A mass conservative TR-BDF2 semi-implicit semi-Lagrangian DG discretization of the shallow water equations on general structured meshes of quadrilaterals

Giovanni Tumolo*1

1Earth System Physics Section,
The Abdus Salam International Centre for Theoretical Physics, Trieste, Italy

*Email address for correspondence: gtumolo@ictp.it

Communicated by Luca Bonaventura
Received on June 12, 2015. Accepted on July 29, 2015.

Abstract

As an extension of a previous work considering a fully advective formulation on Cartesian meshes, a mass conservative discretization approach is presented here for the shallow water equations, based on discontinuous finite elements on general structured meshes of quadrilaterals. A semi-implicit time integration is performed by employing the TR-BDF2 scheme and is combined with the semi-Lagrangian technique for the momentum equation only. Indeed, in order to simplify the derivation of the discrete linear Helmholtz equation to be solved at each time-step, a non-conservative formulation of the momentum equation is employed. The Eulerian flux form is considered instead for the continuity equation in order to ensure mass conservation. Numerical results show that on distorted meshes and for relatively high polynomial degrees, the proposed numerical method fully conserves mass and presents a higher level of accuracy than a standard off-centered Crank Nicolson approach. This is achieved without any significant imprinting of the mesh distortion on the solution.

Keywords: discontinuous Galerkin methods, semi-implicit discretizations, semi-Lagrangian discretizations, shallow water equations, quadrilateral meshes.

AMS subject classification: 05A16, 65N38, 78M50.

1. Introduction

Two extensions of the technique proposed in [1] are presented for the solution of the shallow water equations, enabling mass conservation and the possibility to use generic structured meshes of quadrilaterals (in principle not aligned with the coordinate directions).
Conservation of mass was sacrificed in [1] in order to improve efficiency through the use of a non-conservative semi-Lagrangian advection scheme. However, mass conservation property is crucial not only for climate modelling (where shallow water equations represent a standard test bed for the horizontal discretization of the dynamics equations), e.g. [2], [3], [4], [5] but also for hydraulic applications, e.g. [6].

For the latter case, the ability of handling more complex geometries is also very important. In this sense, the extension to generical, only logically Cartesian meshes of quadrilaterals represents a first step towards a broader spectrum of more realistic low Mach number and low Froud number fluid dynamics applications.

Mass conservation is achieved here by the use of the Eulerian flux form for the continuity equation, which inherently leads to such a conservation property, independently on the choice of the (consistent) numerical fluxes. As a result, following e.g. [7–10], a semi-implicit time discretization, based here on the TR-BDF2 scheme, is combined with the semi-Lagrangian technique for the momentum equation only, hence departing from [1].

The flexibility of employing general structured meshes of distorted quadrilaterals is obtained through the introduction of a bilinear mapping to transform the generic quadrilateral element of the mesh in the physical space into the master element.

It has to be stressed that both of these extensions are only a first attempt to generalize the numerical formulation of [1], while a more ambitious extension to a fully conservative and semi-Lagrangian approach on the lines of [11], as well as an implementation on generical nonstructured meshes, is left for further studies.

In section 2 the governing equations under consideration are reviewed, in section 3 the semi-implicit semi-Lagrangian time discretization of such equations is introduced in spherical geometry, while details of the discontinuous Galerkin spatial discretization are presented in sections 4 and 5. Numerical results are presented in section 6, while in section 7 we draw some conclusions and outline the next development steps of the proposed approach.

2. Governing equations

The model problem considered in this paper is given by the two-dimensional shallow water equations on a rotating sphere. These equations are a standard test bed for numerical methods to be applied to the full equations of motion of atmospheric or oceanic circulation models, see e.g. [12].
G. Tumolo

The advective, vector form is used for the momentum equation and, departing from [1], the Eulerian flux form is employed for the continuity equation, in order to ensure mass conservation, along the lines of the approach pursued in [7–10]:

\[ \frac{\partial \eta}{\partial t} = -\nabla \cdot (h \mathbf{u}), \]

\[ \frac{D\mathbf{u}}{Dt} = -g\nabla \eta - f \mathbf{k} \times \mathbf{u}. \]

Here \( \eta \) represents the free surface elevation, \( h \) indicates the fluid depth, \( g \) is the gravity force per unit mass on the Earth’s surface, \( f \) the Coriolis parameter and \( \mathbf{k} \) the unit vector locally normal to the Earth’s surface. Assuming that \( x,y \) are orthogonal curvilinear coordinates on the sphere (or on a portion of it), we denote by \( m_x \) and \( m_y \) the components of the (diagonal) metric tensor. Furthermore, we set \( \mathbf{u} = (u,v)^T \), where \( u \) and \( v \) are the contravariant components of the velocity vector in the coordinate direction \( x \) and \( y \) respectively, multiplied by the corresponding metric tensor components. We also denote by \( D/Dt \) the Lagrangian derivative

\[ \frac{D}{Dt} = \frac{\partial}{\partial t} + u \frac{\partial}{m_x \partial x} + v \frac{\partial}{m_y \partial y}, \]

so that \( u = m_x D_x/Dt, v = m_y D_y/Dt \). In particular, standard spherical coordinates will be employed in this paper.

3. A TR-BDF2 time integration approach for the shallow water equations on the sphere

The TR-BDF2 method was already successfully applied in a previous work for the semi-implicit semi-Lagrangian (SISL) time discretization of the shallow water equations in advective form [1]. For a review of the TR-BDF2 method, which was first introduced in [13] and further analyzed in [14] in the context of the solution of Cauchy problems for ordinary differential system, see section 3 of [1].

For a more comprehensive description of the semi-Lagrangian technique see e.g. [15] and references therein.

The mixed Eulerian semi-implicit and SISL discretization of equations (1),(2) respectively, based on the TR-BDF2 scheme, is obtained by performing the two stages in TR-BDF2, after semi-Lagrangian reinterpretation of the intermediate values in equation (2): this is achieved through the introduction of proper discrete semi-Lagrangian evolution operators, [16,17], in
the following called $E$. If $G = G(x, t)$ denotes a generic function of space and time, then the expression $[E(t^n, \Delta t)G](x)$ refers to a numerical approximation of $G(x_D, t^n)$ where $x_D$ is the position at time $t^n$ of the fluid parcel reaching location $x$ at time $t^{n+1}$ and, according to standard terminology, it is called the departure point associated with the arrival point $x$. For a detailed review of the definition of such operators on vector fields on the sphere, see section 4 of [1]. The approximation of $x_D$ and the computation of $G(x_D, t^n)$ follows the same approach employed in [1]. The only difference with respect to [1] is given by the fact that the searching algorithm, to be employed to determine in which element a particular departure point lies, becomes nontrivial for distorted meshes. Many approaches are available [18–21], and here we follow the one proposed by [21].

