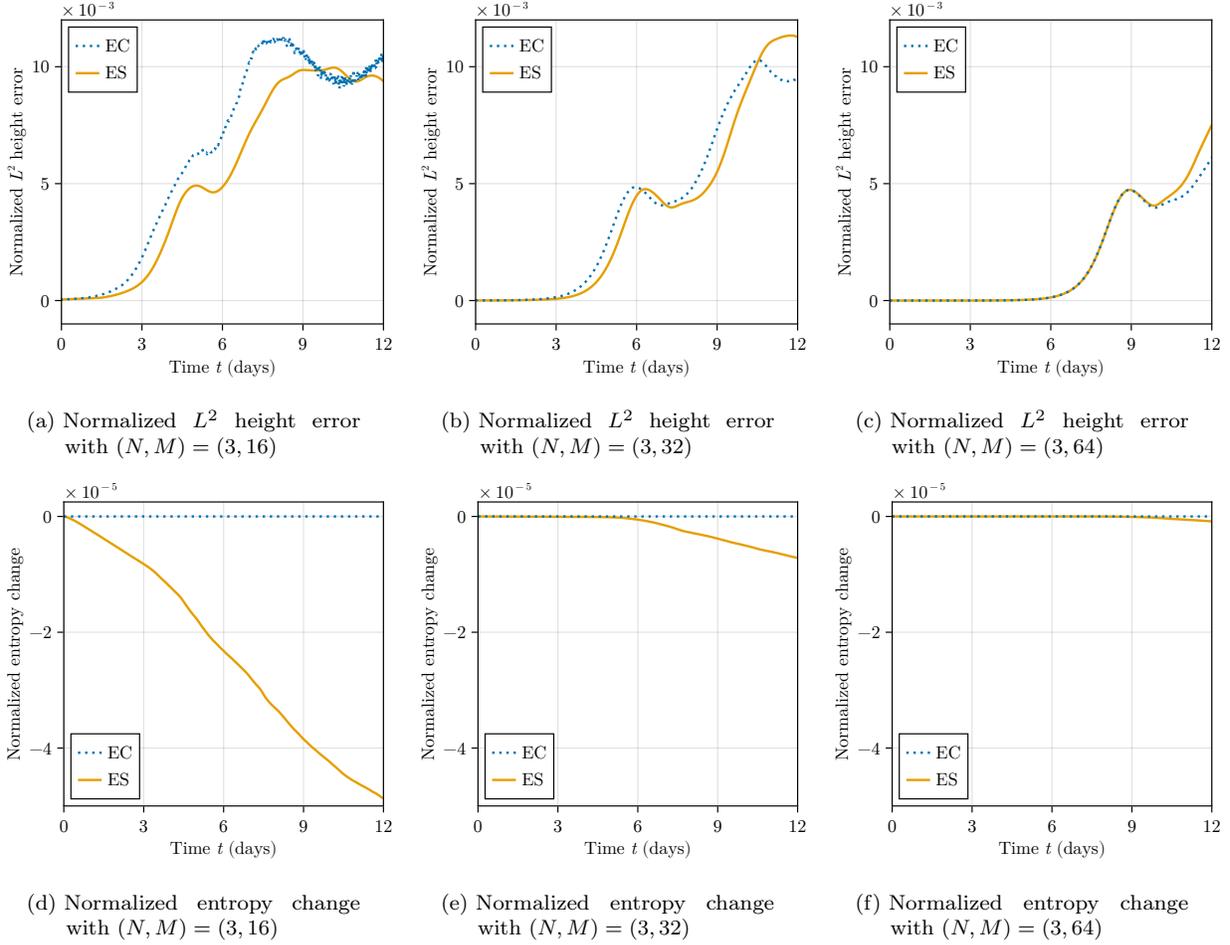
25

(a) Normalized L^2 height error with $(N, M) = (3, 16)$

(b) Normalized L^2 height error with $(N, M) = (3, 32)$

(c) Normalized L^2 height error with $(N, M) = (3, 64)$

(d) Normalized entropy change with $(N, M) = (3, 16)$

(e) Normalized entropy change with $(N, M) = (3, 32)$

(f) Normalized entropy change with $(N, M) = (3, 64)$

Figure 5: Error growth and total energy (i.e., mathematical entropy) conservation or dissipation for the unperturbed barotropic instability problem

well as the representation of the initial condition within a finite-dimensional approximation space. However, the onset of such rapid error growth is shifted later in time as the mesh is refined, and the integrated total energy $I_S[\eta]$ is nevertheless conserved for the EC scheme and dissipated for the ES scheme, in accordance with Theorem 2.

While we have thus far used a small time step so as to neglect temporal discretization error in our numerical experiments, it should be emphasized that the proposed discretizations only guarantee entropy conservation or dissipation at the semi-discrete level, requiring additional techniques (e.g., relaxation Runge–Kutta methods [62, 63]) to ensure fully-discrete entropy stability. To examine the influence of the temporal discretization on mass and entropy conservation, we ran the unperturbed barotropic instability problem using an EC spatial discretization with $(N, M) = (3, 16)$ for fixed time step sizes of $\Delta t \in \{10 \text{ s}, 20 \text{ s}, 40 \text{ s}, 80 \text{ s}, 160 \text{ s}, 320 \text{ s}\}$. Plotting the absolute values of the quantities defined in (85), we see in Figures 6a and 6b that although mass is conserved up to roundoff error for all time step sizes, as expected for linear invariants of ODEs discretized with Runge–Kutta methods [64], the entropy conservation error (which we note to be primarily dissipative here) increases with time more rapidly for larger time step sizes. Plotting the normalized absolute change in both