Furthermore, in order to avoid the solution of a nonlinear system at each timestep, the dependency on $h$ in $\nabla \cdot (hu)$ is frozen in time to the latest computed value [7–10].

The TR stage of the SISL time discretization of the equations in vector form (1)-(2) is given by

$$
\eta^{n+2\gamma} + \gamma \Delta t \nabla \cdot (h^n u^{n+2\gamma}) = \eta^n - \gamma \Delta t \nabla \cdot (h^n u^n),
$$

$$
\eta^{n+2\gamma} + \gamma \Delta t \left( g \nabla \eta^{n+2\gamma} + f k \times u^{n+2\gamma} \right) = E(t^n, 2\gamma \Delta t) [u - \gamma \Delta t (g \nabla \eta + f k \times u)].
$$

The TR stage is then followed by the BDF2 stage:

$$
\eta^{n+1} + 2\gamma \Delta t \nabla \cdot (h^{n+2\gamma} u^{n+1}) = (1 - \gamma_3) \eta^n + \gamma_3 \eta^{n+2\gamma},
$$

$$
\eta^{n+1} + 2\Delta t \left( g \nabla \eta^{n+1} + f k \times u^{n+1} \right) = \left( 1 - \gamma_3 \right) E(t^n, \Delta t) u + \gamma_3 E(t^n + 2\gamma \Delta t, (1 - 2\gamma) \Delta t) u.
$$

For each of the two stages, the spatial discretization can be performed along the lines described in detail in sections 4, 5. After spatial discretization has been performed, the discrete degrees of freedom representing velocity unknowns can be replaced in the respective discrete free surface elevation equations, yielding in each case a linear system whose structure is entirely analogous to that obtained in [10]. The non-symmetric linear systems obtained from the TR-BDF2 stages are solved in our implementation by the GMRES method [22]. Once $\eta^{n+2\gamma}$ (or $\eta^{n+1}$) degrees of freedom have been computed by solving this linear system, $u^{n+2\gamma}$ (or $u^{n+1}$) degrees of freedom can then be recovered by back substituting into the (discrete) momentum equation.
4. Discontinuous finite element spaces

The discontinuous finite element discretization approach employed in this work is independent of the nature of the mesh and could also be implemented for fully unstructured and even non conforming meshes. For simplicity, however, in this paper we only consider implementations on structured, logically Cartesian meshes, consisting of quadrilateral elements. Therefore, we assume that equations (1)-(2) are to be solved with appropriate initial and boundary conditions on a domain $\Omega$ which is partitioned in $N$ non overlapping quadrilateral elements $K_I, I = 1, \ldots, N$, i.e. $\Omega = \bigcup_{I=1}^{N} K_I$. We denote the domain partition as $T_h = \{ K_I : I = 1, \ldots, N \}$, where the subscript $h$, not to be confused with the fluid depth, denotes the maximum diameter of the elements $K_I, I = 1, \ldots, N$. The vertices of the generic element $K_I$ are denoted by $(x^A_I, y^A_I), (x^B_I, y^B_I), (x^C_I, y^C_I), (x^D_I, y^D_I)$, while the four edges of the element $K_I$ are labelled as $e_{S,I}, e_{E,I}, e_{N,I}, e_{W,I}$ (see figure 1).

![Figure 1. Notation for denoting edges and vertices, reference element $\hat{K} = [-1,1]^2$.](image)

It is immediate that each $K_I$ is the image of the master element $\hat{K} = [-1,1] \times [-1,1]$ via the bilinear local map $F_I$, such that

$$x = F_{I,1}(\xi_1, \xi_2) = x_I^A \frac{1 - \xi_1}{2} \frac{1 - \xi_2}{2} + x_I^B \frac{1 + \xi_1}{2} \frac{1 - \xi_2}{2} +$$

$$+ x_I^C \frac{1 + \xi_1}{2} \frac{1 + \xi_2}{2} + x_I^D \frac{1 - \xi_1}{2} \frac{1 + \xi_2}{2},$$

$$y = F_{I,2}(\xi_1, \xi_2) = y_I^A \frac{1 - \xi_1}{2} \frac{1 - \xi_2}{2} + y_I^B \frac{1 + \xi_1}{2} \frac{1 - \xi_2}{2} +$$

$$+ y_I^C \frac{1 + \xi_1}{2} \frac{1 + \xi_2}{2} + y_I^D \frac{1 - \xi_1}{2} \frac{1 + \xi_2}{2},$$

where $(x,y) \in K_I$ and $(\xi_1, \xi_2) \in \hat{K}$. For a non-negative integer $p$, we denote by $Q_p$ the set of all polynomials of degree less or equal to $p$ in each
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

coordinate on $\hat{K}$. We will also define

$$Q_p(K_I) = \{ w : w = v \circ F_I^{-1}, \quad v \in Q_p \}.$$  

For each $I$, we denote by the non-negative integer $p_I$ the local polynomial degree on $K_I$ and we set $p = \{ p_I : I = 1, \ldots, N \}$. We then consider the finite element space:

$$V_{h,p} = \left\{ v \in L^2(\Omega) : v|_{K_I} \in Q_{p_I}(K_I) \quad I = 1, \ldots, N \right\}.  \quad (7)$$

We will denote by $E_h$ the set of all element boundary edges, which is naturally decomposed as $E_h = E_{h,int} \cup E_b$, where $E_{h,int}$ denotes the element boundary edges that belong to the interior of $\Omega$ and $E_b$ denotes the element boundary edges that belong to $\partial\Omega$. Functions in $V_{h,p}$ are in general discontinuous across element boundaries. Thus, it is convenient to define jump and average operators on each inter-element edge. For a given edge $e \in E_{h,int}$ there exist two elements $K, K' \in T_h$ such that $e = \partial K \cap \partial K'$, hence for each point $x \in e$ and $v_h \in V_{h,p}$, being $n_{\partial K,e}$ the restriction of $n_{\partial K}$ on the edge $e \subset \partial K$, we can define

$$\{(v_h)\} (x) = \frac{1}{2} (v_h|_K(x) + v_h|_{K'}(x)) \quad \quad (8)$$

$$[[v_h]] (x) = v_h|_K(x)n_{\partial K,e} + v_h|_{K'}(x)n_{\partial K',e}. \quad (9)$$

Notice that these definitions, following e.g. [23], are symmetric with respect to the two elements $K, K'$ sharing the boundary point. If instead $e \in E_{h,b}$, there is a unique $K \in T_h$ such that $e \in \partial K$ and hence we set

$$\{(v_h)\} (x) = v_h|_K(x) \quad \quad [[v_h]] (x) = v_h|_K(x)n_{\partial K,e}. \quad (10)$$

Notice that, actually, since the discretization we propose only employs centered fluxes, the jump operators are never used in this work. We now consider the issue of defining a basis for $V_{h,p}$. Thanks to the definition (7), this basis can be defined by first choosing a basis for $Q_p$ which induces a set of local bases on the spaces $Q_{p_I}(K_I)$. The basis of $V_{h,p}$ is then obtained by collecting the extensions of the latter functions to zero outside of their corresponding element. In principle, either Lagrangian or hierarchical, Legendre bases could be employed. We will work mostly with hierarchical bases, because they provide a natural environment for the implementation of a $p-$adaptation algorithm. Hierarchical bases can be defined starting
with the one dimensional case. More specifically, let \( \hat{K}^{1D} = [-1, 1] \) and consider, for \( \xi \in \hat{K}^{1D} \), the \( k \)-th Legendre polynomial, defined recursively by the three-term recurrence relation:

\[
L_{k+1}(\xi) = \frac{2k + 1}{k + 1} \xi L_k(\xi) - \frac{k}{k + 1} L_{k-1}(\xi), \quad k = 1, 2, \ldots
\]

\[
L_0(\xi) = 1, \quad L_1(\xi) = \xi.
\]

(11)

The Legendre polynomials form an orthogonal basis for polynomials on \( \hat{K}^{1D} \), since

\[
(L_p, L_q)_{L^2} = \int_{-1}^1 L_p(\xi)L_q(\xi)d\xi = \frac{2}{2p + 1}\delta_{pq}.
\]

Notice that this orthogonality property has important numerical implications for Galerkin type methods, since it not only implies that all the mass matrices are diagonal, but also gives in general improved conditioning of the resulting discretizations, see e.g. [24] for more details. Turning now to the two-dimensional case, on quadrilateral meshes the most natural and straightforward way to construct the two-dimensional basis is by taking products of the one-dimensional bases. For \((\xi_1, \xi_2) \in \hat{K}\), a hierarchical basis of \( \mathbb{Q}_p \) can then be constructed by taking tensor products of the Legendre polynomials introduced for the one-dimensional case, so that our basis functions have the the form \( L_k(\xi_1)L_l(\xi_2) \) for \( 0 \leq k, l \leq p \).

Finally, a central issue in finite element formulations for fluid problems is the choice of appropriate approximation spaces for the velocity and pressure (here replaced by the free surface elevation) variables. Indeed, an inconsistent choice of the two approximation spaces may result in a solution that is polluted by spurious modes, see for example [25] and the more recent and comprehensive analysis in [26].

In this paper, the focus is on the proposal of an efficient time discretization for DG methods, rather than on an investigation of the optimal approach for spatial discretization by discontinuous elements. As a consequence, no attempt is made here to analyze these issues at a theoretical level. In order to allow for numerical experiments with different possible choices, in the derivation presented in the following we consider potentially different degrees \( p^u, p^\eta \) for the velocity and free surface elevation variables, respectively. The notation \( \{\psi_{I,m}^{(p^u+1)^2}\} \) and \( \{\phi_{I,l}^{(p^\eta+1)^2}\} \) will be used to denote the corresponding basis functions of \( \mathbb{Q}_{p^u}(K_I) \) and \( \mathbb{Q}_{p^\eta}(K_I) \) respectively. The corresponding global finite element spaces will be denoted by \( V_{h,p^u}, V_{h,p^\eta} \), respectively.
5. Spatial discretization

The problem of a DG discretization of the SISL semidiscrete equations (1)-(2) is now addressed on a general conforming mesh composed of quadrilaterals, whose edges are not necessarily aligned with the coordinate axes. We discretize in space by taking the $L^2$ projection against the test functions on a given element $K_I$. In particular, following [27] or the direct Lagrange Galerkin method proposed in [28], we use as test functions the basis functions $\{\varphi_{I,l}\}_{l=1}^{(p_\eta^\delta+1)^2}$, $\{\psi_{I,m}\}_{m=1}^{(p_u^\delta+1)^2}$.

Notice that in the TR substep of the SISLDG formulation of the shallow water equations, we have to consider the numerical evolution operator $E(t^n,2\gamma\Delta t)$ applied to the gradient of the free surface elevation $\eta_h$ and to the divergence of the velocity $u_h$, where both $\eta_h$ and $u_h$ are piecewise polynomial functions with no global continuity constraints. To compute these terms, we follow the standard local discontinuous Galerkin method and define first the discrete gradient operator $D_h\eta_h$ and the discrete divergence operator $\nabla_h \cdot u_h$ as follows. Given $\eta_h \in V_{h,\eta}$ the discrete free surface elevation gradient $D_h\eta_h = (D^x_h\eta_h, D^y_h\eta_h)^T$ is the unique element of $[V_{h,\eta}]^2$ such that, for $I = 1, \cdots, N$:

$$\int_{K_I} D_h\eta_h \cdot \Psi \, dx = - \int_{K_I} \eta_h \nabla \cdot \Psi \, dx$$

$$+ \int_{\partial K_I} \hat{\eta}_h \Psi \cdot n \, ds \quad \forall \Psi \in [V_{h,\eta}]^2,$$

where $n$ is the unit normal vector pointing outward from $\partial K_I$. Notice that elements of $V_{h,\eta}$ are not uniquely defined on the elements of $\mathcal{E}_{h,int}$. Therefore, the resulting boundary terms in (12) are indeed replaced by appropriate numerical fluxes $\hat{\eta}_h$, which are single valued functions of both interface states and thus introduce the necessary coupling between the unknowns of neighboring elements. $\hat{\eta}_h$ will be defined in the following formula (17).