26

(a) Mass conservation   (b) Entropy conservation   (c) Convergence with Δt
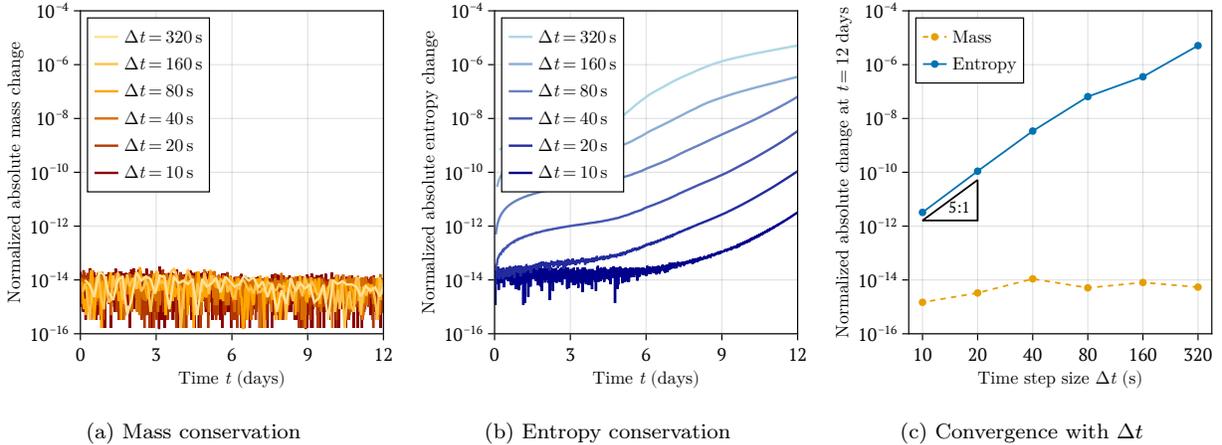
Figure 6: Effect of time step size on mass and entropy conservation using an EC spatial discretization with $(N, M) = (3, 16)$ for the unperturbed barotropic instability problem

quantities at $t = 12$ days in Figure 6c, we observe fifth-order temporal convergence of the entropy error, which is consistent with Chan's observations in [65, Fig. 2c] for an entropy-conservative spatial discretization of the compressible Euler equations using the same fourth-order low-storage explicit Runge–Kutta method employed in this work.

In Figure 7, we plot the relative vorticity field given by (86) at $t = 6$ days for an unperturbed initial condition, viewed as an orthographic projection of the northern hemisphere with the $\lambda = 0$ meridian pointed downwards. For the lower-resolution simulations with $M = 16$ and $M = 32$, nonphysical flow features (i.e., grid imprinting) associated with numerical perturbation growth can be seen, with significantly more high-frequency noise observed for the EC scheme than for the ES scheme. At $M = 64$, the initial condition is preserved quite well, suggesting that such a resolution would likely be sufficient for qualitatively preserving the physical flow features arising from a perturbed initial condition. Comparing the contours in Figure 8 resulting from the perturbed initial condition to reference solutions such as those in [2], we see that this is indeed the case, with $M = 64$ being sufficient to obtain vorticity fields which exhibit the qualitative features observed in the literature, without interference by artificial flow features initiated by numerical perturbations absent in the analytical definition of the initial condition.

Although nominal resolution as defined in (76) is not a perfectly generalizable measure of a scheme's capability to capture fine-scale flow features, particularly when comparing methods with different orders of accuracy, we note that the 52 km nominal resolution obtained for $(N, M) = (3, 64)$ is approximately the same as that used for a standard DG method in [31, Section 4.2], where their choice of $(N, M) = (7, 30)$ corresponds to a nominal resolution of approximately 48 km. Furthermore, the elimination of visible grid imprinting between resolutions of 104 km and 52 km is roughly consistent with the findings of St-Cyr *et al.* [66, Fig. 13], wherein grid imprinting is visible for a continuous spectral-element method at a reported resolution of 1.25° (approximately 139 km) but not at 0.625° (approximately 69 km).

To examine the influence of the discretization on the solution quality more closely, we show in Figure 9 close-up views of a portion of the vorticity field for the EC and ES schemes, as well as for a standard DG method using the same cubed-sphere grid, tensor-product polynomial approximation space, and collocated LGL quadrature. The standard DG method discretizes (11) using the weak
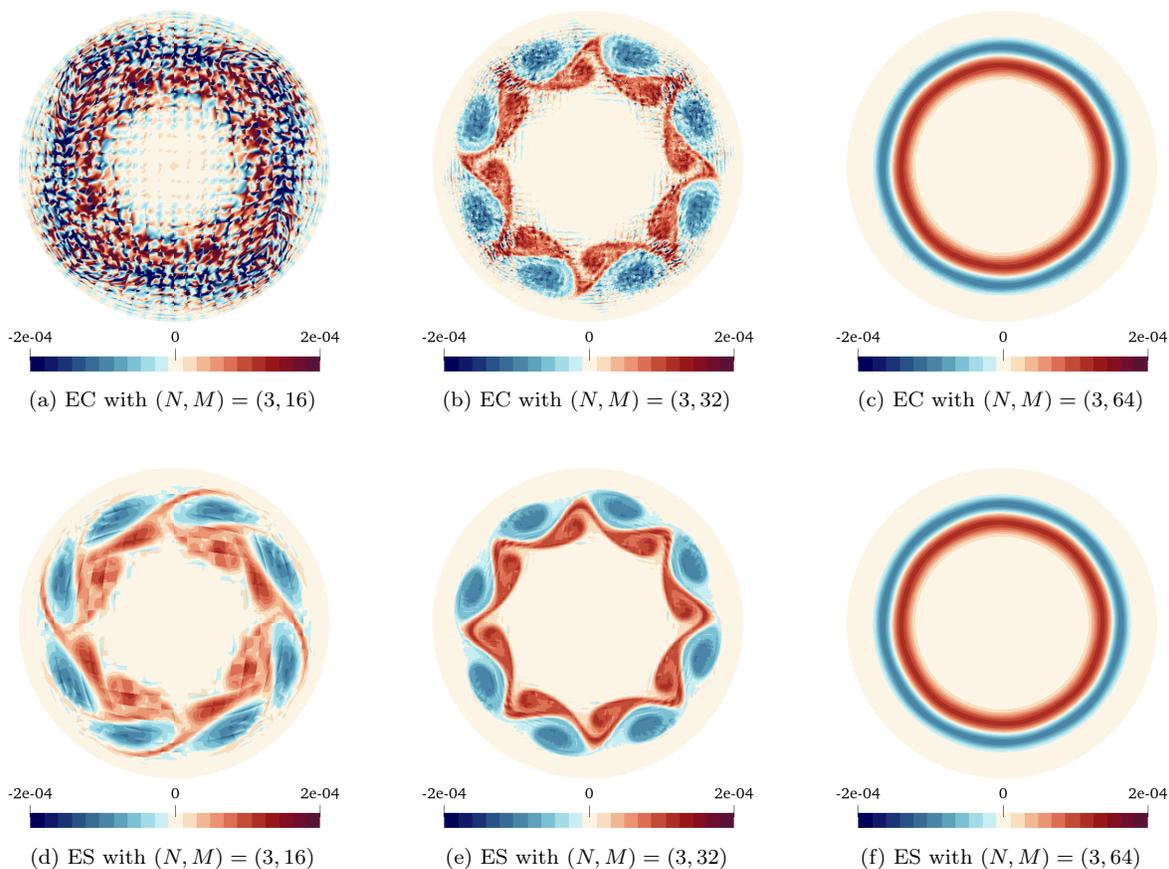
(a) EC with $(N, M) = (3, 16)$ (b) EC with $(N, M) = (3, 32)$ (c) EC with $(N, M) = (3, 64)$