Notice that, by integration by parts in (12), we also have for $I = 1, \cdots, N$:

$$\int_{K_I} D_h\eta_h \cdot \Psi \, dx = \int_{K_I} \nabla \eta_h \cdot \Psi \, dx +$$

$$\int_{\partial K_I} (\hat{\eta}_h - \eta_h) \Psi \cdot n \, ds \quad \forall \Psi \in [V_{h,\eta}]^2.$$

Hence, as observed also in [29], it is possible to recast the discrete gradient operator as the sum of two contributions, the first one taking into account the elementwise gradient of $\eta_h$, while the second one taking into account its jumps across the element interfaces. As a consequence, the components of the pressure gradient term are defined as:
\begin{align}
\int_{K_i} D_h^x \eta_h \psi_{I,m} \, dx &= \\
&- \int_{K_i} \eta_h \frac{\partial}{\partial x} \psi_{I,m} \, dx + \int_{\partial K_i} \hat{\eta}_h \psi_{I,m} i \cdot n \, ds = \\
&\int_{K_i} \frac{\partial \eta_h}{\partial x} \psi_{I,m} \, dx + \int_{\partial K_i} (\hat{\eta}_h - \eta_h) \psi_{I,m} i \cdot n \, ds,
\end{align}

for \( m = 1, \ldots, (p_u^I + 1)^2 \) and

\begin{align}
\int_{K_i} D_h^y \eta_h \psi_{I,m} \, dx &= \\
&- \int_{K_i} \eta_h \frac{\partial}{\partial y} \psi_{I,m} \, dx + \int_{\partial K_i} \hat{\eta}_h \psi_{I,m} k \cdot n \, ds = \\
&\int_{K_i} \frac{\partial \eta_h}{\partial y} \psi_{I,m} \, dx + \int_{\partial K_i} (\hat{\eta}_h - \eta_h) \psi_{I,m} j \cdot n \, ds,
\end{align}

for \( m = 1, \ldots, (p_v^I + 1)^2 \), where \( i, j \) are the unit vectors along the coordinate axes. The coefficients of the discrete partial derivatives of \( \eta_h \) with respect to the chosen local polynomial basis are computed via the solution of (14) and (15). If orthogonal hierarchical bases \( \{ \psi_{I,m} \}_{m=1}^{(p_u^I+1)^2} \) of \( Q_{p_u^I}(K_i) \) are used, this only requires the inversion of diagonal local mass matrices. Notice that \( D_h \eta_h \) could in principle also be defined as an element of \( [V_{h,p^u}]^2 \). Numerical experiments have shown that this does not lead to significant changes in the results, so that we adopted definition (12) for convenience.

Given \( u_h \in [V_{h,p^u}]^2 \), and a weight function \( h_h \in V_{h,p^w} \), the discrete divergence \( \nabla_h \cdot u_h \) is the unique element of \( V_{h,p^w} \) such that, for \( I = 1, \ldots, N \):

\begin{align}
\int_{K_i} \Phi \nabla_h \cdot (h_h u_h) \, dx &= \\
&- \int_{K_i} h_h u_h \cdot \nabla \Phi \, dx + \int_{\partial K_i} \Phi \hat{h}_h \hat{u}_h \cdot n \, ds \quad \forall \Phi \in V_{h,p^w},
\end{align}

where \( n \) is the unit normal vector pointing outward from \( \partial K_i \). Notice that elements of \( V_{h,p^w}, V_{h,p^\eta} \), are not uniquely defined on the elements of \( E_{h,int} \). Therefore, the resulting boundary terms in (16) are indeed replaced by appropriate numerical fluxes \( \hat{u}_a, \hat{h}_h \), which are single valued functions of both interface states and thus introduce the necessary coupling between the unknowns of neighboring elements. \( \hat{u}_h, \hat{\eta}_h, \hat{\eta}_h \), will be defined in the following formula (17). The coefficients of the discrete divergence \( \nabla_h \cdot u_h \) with respect to the chosen local polynomial basis are computed via the solution
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

of (16) where we set \( h_h = 1 \). If orthogonal hierarchical bases \( \{ \varphi_{I,l} \}_{l=1}^{(p_h+1)^2} \) of \( Q_{p_h}(K_I) \) are used, this only requires the inversion of diagonal local mass matrices. Notice that \( \nabla_h \cdot (h_h u_h) \) could in principle also be defined as an element of \( [V_{h,p^v}]^2 \). Numerical experiments in [1,10] have shown that this does not lead to significant changes in the results, so that we adopted definition (16) for convenience. In general it is customary in the framework of DG methods for hyperbolic problems to employ upwind fluxes like the Lax-Friedrichs or Rusanov flux, see e.g. [23,30,31]. In this work, however, in order to simplify the structure of the linear system resulting from the semi-implicit discretization, we will use centered fluxes as in [29]. Therefore, using the notation in (9), for each dependent variable \( a_h \in V_{h,p^v} \) or \( a \in V_{h,p^v} \) we define:

\[
\hat{a}_h|_e = \{ a_h \}, \quad \forall e \in \mathcal{E}_{h,int}.
\]

We then obtain for each element \( K_I, I = 1, \ldots, N \):

\[
\int_{K_I} \varphi_{I,l} u_h^{n+2\gamma} \ dx + \gamma \Delta t \int_{K_I} \varphi_{I,l} \nabla_h \cdot \left( h_h^n u_h^{n+2\gamma} \right) \ dx = \int_{K_I} \varphi_{I,l} \left[ \eta_h^n - \gamma \Delta t \nabla_h \cdot (h_h^n u_h^n) \right] \ dx,
\]

\[
\int_{K_I} \psi_{I,m} u_h^{n+2\gamma} \ dx + \gamma \Delta t \int_{K_I} \psi_{I,m} \left[ \mathcal{D}^h_{u_h} \right]^{n+2\gamma} \ dx - \Delta t \int_{K_I} \psi_{I,m} f u_h^{n+2\gamma} \ dx
\]

\[
= \int_{K_I} \psi_{I,m} i \cdot \left\{ E(t^n, 2\gamma \Delta t) \left[ u_h - \gamma \Delta t (g \mathcal{D} \eta_h + f k \times u_h) \right] \right\} \ dx,
\]

\[
\int_{K_I} \psi_{I,m} v_h^{n+2\gamma} \ dx + \gamma \Delta t \int_{K_I} \psi_{I,m} \left[ \mathcal{D}^h_{v_h} \right]^{n+2\gamma} \ dx + \gamma \Delta t \int_{K_I} \psi_{I,m} f v_h^{n+2\gamma} \ dx
\]

\[
= \int_{K_I} \psi_{I,m} j \cdot \left\{ E(t^n, 2\gamma \Delta t) \left[ u_h - \gamma \Delta t (g \mathcal{D} \eta_h + f k \times u_h) \right] \right\} \ dx,
\]

After use of definitions (16) and (14), (15) in the continuity and momentum equations respectively, we have:
\[
\int_{K_I} \varphi_{I,l} h^n_I d\mathbf{x} - \gamma \Delta t \int_{K_I} h^{n+2\gamma}_I u^{n+2\gamma}_I \cdot \nabla \varphi_{I,l} d\mathbf{x} + \gamma \Delta t \int_{\partial K_I} \varphi_{I,l} \hat{h}_I^n \hat{u}_h^{n+2\gamma} \cdot \mathbf{n} d\mathbf{s} = \int_{K_I} \varphi_{I,l} \eta^n_I - \gamma \Delta t \nabla \hat{h}_I \cdot (h^n_I u^n_I) d\mathbf{x},
\]

(18)

\[
\int_{K_I} \psi_{I,m} h^n_I d\mathbf{x} - \gamma \Delta t \int_{K_I} \psi_{I,m} f h^{n+2\gamma}_I d\mathbf{x} + \gamma \Delta t g \int_{K_I} \psi_{I,m} d\mathbf{x} + \gamma \Delta t g \int_{\partial K_I} \left( \eta^{n+2\gamma}_I - \eta^{n+2\gamma}_I \right) \psi_{I,m} \mathbf{i} \cdot \mathbf{n} d\mathbf{s} = \int_{K_I} \psi_{I,m} \mathbf{i} \cdot \left\{ E(t^n, 2\gamma \Delta t) \left[ u_h - \gamma \Delta t (gD_h \eta_h + f k \times u_h) \right] \right\} d\mathbf{x},
\]

(19)

where \( \mathbf{i}, \mathbf{j} \) are the unit vectors along the coordinate axes, while \( \mathbf{n} \) is the unit normal vector outward from \( \partial K_I \). Expanding the dependent variables in terms of the basis functions, one has

\[
\eta^n_{I,I} = \left( p_{\eta}^I + 1 \right)^2 \sum_{r=1}^{P} \varphi_{I,r} \eta_{I,r}^n,
\]

(21)

\[
u^n_{I,I} = \left( p_{\nu}^I + 1 \right)^2 \sum_{r=1}^{P} \psi_{I,r} u_{I,r}^n, \quad v^n_{I,I} = \left( p_{\nu}^I + 1 \right)^2 \sum_{r=1}^{P} \psi_{I,r} v_{I,r}^n.
\]

(22)

The issue of the appropriate and solution-dependent choice of the values \( p_{\eta}^I \) and \( p_{\nu}^I \) is addressed in [10], where a simple but effective \( p \)-adaptivity criterion has been described.
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

Substituting now expressions (21)-(22) into the numerical fluxes definitions (17) and then also into equations (18)-(20), one obtains the full space and time (for the TR substep) discretization of system (1), (2) that can be conveniently presented employing a vector notation. For a generic variable \( \zeta \), we denote by \( \zeta^n_I = (\zeta^n_{I,l}) \), \( l = 1, \ldots, (p^\zeta_I + 1)^2 \) the vectors collecting all the discrete degrees of freedom associated to element \( K_I \). Furthermore, we introduce the local mass matrices associated to \( \eta_h, u_h, v_h \), denoted by \( M^n_\eta I, M^u_I \) respectively, whose entries are given by

\[
(M^n_\eta)_l,r = \int_{K_I} \varphi_{I,l} \varphi_{I,r} \, dx,
\]

\[
(M^u_I)_{m,r} = \int_{K_I} \psi_{I,m} \psi_{I,r} \, dx,
\]

where one has \( l, r = 1, \ldots, (p^\eta_I + 1)^2 \) for \( \eta_h \), and \( m, r = 1, \ldots, (p^u_I + 1)^2 \) for \( u, v \). Notice that, if hierarchical basis functions are used, \( M^n_\eta I, M^u_I \) matrices are diagonal. If we consider the partition \( \partial K_I = e_{S,I} \cup e_{E,I} \cup e_{N,I} \cup e_{W,I}, \quad I = 1, \ldots, N \), according to the convention of figure 1, we also define a number of auxiliary matrices. Specifically, if \( n_x = n \cdot i, n_y = n \cdot j \), for the continuity equation we set:

\[
(L^n_{\eta u})_{l,r} = -\int_{K_I} \frac{\partial \varphi_{I,l}}{\partial x} h^n_{h} \psi_{I,r} \, dx
\]

\[
+ \frac{1}{2} \int_{e_{S,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,r} n_x \, ds + \frac{1}{2} \int_{e_{E,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,r} n_x \, ds
\]

\[
+ \frac{1}{2} \int_{e_{N,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,r} n_x \, ds + \frac{1}{2} \int_{e_{W,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,r} n_x \, ds,
\]

\[
(S^n_{\eta u})_{l,r} = \frac{1}{2} \int_{e_{S,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,S,r} n_x \, ds,
\]

\[
(E^n_{\eta u})_{l,r} = \frac{1}{2} \int_{e_{E,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,E,r} n_x \, ds,
\]

\[
(N^n_{\eta u})_{l,r} = \frac{1}{2} \int_{e_{N,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,N,r} n_x \, ds,
\]

\[
(W^n_{\eta u})_{l,r} = \frac{1}{2} \int_{e_{W,I}} \varphi_{I,l} \hat{h}^n_{h} \psi_{I,W,r} n_x \, ds,
\]
\[
(L^{\eta u}_{I})_{I,r} = - \int_{K_I} \left( \frac{\partial}{\partial y} \varphi_{I,l} \right) h^n_k \psi_{I,r} \, dx \\
+ \frac{1}{2} \int_{e_{S, I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,r} n_z \, ds \\
+ \frac{1}{2} \int_{e_{E, I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,r} n_z \, ds \\
+ \frac{1}{2} \int_{e_{N, I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,r} n_z \, ds \\
+ \frac{1}{2} \int_{e_{W, I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,r} n_z \, ds,
\]

\( (S^{\eta u}_{I})_{I,r} = \frac{1}{2} \int_{e_{S,I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,S,r} n_z \, ds, \)

\( (E^{\eta u}_{I})_{I,r} = \frac{1}{2} \int_{e_{E,I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,E,r} n_z \, ds, \)

\( (N^{\eta u}_{I})_{I,r} = \frac{1}{2} \int_{e_{N,I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,N,r} n_z \, ds, \)

\( (W^{\eta u}_{I})_{I,r} = \frac{1}{2} \int_{e_{W,I}} \varphi_{I,l} \hat{h}^n_k \psi_{I,W,r} n_z \, ds, \)

\[
(P^{TR}_{I})_{I} = M_I^{\eta} \eta_I^u + \gamma \Delta t \left( L^{\eta u}_{I} u_I^u + S^{\eta u}_{I} u_I^{m u} + E^{\eta u}_{I} u_I^{E} + N^{\eta u}_{I} u_I^{N} + W^{\eta u}_{I} u_I^{W} + L^{\eta v}_{I} v_I^u + S^{\eta v}_{I} v_I^{m u} + E^{\eta v}_{I} v_I^{E} + N^{\eta v}_{I} v_I^{N} + W^{\eta v}_{I} v_I^{W} \right) \]