(d) ES with $(N, M) = (3, 16)$ (e) ES with $(N, M) = (3, 32)$ (f) ES with $(N, M) = (3, 64)$

Figure 7: Relative vorticity field (s^{-1}) at $t = 6$ days for the unperturbed barotropic instability problem

formulation in (43), without skew-symmetric splitting or flux differencing, and employs a local Lax–Friedrichs interface flux given by

$$(J\boldsymbol{f}^j)^*_{L,R} := \frac{1}{2}\left[(J\boldsymbol{f}^j)_L + (J\boldsymbol{f}^j)_R\right] - \frac{J_L}{2}\max(\lambda_L^j, \lambda_R^j)(\boldsymbol{u}_R - \boldsymbol{u}_L), \tag{90}$$

where we note that the dissipation term is identical to that employed for the ES scheme in (52). While the EC scheme exhibits oscillatory behaviour of a similar nature to that observed in the isolated mountain case due to the lack of any mechanism to dissipate high-frequency noise, it is nevertheless remarkable that a dissipation-free numerical method is able to remain stable for a relatively complex model problem for a geophysical flow exhibiting multiscale behaviour and steep solution gradients. These oscillations are greatly reduced through the introduction of the dissipation term to the EC scheme to obtain the ES formulation. While the standard DG method was able to run to completion for this case, we nevertheless observe some spurious behaviour in Figure 9c which is not present in Figure 9b, qualitatively illustrating the benefit of the flux-differencing formulation in this particular case. We will present quantitative robustness comparisons in the following subsection.
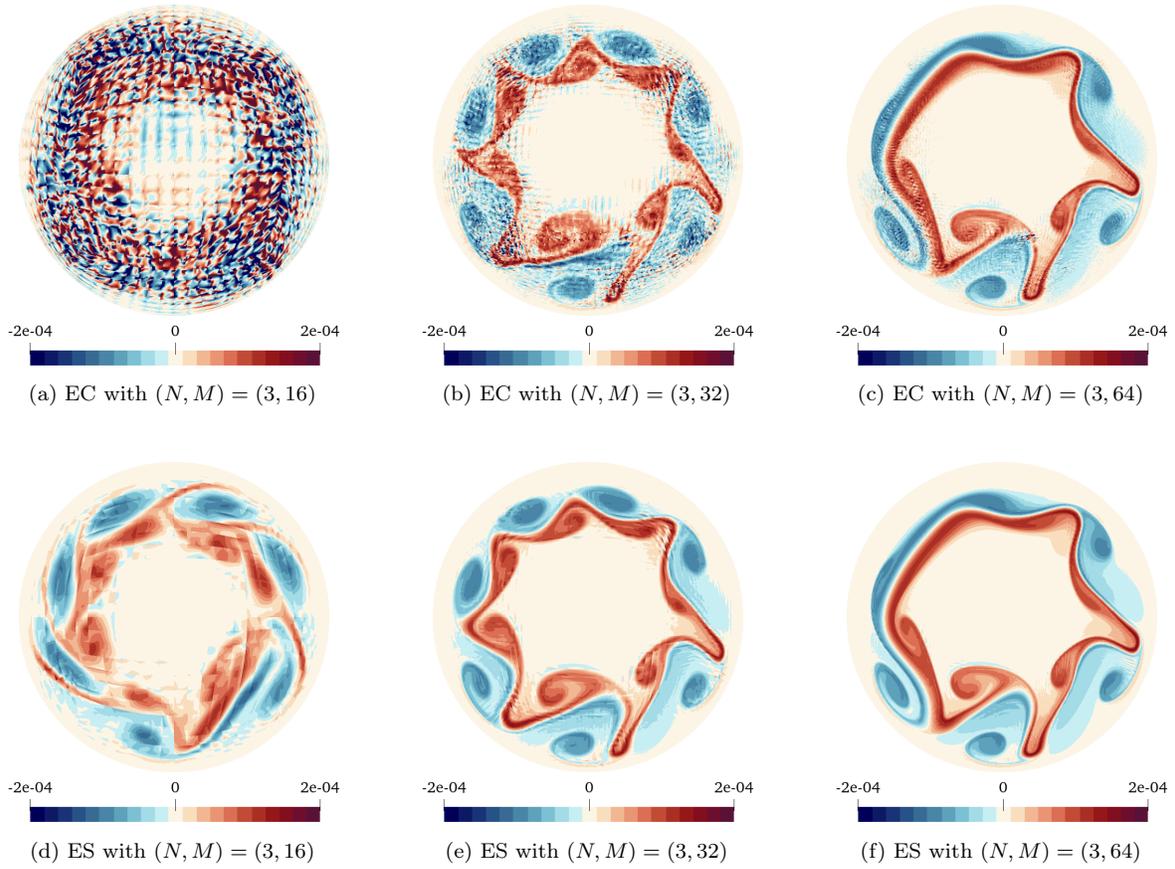
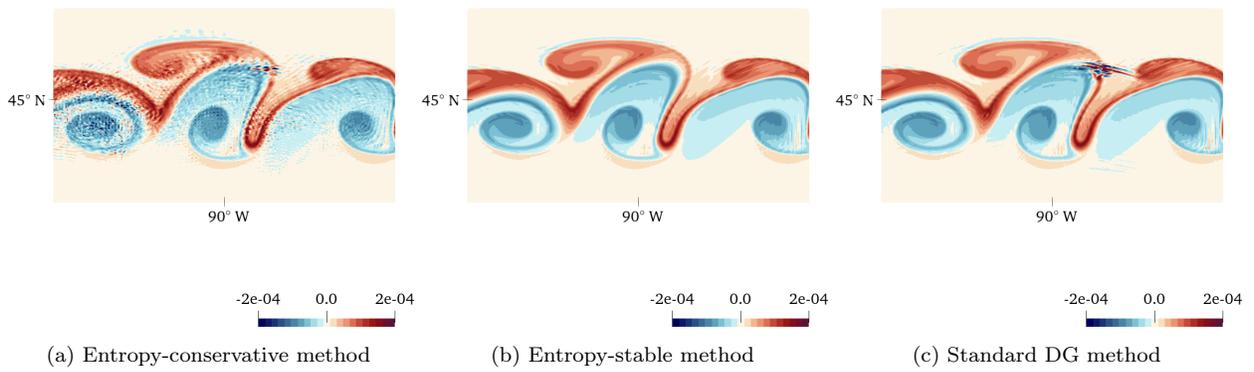Figure 8: Relative vorticity field (s^{-1}) at $t = 6$ days for the perturbed barotropic instability problem

Figure 9: Comparison of relative vorticity fields (s^{-1}) at $t = 6$ days with $(N, M) = (3, 64)$

5.4. Rossby–Haurwitz wave

As a fourth and final test case, we consider a Rossby–Haurwitz wave with wavenumber four, as described in [1, Case 6]. Taking parameter values of $\omega = K = 7.848 \times 10^{-6}$ s^{-1}, $R = 4$, and $h_{\text{ref}} = 8000$ m, the functions

$$A(\theta) := \frac{\omega}{2}(2\Omega + \omega)\cos^2\theta + \frac{K^2}{4}\cos^{2R}\theta\left((R+1)\cos^2\theta + (2R^2 - R - 2) - 2\frac{R^2}{\cos^2\theta}\right), \quad (91a)$$

$$B(\theta) := \frac{2(\Omega + \omega)K}{(R+1)(R+2)}\cos^R\theta\left((R^2 + 2R + 2) - (R+1)^2\cos^2\theta\right), \quad (91b)$$

$$C(\theta) := \frac{1}{4}K^2\cos^{2R}\theta\left((R+1)\cos^2\theta - (R+2)\right), \quad (91c)$$

are used to prescribe the initial depth and spherical velocity components as

$$h_0(\lambda, \theta) = h_{\text{ref}} + \frac{a^2}{g}\left(A(\theta) + B(\theta)\cos(R\lambda) + C(\theta)\cos(2R\lambda)\right), \quad (92a)$$

$$u_0(\lambda, \theta) = a\omega\cos\theta + aK\cos^{R-1}\theta(R\sin^2\theta - \cos^2\theta)\cos(R\lambda), \quad (92b)$$

$$v_0(\lambda, \theta) = -aKR\cos^{R-1}\theta\sin\theta\sin(R\lambda). \quad (92c)$$

While it is common to run such a test case until $t = 7$ days or $t = 14$ days, we run for 28 days as a more severe stress test of the schemes' robustness for predictions over longer time horizons. Furthermore, as a baseline for such robustness comparisons, we also report results using standard collocated DG methods employing LGL quadrature and a local Lax–Friedrichs numerical flux, as described in Section 5.3.

Considering the cases of $(N, M) = (3, 16)$ as well as that of $(N, M) = (6, 8)$, which differ in polynomial degree but correspond to the same nominal resolution of approximately 208 km, we plot the entropy conservation error for the EC, ES, and standard DG schemes in Figures 10a and 10b, with vertical lines indicating crash times. When applied to the spherical shallow water equations, the standard DG scheme is conservative of mass, but does not satisfy any discrete entropy conservation or stability properties. We accordingly observe uncontrolled growth in the total energy for the standard DG schemes, with crashes occurring at $t = 18.71$ days and $t = 13.39$ days for the $(N, M) = (3, 16)$ and $(N, M) = (6, 8)$ simulations respectively. For both polynomial degrees considered, the EC method is able to run for longer than the corresponding standard DG scheme, but nevertheless eventually crashes due to its inability to damp oscillations, which eventually leads to negative values for the layer depth. As with the standard DG method, the $N = 3$ EC scheme runs for longer than the $N = 6$ EC scheme at an equal nominal resolution, with crashes occurring at $t = 26.03$ days and $t = 18.01$ days. Unlike the standard and EC formulations, the ES schemes were capable of running the full 28-day simulations to completion for both $(N, M) = (3, 16)$ and $(N, M) = (6, 8)$, indicating that the local Lax–Friedrichs interface dissipation term in (52) provided sufficient damping of the oscillations that led to negative depth values when the EC scheme was applied to this problem. Examining the evolution of the discrete integral of the potential enstrophy $Z := (\zeta + f)^2/h$, where ζ is the relative vorticity defined in (86), we see in Figures 10c and 10d that the aforementioned oscillatory behaviour manifests for the standard DG and EC schemes in the form of a gradual buildup of grid-scale vorticity, beginning well before the simulations crashed. While the ES scheme does not guarantee strict conservation or dissipation of potential enstrophy or other vorticity-derived invariants, the integrated potential enstrophy deviated by no more than approximately 0.1% over
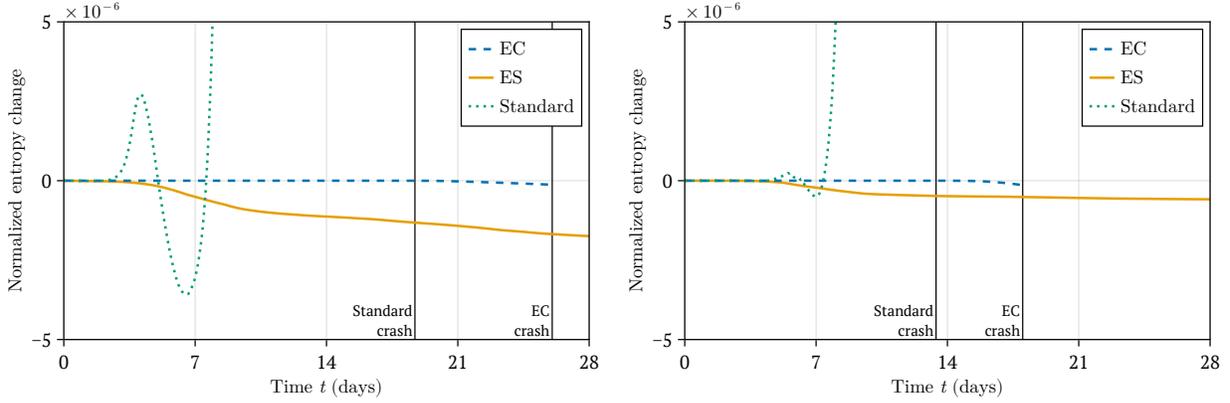
(a) Normalized entropy change and crash times with $(N, M) = (3, 16)$