(26)

for the momentum equation along \( x \):

\[
(L^{\eta u}_{I})_{m,r} = g \left( \int_{K_I} \psi_{I,m} \frac{\partial \varphi_{I,r}}{\partial x} \, dx \right.
\\
- \frac{1}{2} \int_{e_{S,I}} \psi_{I,m} \varphi_{I,r} n_x \, ds \left. - \frac{1}{2} \int_{e_{E,I}} \psi_{I,m} \varphi_{I,r} n_x \, ds \right)
\\
- \frac{1}{2} \int_{e_{N,I}} \psi_{I,m} \varphi_{I,r} n_x \, ds - \frac{1}{2} \int_{e_{W,I}} \psi_{I,m} \varphi_{I,r} n_x \, ds),
\]

(28)
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

\[
\left( S_{T}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{S,I}} \psi_{I,m} \varphi_{I_{S},r} n_{x} \, ds,
\]

\[
\left( E_{T}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{E,I}} \psi_{I,m} \varphi_{I_{E},r} n_{x} \, ds,
\]

\[
\left( N_{T}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{N,I}} \psi_{I,m} \varphi_{I_{N},r} n_{x} \, ds,
\]

\[
\left( W_{T}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{W,I}} \psi_{I,m} \varphi_{I_{W},r} n_{x} \, ds,
\]

\[
(U^{TR})_{m} = \int_{K_{I}} \psi_{I,m} \mathbf{i} \cdot \left\{ E(t^n, 2\gamma \Delta t) \left[ \mathbf{u}_h - \gamma \Delta t \left( g \mathbf{D}_h \eta_h + f \mathbf{k} \times \mathbf{u}_h \right) \right] \right\} \, d\mathbf{x};
\]

and for the momentum equation along \( y \):

\[
\left( L_{I}^{yn} \right)_{m,r} = \frac{g}{2} \int_{K_{I}} \psi_{I,m} \frac{\partial \varphi_{I_{r}}}{\partial y} \, d\mathbf{x}
\]

\[
- \frac{1}{2} \int_{e_{S,I}} \psi_{I,m} \varphi_{I_{S},r} n_{y} \, ds
- \frac{1}{2} \int_{e_{E,I}} \psi_{I,m} \varphi_{I_{E},r} n_{y} \, ds
- \frac{1}{2} \int_{K_{I}} \psi_{I,m} \varphi_{I_{r},y} \, ds,
\]

\[
(30)
\]

\[
\left( S_{I}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{S,I}} \psi_{I,m} \varphi_{I_{S},r} n_{y} \, ds,
\]

\[
\left( E_{I}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{E,I}} \psi_{I,m} \varphi_{I_{E},r} n_{y} \, ds,
\]

\[
\left( N_{I}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{N,I}} \psi_{I,m} \varphi_{I_{N},r} n_{y} \, ds,
\]

\[
\left( W_{I}^{yn} \right)_{m,r} = \frac{g}{2} \int_{e_{W,I}} \psi_{I,m} \varphi_{I_{W},r} n_{y} \, ds,
\]

\[
(V^{TR})_{m} = \int_{K_{I}} \psi_{I,m} \mathbf{j} \cdot \left\{ E(t^n, 2\gamma \Delta t) \left[ \mathbf{u}_h - \gamma \Delta t \left( g \mathbf{D}_h \eta_h + f \mathbf{k} \times \mathbf{u}_h \right) \right] \right\} \, d\mathbf{x};
\]

Notice that all the local matrices defined in (25), (27) and (30) are in general rectangular, of dimensions \((p_{\eta}^J+1)^2 \times (p_{u}^J+1)^2\), where \( J \in \{ I, IS, IE, IN, IW \} \) for matrices defined in (25), while of dimension \((p_{\eta}^J+1)^2 \times (p_{u}^J+1)^2\), where \( J \in \{ I, IS, IE, IN, IW \} \) for the matrices defined in (27) and (30).
After the definitions (25)-(31), the DG discretization of equations (3), (4) reads in matrix form:

\[
M^\eta_{Ij} \eta^{n+2\gamma} + 
\gamma \Delta t \left( L^u_{Ij} u^{n+2\gamma} + S^u_{Ij} u^{n+2\gamma} + E^u_{Ij} u^{n+2\gamma} + N^u_{Ij} u^{n+2\gamma} + W^u_{Ij} u^{n+2\gamma} +
L^v_{Ij} v^{n+2\gamma} + S^v_{Ij} v^{n+2\gamma} + E^v_{Ij} v^{n+2\gamma} + N^v_{Ij} v^{n+2\gamma} + W^v_{Ij} v^{n+2\gamma} \right)
= \mathbf{P}_I^R
\]

\[
M^\eta_{Ij} u^{n+2\gamma} + \gamma \Delta t \left( L^u_{Ij} u^{n+2\gamma} + S^u_{Ij} u^{n+2\gamma} + E^u_{Ij} u^{n+2\gamma} + N^u_{Ij} u^{n+2\gamma} + W^u_{Ij} u^{n+2\gamma} - M^f_{Ij} u^{n+2\gamma} \right) = \mathbf{U}_I^R,
\]

\[
N^\eta_{Ij} v^{n+2\gamma} + W^\eta_{Ij} v^{n+2\gamma} - M^f_{Ij} v^{n+2\gamma} = \mathbf{V}_I^R,
\]

As customary in semi-implicit methods, see e.g. [7,32,33], \( u^{n+2\gamma} \) and \( v^{n+2\gamma} \) are expressed in terms of \( \eta^{n+2\gamma} \) and the resulting expressions substituted into the discrete continuity equation (26), in order to obtain a discrete Helmholtz equation in the \( \eta^{n+2\gamma} \) unknown only. If the elements surrounding \( K_I \) are labelled as S, E, N, W in counterclockwise order, see figure 1, this equation can be rewritten as follows:

\[
K^W_{Ij} \eta^{n+2\gamma} + K^S_{Ij} \eta^{n+2\gamma} + K^W_{Ij} \eta^{n+2\gamma} +
K^N_{Ij} \eta^{n+2\gamma} + K^S_{Ij} \eta^{n+2\gamma} + K^N_{Ij} \eta^{n+2\gamma} +
K^E_{Ij} \eta^{n+2\gamma} + K^N_{Ij} \eta^{n+2\gamma} + K^E_{Ij} \eta^{n+2\gamma} +
K^N_{Ij} \eta^{n+2\gamma} + K^E_{Ij} \eta^{n+2\gamma} + K^N_{Ij} \eta^{n+2\gamma} +
K^E_{Ij} \eta^{n+2\gamma} = \mathbf{R}_I.
\]
Here, we have set:

\[ K_{WW}^I = c_2^{TR} \left[ W_{I}^{nu}(M_{Is}^u)^{-1} W_{I}^{nu} + W_{I}^{nu}(M_{Iw}^u)^{-1} W_{I}^{nu} \right], \]

\[ K_{SW}^I = c_2^{TR} \left[ S_{I}^{nu}(M_{Is}^u)^{-1} W_{I}^{nu} + S_{I}^{nu}(M_{Iw}^u)^{-1} S_{I}^{nu} \right], \]

\[ K_{W}^I = c_2^{TR} \left[ L_{I}^{nu}(M_{I}^u)^{-1} W_{I}^{nu} + L_{I}^{nu}(M_{Iw}^u)^{-1} L_{I}^{nu} \right], \]

\[ K_{NW}^I = c_2^{TR} \left[ N_{I}^{nu}(M_{Is}^u)^{-1} W_{I}^{nu} + N_{I}^{nu}(M_{Iw}^u)^{-1} N_{I}^{nu} \right], \]

\[ K_{SS}^I = c_2^{TR} \left[ S_{I}^{nu}(M_{Is}^u)^{-1} S_{I}^{nu} + S_{I}^{nu}(M_{Iw}^u)^{-1} S_{I}^{nu} \right], \]

\[ K_{S}^I = c_2^{TR} \left[ L_{I}^{nu}(M_{I}^u)^{-1} S_{I}^{nu} + L_{I}^{nu}(M_{Iw}^u)^{-1} L_{I}^{nu} \right], \]

\[ K_{I} = M_{I}^u + \]

\[ E_{I}^{nu}(M_{I}^u)^{-1} W_{I}^{nu} + N_{I}^{nu}(M_{Is}^u)^{-1} S_{I}^{nu} + \]

\[ W_{I}^{nu}(M_{Iw}^u)^{-1} L_{I}^{nu} + L_{I}^{nu}(M_{I}^u)^{-1} L_{I}^{nu} \]

\[ S_{I}^{nu}(M_{Is}^u)^{-1} N_{I}^{nu} + E_{I}^{nu}(M_{I}^u)^{-1} W_{I}^{nu} + \]

\[ N_{I}^{nu}(M_{Iw}^u)^{-1} S_{I}^{nu} + W_{I}^{nu}(M_{Iw}^u)^{-1} L_{I}^{nu} \]

\[ K_{I}^N = c_2^{TR} \left[ L_{I}^{nu}(M_{I}^u)^{-1} N_{I}^{nu} + N_{I}^{nu}(M_{Iw}^u)^{-1} N_{I}^{nu} \right], \]

\[ K_{NN}^I = c_2^{TR} \left[ N_{I}^{nu}(M_{Is}^u)^{-1} N_{I}^{nu} + N_{I}^{nu}(M_{Iw}^u)^{-1} N_{I}^{nu} \right], \]

\[ K_{SE}^I = c_2^{TR} \left[ S_{I}^{nu}(M_{Is}^u)^{-1} E_{I}^{nu} + E_{I}^{nu}(M_{I}^u)^{-1} S_{I}^{nu} \right], \]

\[ S_{I}^{nu}(M_{Iw}^u)^{-1} E_{I}^{nu} + E_{I}^{nu}(M_{Iw}^u)^{-1} S_{I}^{nu} \]
G. Tumolo

\[
K_I^E = c_2^{TR} \left[ L_1^{\eta u}(M_I^\eta)^{-1} E_I^{\eta u} + E_I^{\eta u}(M_I^{1\eta})^{-1} L_1^{\eta u} + L_1^{\eta u}(M_I^\eta)^{-1} E_I^{\eta u} + E_I^{\eta u}(M_I^{1\eta})^{-1} L_1^{\eta u} \right],
\]

\[
K_I^{NE} = c_2^{TR} \left[ N_I^{\eta u}(M_I^{1\eta})^{-1} E_I^{\eta u} + E_I^{\eta u}(M_I^{1\eta})^{-1} N_I^{\eta u} + N_I^{\eta u}(M_I^{1\eta})^{-1} E_I^{\eta u} + E_I^{\eta u}(M_I^{1\eta})^{-1} N_I^{\eta u} \right],
\]

\[
K_I^{EE} = c_2^{TR} \left[ E_I^{\eta u}(M_I^{1\eta})^{-1} E_I^{\eta e} + E_I^{\eta u}(M_I^{1\eta})^{-1} E_I^{\eta e} \right],
\]

and defined the right hand side as:

\[
(38) \quad \mathcal{R}_I = \mathcal{P}^{TR}_I + c_1^{TR} \left[ L_1^{\eta u}(M_I^\eta)^{-1} \mathbf{u}_I^{TR} + S_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{u}_I^{TR} + E_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{u}_I^{TR} + W_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{v}_I^{TR} + L_1^{\eta u}(M_I^\eta)^{-1} \mathbf{v}_I^{TR} + N_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{v}_I^{TR} + S_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{v}_I^{TR} + E_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{v}_I^{TR} + N_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{v}_I^{TR} + W_I^{\eta u}(M_I^{1\eta})^{-1} \mathbf{v}_I^{TR} \right],
\]

while constants $c_1^{TR}, c_2^{TR}$ are defined by $c_1^{TR} = -\gamma \Delta t$, $c_2^{TR} = -(\gamma \Delta t)^2$. All the integrals appearing in the elemental equations are evaluated by means of Gaussian numerical quadrature formulae with a number of quadrature nodes consistent with the local polynomial degree being used. In particular, notice that integrals of terms in the image of the evolution operator $E$, i.e. functions evaluated at the departure points of the trajectories arriving at the quadrature nodes, cannot be computed exactly (see e.g. [28,34]), since such functions are not polynomials. Therefore, a sufficiently accurate approximation of these integrals is needed, which may entail the need to employ numerical quadrature formulae with more nodes than the minimal requirement implied by the local polynomial degree.

Furthermore, notice that we have written here the discrete equations in the general case case of a mesh of distorted quadrilateral elements with edges not aligned with the coordinate axes. The resulting coupling among each element and its corner neighbours leads to a computational stencil identical to that of the two dimensional case with rotation considered in [10] and also employed for the shallow water equations on the sphere.