(b) Normalized entropy change and crash times with $(N, M) = (6, 8)$
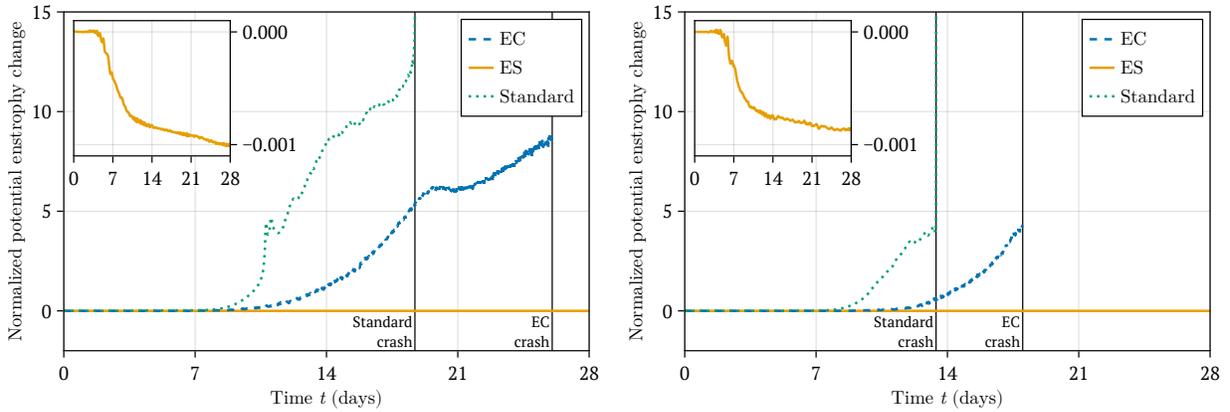
(c) Normalized potential enstrophy change and crash times with $(N, M) = (3, 16)$

(d) Normalized potential enstrophy change and crash times with $(N, M) = (6, 8)$

Figure 10: Robustness of EC, ES, and standard collocated DG schemes for the Rossby–Haurwitz wave

the course of each of the 28-day simulations, with such deviations being almost entirely dissipative in nature. These results indicate that the interface dissipation term was sufficient in this case to suppress spurious potential enstrophy production.

6. Conclusions

A new discontinuous spectral-element formulation was devised for the rotating shallow water equations with variable bottom topography expressed in covariant form on general unstructured quadrilateral tessellations of curved two-dimensional manifolds. The approach relies on the use of tensor-product summation-by-parts operators within a flux-differencing formulation in reference coordinates derived from a skew-symmetric splitting of the divergence of the momentum flux tensor in covariant form. The resulting methods are shown to be conservative of mass, and, depending on the choice of numerical interface flux, conservative or dissipative of total energy, which serves as a mathematical entropy function for the shallow water system.

The proposed schemes were applied to a series of standard model problems for idealized atmospheric dynamics based on the shallow water equations on the sphere, including an unsteady solid-body rotation, a zonal flow over an isolated mountain, a barotropic instability, and a Rossby–Haurwitz wave. Using an entropy-conservative numerical flux at element interfaces, the resulting schemes conserve mass as well as total energy, serving as a robust yet dissipation-free baseline to which numerical dissipation can be added judiciously, for example, to reduce oscillations. When such dissipation is introduced through an entropy-stable interface flux, the resulting schemes guarantee conservation of mass and dissipation of total energy, attain optimal convergence (i.e., order $N + 1$ for a tensor-product approximation of polynomial degree N), and exhibit excellent robustness even under challenging conditions including those involving under-resolved vortical structures and integration over long time horizons.

The advances in this paper represent a first step in the process of developing a new atmospheric dynamical core for numerical weather prediction and climate modelling. Future work will involve the extension of the proposed approach to the three-dimensional nonhydrostatic compressible Euler equations using an appropriate vertical discretization and a horizontally-explicit vertically-implicit temporal discretization [22]. Additionally, we are investigating extensions to triangular and prismatic elements through the use of multidimensional SBP formulations (as introduced by Hicken *et al.* [67] and reviewed within an entropy-stable setting by Chen and Shu [68]) as well as entropy-stable tensor-product formulations in collapsed coordinates, which were introduced by Montoya and Zingg [69] and extended to prismatic elements by Keim *et al.* [70]. We would also like to investigate techniques for preserving vorticity-based invariants such as potential enstrophy within the present framework. Additionally, we hope to extend the proposed methodology beyond the context of geophysical fluid dynamics, for example, to covariant formulations of hyperbolic partial differential equations arising in general relativity and relativistic hydrodynamics, as well as to the treatment of curved elements in computational fluid dynamics for engineering applications.

Declaration of generative AI in scientific writing

During proofreading of this article, ChatGPT (OpenAI) was used to assist in identifying spelling, grammar, and clarity issues. All manuscript text was written and thoroughly reviewed by the authors, who take full responsibility for its content.

References

[1] D. L. Williamson, J. B. Drake, J. J. Hack, R. Jakob, P. N. Swarztrauber, A standard test set for numerical approximations to the shallow water equations in spherical geometry, Journal of Computational Physics 102 (1) (1992) 211–224. `doi:10.1016/s0021-9991(05)80016-6`.

[2] J. Galewsky, R. K. Scott, L. M. Polvani, An initial-value problem for testing numerical models of the global shallow-water equations, Tellus A 56 (5) (2004) 429–440. `doi:10.1111/j.1600-0870.2004.00071.x`.

[3] M. Läuter, D. Handorf, K. Dethloff, Unsteady analytical solutions of the spherical shallow water equations, Journal of Computational Physics 210 (2) (2005) 535–553. `doi:10.1016/j.jcp.2005.04.022`.

[4] M. Taylor, J. Tribbia, M. Iskandarani, The spectral element method for the shallow water equations on the sphere, Journal of Computational Physics 130 (1) (1997) 92–108. `doi:10.1006/jcph.1996.5554`.

[5] F. X. Giraldo, A spectral element shallow water model on spherical geodesic grids, International Journal for Numerical Methods in Fluids 35 (8) (2001) 869–901. `doi:10.1002/1097-0363(20010430)35:8<869::aid-fld116>3.0.co;2-s`.

[6] S. J. Thomas, R. D. Loft, Semi-implicit spectral element atmospheric model, Journal of Scientific Computing 17 (1-4) (2002) 339–350. `doi:10.1023/a:1015129420882`.

[7] F. X. Giraldo, J. S. Hesthaven, T. Warburton, Nodal high-order discontinuous Galerkin methods for the spherical shallow water equations, Journal of Computational Physics 181 (2) (2002) 499–525. `doi:10.1006/jcph.2002.7139`.

[8] R. D. Nair, S. J. Thomas, R. D. Loft, A discontinuous Galerkin global shallow water model, Monthly Weather Review 133 (4) (2005) 876–888. `doi:10.1175/mwr2903.1`.

[9] S. Marras, J. F. Kelly, M. Moragues, A. Müller, M. A. Kopera, M. Vázquez, F. X. Giraldo, G. Houzeaux, O. Jorba, A review of element-based Galerkin methods for numerical weather prediction: Finite elements, spectral elements, and discontinuous Galerkin, Archives of Computational Methods in Engineering 23 (4) (2016) 673–722. `doi:10.1007/s11831-015-9152-1`.

[10] R. M. Kirby, G. E. Karniadakis, De-aliasing on non-uniform grids: Algorithms and applications, Journal of Computational Physics 191 (1) (2003) 249–264. `doi:10.1016/s0021-9991(03)00314-0`.

[11] G. Mengaldo, D. De Grazia, D. Moxey, P. E. Vincent, S. J. Sherwin, Dealiasing techniques for high-order spectral element methods on regular and irregular grids, Journal of Computational Physics 299 (2015) 56–81. `doi:10.1016/j.jcp.2015.06.032`.

[12] G. J. Gassner, A. R. Winters, D. A. Kopriva, Split form nodal discontinuous Galerkin schemes with summation-by-parts property for the compressible Euler equations, Journal of Computational Physics 327 (2016) 39–66. `doi:10.1016/j.jcp.2016.09.013`.

[13] A. R. Winters, R. C. Moura, G. Mengaldo, G. J. Gassner, S. Walch, J. Peiro, S. J. Sherwin, A comparative study on polynomial dealiasing and split form discontinuous Galerkin schemes for under-resolved turbulence computations, Journal of Computational Physics 372 (2018) 1–21. `doi:10.1016/j.jcp.2018.06.016`.

[14] H.-O. Kreiss, G. Scherer, Finite element and finite difference methods for hyperbolic partial differential equations, in: C. de Boor (Ed.), Mathematical Aspects of Finite Elements in Partial Differential Equations, Academic Press, 1974, pp. 195–212. `doi:10.1016/b978-0-12-208350-1.50012-1`.

[15] E. Tadmor, The numerical viscosity of entropy stable schemes for systems of conservation laws. I, Mathematics of Computation 49 (179) (1987) 91–103. `doi:10.2307/2008251`.

[16] P. G. LeFloch, J.-M. Mercier, C. Rohde, Fully discrete, entropy conservative schemes of arbitrary order, SIAM Journal on Numerical Analysis 40 (5) (2002) 1968–1992. `doi:10.1137/s003614290240069x`.

[17] T. C. Fisher, High-order L^2 stable multi-domain finite difference method for compressible flows, Ph.D. thesis, Purdue University (2012).

[18] G. J. Gassner, A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods, SIAM Journal on Scientific Computing 35 (3) (2013) A1233–A1253. `doi:10.1137/120890144`.

[19] G. J. Gassner, A. R. Winters, A novel robust strategy for discontinuous Galerkin methods in computational fluid mechanics: Why? when? what? where?, Frontiers in Physics 8 (2021). `doi:10.3389/fphy.2020.500690`.

[20] K. Ricardo, D. Lee, K. Duru, Conservation and stability in a discontinuous Galerkin method for the vector invariant spherical shallow water equations, Journal of Computational Physics 500 (2024) 112763. `doi:10.1016/j.jcp.2024.112763`.

[21] M. Waruszewski, J. E. Kozdon, L. C. Wilcox, T. H. Gibson, F. X. Giraldo, Entropy stable discontinuous Galerkin methods for balance laws in non-conservative form: Applications to the Euler equations with gravity, Journal of Computational Physics 468 (2022) 111507. `doi:10.1016/j.jcp.2022.111507`.

[22] M. Baldauf, A horizontally explicit, vertically implicit (HEVI) discontinuous Galerkin scheme for the 2-dimensional Euler and Navier–Stokes equations using terrain-following coordinates, Journal of Computational Physics 446 (2021) 110635. `doi:10.1016/j.jcp.2021.110635`.

[23] M. Baldauf, Discontinuous Galerkin solver for the shallow-water equations in covariant form on the sphere and the ellipsoid, Journal of Computational Physics 410 (2020). `doi:10.1016/j.jcp.2020.109384`.

[24] S. Gaudreault, M. Charron, V. Dallerit, M. Tokman, High-order numerical solutions to the shallow-water equations on the rotated cubed-sphere grid, Journal of Computational Physics 449 (2022) 110792. `doi:10.1016/j.jcp.2021.110792`.

[25] Y. Kawai, H. Tomita, Development of a high-order global dynamical core using the discontinuous Galerkin method for an atmospheric large-eddy simulation (LES) and proposal of test cases: SCALE-DG v0.8.0, Geoscientific Model Development 18 (3) (2025) 725–762. `doi:10.5194/gmd-18-725-2025`.

[26] G. J. Gassner, A. R. Winters, D. A. Kopriva, A well balanced and entropy conservative discontinuous Galerkin spectral element method for the shallow water equations, Applied Mathematics and Computation 272 (2016) 291–308. `doi:10.1016/j.amc.2015.07.014`.

[27] N. Wintermeyer, A. R. Winters, G. J. Gassner, D. A. Kopriva, An entropy stable nodal discontinuous Galerkin method for the two-dimensional shallow water equations on unstructured curvilinear meshes with discontinuous bathymetry, Journal of Computational Physics 340 (2017) 200–242. `doi:10.1016/j.jcp.2017.03.036`.

[28] R. Aris, Vectors, Tensors, and the Basic Equations of Fluid Mechanics, Dover Publications, 1962.

[29] C. W. Misner, K. S. Thorne, J. A. Wheeler, Gravitation, W. H. Freeman and Company, 1973.

[30] M. Läuter, F. X. Giraldo, D. Handorf, K. Dethloff, A discontinuous Galerkin method for the shallow water equations in spherical triangular coordinates, Journal of Computational Physics 227 (24) (2008) 10226–10242. `doi:10.1016/j.jcp.2008.08.019`.

[31] L. Bao, R. D. Nair, H. M. Tufo, A mass and momentum flux-form high-order discontinuous Galerkin shallow water model on the cubed-sphere, Journal of Computational Physics 271 (2014) 224–243. `doi:10.1016/j.jcp.2013.11.033`.

[32] Y. Kuang, K. Wu, H. Tang, Runge–Kutta discontinuous local evolution Galerkin methods for the shallow water equations on the cubed-sphere grid, Numerical Mathematics: Theory, Methods and Applications 10 (2) (2017) 373–419. `doi:10.4208/nmtma.2017.s09`.

[33] J. A. Rossmanith, A wave propagation method for hyperbolic systems on the sphere, Journal of Computational Physics 213 (2) (2006) 629–658. `doi:10.1016/j.jcp.2005.08.027`.

[34] P. A. Ullrich, C. Jablonowski, B. van Leer, High-order finite-volume methods for the shallow-water equations on the sphere, Journal of Computational Physics 229 (17) (2010) 6104—6134. `doi:10.1016/j.jcp.2010.04.044`.

[35] E. Tadmor, Skew-selfadjoint form for systems of conservation laws, Journal of Mathematical Analysis and Applications 103 (2) (1984) 428–442. `doi:10.1016/0022-247x(84)90139-2`.

[36] M. Ben-Artzi, P. G. LeFloch, Well-posedness theory for geometry-compatible hyperbolic conservation laws on manifolds, Annales de l'I.H.P. Analyse non linéaire 24 (6) (2007) 989–1008. `doi:10.1016/j.anihpc.2006.10.004`.

[37] O. Guba, M. A. Taylor, P. A. Ullrich, J. R. Overfelt, M. N. Levy, The spectral element method (SEM) on variable-resolution grids: Evaluating grid sensitivity and resolution-aware numerical viscosity, Geoscientific Model Development 7 (6) (2014) 2803–2816. `doi:10.5194/gmd-7-2803-2014`.

[38] D. C. Del Rey Fernández, J. E. Hicken, D. W. Zingg, Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations, Computers & Fluids 95 (2014) 171–196. `doi:10.1016/j.compfluid.2014.02.016`.

[39] M. Svärd, J. Nordström, Review of summation-by-parts schemes for initial-boundary-value problems, Journal of Computational Physics 268 (2014) 17–38. `doi:10.1016/j.jcp.2014.02.031`.

[40] F. Renac, Entropy stable DGSEM for nonlinear hyperbolic systems in nonconservative form with application to two-phase flows, Journal of Computational Physics 382 (2019) 1–26. `doi:10.1016/j.jcp.2018.12.035`.

[41] A. M. Rueda-Ramírez, F. Hindenlang, J. Chan, G. J. Gassner, Entropy-stable Gauss collocation methods for ideal magneto-hydrodynamics, Journal of Computational Physics 475 (2023) 111851. `doi:10.1016/j.jcp.2022.111851`.

[42] M. H. Carpenter, D. Gottlieb, S. Abarbanel, Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes, Journal of Computational Physics 111 (2) (1994) 220–236. `doi:10.1006/jcph.1994.1057`.

[43] J. Chan, K. Shukla, X. Wu, R. Liu, P. Nalluri, High-order entropy stable schemes for the quasi-one-dimensional shallow water and compressible Euler equations, Journal of Computational Physics 504 (2024) 112876. `doi:10.1016/j.jcp.2024.112876`.

[44] H. Ranocha, Comparison of some entropy conservative numerical fluxes for the Euler equations, Journal of Scientific Computing 76 (1) (2018) 216–242. `doi:10.1007/s10915-017-0618-1`.

[45] H. Ranocha, Shallow water equations: Split-form, entropy stable, well-balanced, and positivity preserving numerical methods, GEM-International Journal on Geomathematics 8 (1) (2017) 85–133. `doi:10.1007/s13137-016-0089-9`.

[46] T. C. Fisher, M. H. Carpenter, J. Nordström, N. K. Yamaleev, C. Swanson, Discretely conservative finite-difference formulations for nonlinear conservation laws in split form: Theory and boundary conditions, Journal of Computational Physics 234 (2013) 353–375. `doi:10.1016/j.jcp.2012.09.026`.

[47] T. C. Fisher, M. H. Carpenter, High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains, Journal of Computational Physics 252 (2013) 518–557. `doi:10.1016/j.jcp.2013.06.014`.

[48] M. H. Carpenter, T. C. Fisher, E. J. Nielsen, S. H. Frankel, Entropy stable spectral collocation schemes for the Navier–Stokes equations: Discontinuous interfaces, SIAM Journal on Scientific Computing 36 (5) (2014) B835–B867. `doi:10.1137/130932193`.

[49] A. M. Rueda-Ramírez, W. Pazner, G. J. Gassner, Subcell limiting strategies for discontinuous Galerkin spectral element methods, Computers & Fluids 247 (2022) 105627. `doi:10.1016/j.compfluid.2022.105627`.

[50] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, B. Perthame, A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows, SIAM Journal on Scientific Computing 25 (6) (2004) 2050–2065. `doi:10.1137/s1064827503431090`.

[51] U. S. Fjordholm, S. Mishra, E. Tadmor, Well-balanced and energy stable schemes for the shallow water equations with discontinuous topography, Journal of Computational Physics 230 (14) (2011) 5587–5609. `doi:https://doi.org/10.1016/j.jcp.2011.03.042`.

[52] J. Bezanson, A. Edelman, S. Karpinski, V. B. Shah, Julia: A fresh approach to numerical computing, SIAM Review 59 (1) (2017) 65–98. `doi:10.1137/141000671`.

[53] H. Ranocha, M. Schlottke-Lakemper, A. R. Winters, E. Faulhaber, J. Chan, G. J. Gassner, Adaptive numerical simulations with Trixi.jl: A case study of Julia for scientific computing, Proceedings of the JuliaCon Conferences 1 (1) (2022) 77. `doi:10.21105/jcon.00077`.

[54] T. Montoya, A. M. Rueda-Ramírez, G. J. Gassner, Reproducibility repository for "Entropy-stable discontinuous spectral-element methods for the spherical shallow water equations in covariant form", `https://github.com/tristanmontoya/paper-2025-spherical-shallow-water` (2025). `doi:10.5281/zenodo.17059901`.

[55] C. Ronchi, R. Iacono, P. S. Paolucci, The "cubed sphere": A new method for the solution of partial differential equations in spherical geometry, Journal of Computational Physics 124 (1) (1996) 93–114. `doi:10.1006/jcph.1996.0047`.

[56] R. Sadourny, Conservative finite-difference approximations of the primitive equations on quasi-uniform spherical grids, Monthly Weather Review 100 (2) (1972) 136–144. `doi:10.1175/1520-0493(1972)100<0136:cfaotp>2.3.co;2`.

[57] M. H. Carpenter, C. A. Kennedy, Fourth-order $2N$-storage Runge–Kutta schemes, Tech. rep., NASA Technical Memorandum 109112 (1994).

[58] C. Rackauckas, Q. Nie, DifferentialEquations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia, Journal of Open Research Software 5 (1) (2017) 15. `doi:10.5334/jors.151`.

[59] H. Ranocha, L. Dalcin, M. Parsani, D. I. Ketcheson, Optimized Runge–Kutta methods with automatic step size control for compressible computational fluid dynamics, Communications on Applied Mathematics and Computation 4 (4) (2022) 1191–1228. `doi:10.1007/s42967-021-00159-w`.

[60] H. Ranocha, A. R. Winters, H. G. Castro, L. Dalcin, M. Schlottke-Lakemper, G. J. Gassner, M. Parsani, On error-based step size control for discontinuous Galerkin methods for compressible fluid dynamics, Communications on Applied Mathematics and Computation 7 (1) (2025) 3–39. `doi:10.1007/s42967-023-00264-y`.

[61] S. G. Johnson, QuadGK.jl: Gauss–Kronrod integration in Julia, `https://github.com/JuliaMath/QuadGK.jl` (2013).

[62] D. I. Ketcheson, Relaxation Runge–Kutta methods: Conservation and stability for inner-product norms, SIAM Journal on Numerical Analysis 57 (6) (2019) 2850–2870. `doi:10.1137/19m1263662`.

[63] H. Ranocha, M. Sayyari, L. Dalcin, M. Parsani, D. I. Ketcheson, Relaxation Runge–Kutta methods: Fully discrete explicit entropy-stable schemes for the compressible Euler and Navier–Stokes equations, SIAM Journal on Scientific Computing 42 (2) (2020) A612–A638. `doi:10.1137/19M1263480`.

[64] L. F. Shampine, Conservation laws and the numerical solution of ODEs, Computers & Mathematics with Applications 12 (5–6) (1986) 1287–1296. `doi:10.1016/0898-1221(86)90253-1`.

[65] J. Chan, On discretely entropy conservative and entropy stable discontinuous Galerkin methods, Journal of Computational Physics 362 (2018) 346–374. `doi:10.1016/j.jcp.2018.02.033`.

[66] A. St-Cyr, C. Jablonowski, J. M. Dennis, H. M. Tufo, S. J. Thomas, A comparison of two shallow-water models with nonconforming adaptive grids, Monthly Weather Review 136 (6) (2008) 1898–1922. `doi:10.1175/2007mwr2108.1`.

[67] J. E. Hicken, D. C. Del Rey Fernández, D. W. Zingg, Multidimensional summation-by-parts operators: General theory and application to simplex elements, SIAM Journal on Scientific Computing 38 (4) (2016) A1935–A1958. `doi:10.1137/15m1038360`.

[68] T. Chen, C.-W. Shu, Review of entropy stable discontinuous Galerkin methods for systems of conservation laws on unstructured simplex meshes, CSIAM Transactions on Applied Mathematics 1 (1) (2020) 1–52. `doi:10.4208/csiam-am.2020-0003`.

[69] T. Montoya, D. W. Zingg, Efficient entropy-stable discontinuous spectral-element methods using tensor-product summation-by-parts operators on triangles and tetrahedra, Journal of Computational Physics 516 (2024) 113360. `doi:10.1016/j.jcp.2024.113360`.

[70] J. Keim, A. Schwarz, P. Kopper, M. Blind, C. Rohde, A. Beck, Entropy stable high-order discontinuous Galerkin spectral-element methods on curvilinear, hybrid meshes, Preprint, arXiv:2507.04334 [math.NA] (2025). `doi:10.48550/arXiv.2507.04334`.