The linear system (34) can be solved using e.g. the GMRES method [22]. A classical stopping criterion based on a relative error tolerance of $10^{-10}$ was employed (see e.g. [35]). For the GMRES solver, so far, only a block diagonal preconditioning was employed. Once $\mathbf{u}_I^{n+2\gamma}$ is known, $\mathbf{u}_I^{n+2\gamma}$ and $\mathbf{v}_I^{n+2\gamma}$ can be recovered by substituting $\mathbf{u}_I^{n+2\gamma}$ back into equations (32) and
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

(33), respectively. Once $\eta^{n+2\gamma}_I, u^{n+2\gamma}_I, v^{n+2\gamma}_I$ have been computed from the fully discrete TR substep, then $\eta^{n+1}_I, u^{n+1}_I, v^{n+1}_I$ can be obtained through the DG discretization of the BDF2 substep semidiscrete equations (5), (6), which follows the same lines of the DG discretization of the TR substep semidiscrete equations (3), (4), with the exception of a different definition of the right hand sides:

$$P^{BDF}_I = M^n_I (1 - \gamma_3) \eta^n_I + \gamma_3 \eta^{n+2\gamma}_I$$

(39)

$$U^{BDF}_I (m) = (1 - \gamma_3) \int_{K_I} \varphi_{I,I} i \cdot [E (t^n, \Delta t) u_h] \, dx + \gamma_3 \int_{K_I} \varphi_{I,I} i \cdot [E (t^n + 2\gamma \Delta t, (1 - 2\gamma) \Delta t) u_h] \, dx,$$

$$V^{BDF}_I (m) = (1 - \gamma_3) \int_{K_I} \varphi_{I,I} j \cdot [E (t^n, \Delta t) u_h] \, dx + \gamma_3 \int_{K_I} \varphi_{I,I} j \cdot [E (t^n + 2\gamma \Delta t, (1 - 2\gamma) \Delta t) u_h] \, dx,$$

and the need of replacing constants $c_1^{TR}, c_2^{TR}$ with $c_1^{BDF}, c_2^{BDF}$ in (35)-(38), where $c_1^{BDF}, c_2^{BDF}$ are defined by $c_1^{BDF} = -\gamma_2 \Delta t, c_2^{BDF} = -((\gamma_2 \Delta t)^2$. To improve accuracy, one can also recompute (25) by using $h^{n+2\gamma}_h$ in place of $h^n_h$.

6. Numerical results

In the preliminary tests presented here, a distorted mesh in the longitude-latitude domain is considered as in figure 2 for the case of $20 \times 10$ elements. For higher resolutions, the corresponding mesh can be refined from this reference mesh. Such a choice for the mesh is an adaptation to the sphere of that first proposed on the beta-plane by [36], and is motivated by the intention of mimicking the typical distortions in a mesh of the sphere that occurs, for example, in a cubed sphere. In such a case the mesh should include grid-lines which change direction by at least $\pi/3$, as in figure 2.

Mass conservation has been checked by computing at each time level $t^n$ the following measure of deviation from the initial mass value:

$$E^{n}_{mass} = \frac{\int_{\Omega} h^n_h \, dx - \int_{\Omega} h^0_h \, dx}{\int_{\Omega} h^0_h \, dx}.$$
Values just above the order of machine precision were obtained in all such tests.

Numerical experiments have been run with either equal degree element pairs $Q_p - Q_p$ for the velocity and free surface elevation variables or with polynomials of one degree higher for the velocity variables, without any remarkable difference in the results. To the extent of the tests that we have carried out, the differences between the solutions computed with equal degrees and those computed employing one degree higher polynomials for the velocity variables (as it would be customary for stability reasons in the incompressible framework) are coherent with what one would expect from the different spatial truncation errors of the two approximations. Further investigation will be needed to assess the optimal choice for geophysical scale flows.

6.1. Steady-state geostrophic flow

We first consider the test case 2 of [12], where the solution is a steady state flow with velocity field corresponding to a zonal solid body rotation and $h$ field obtained from the velocity ones through geostrophic balance. All the parameter values are taken as in [12]. The flow orientation parameter has been chosen here as $\alpha = \pi/2 - 0.05$, making the test more challenging on a regular (non-distorted) longitude-latitude mesh.
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

Considering a distorted meshes as in figure 2, a convergence assessment is performed for \( p^h = p^u = 3 \), increasing the number of elements and correspondingly decreasing the value of the time step in order to compare results obtained at the same Courant number. In this case, nevertheless, the maximum Courant numbers vary because of the mesh inhomogeneity, so that \( 2 < C_{cel} < 18, 0.5 < C_{cel} < 4 \). The results are reported in tables 1, 2 and 3 for \( \eta, u \) and \( v \), respectively. The empirical convergence order \( q^\text{emp} \) based on the \( L_2 \) norm errors has also been estimated, showing that in this stationary test convergence rates above the order of the time discretization can be achieved, even in presence of a nonnegligible mesh distortion.

<table>
<thead>
<tr>
<th>( N_x \times N_y )</th>
<th>( \Delta t )</th>
<th>( l_1(\eta) )</th>
<th>( l_2(\eta) )</th>
<th>( l_\infty(\eta) )</th>
<th>( q^\text{emp} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 ( \times ) 10</td>
<td>1800</td>
<td>5.59 ( \times ) 10(^{-3} )</td>
<td>8.39 ( \times ) 10(^{-3} )</td>
<td>1.20 ( \times ) 10(^{-3} )</td>
<td>-</td>
</tr>
<tr>
<td>40 ( \times ) 20</td>
<td>900</td>
<td>5.84 ( \times ) 10(^{-6} )</td>
<td>8.66 ( \times ) 10(^{-6} )</td>
<td>1.67 ( \times ) 10(^{-4} )</td>
<td>3.3</td>
</tr>
<tr>
<td>80 ( \times ) 40</td>
<td>450</td>
<td>7.50 ( \times ) 10(^{-7} )</td>
<td>1.02 ( \times ) 10(^{-6} )</td>
<td>6.88 ( \times ) 10(^{-6} )</td>
<td>3.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( N_x \times N_y )</th>
<th>( \Delta t )</th>
<th>( l_1(u) )</th>
<th>( l_2(u) )</th>
<th>( l_\infty(u) )</th>
<th>( q^\text{emp} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 ( \times ) 10</td>
<td>1800</td>
<td>7.72 ( \times ) 10(^{-3} )</td>
<td>1.51 ( \times ) 10(^{-3} )</td>
<td>1.35 ( \times ) 10(^{-2} )</td>
<td>-</td>
</tr>
<tr>
<td>40 ( \times ) 20</td>
<td>900</td>
<td>7.46 ( \times ) 10(^{-7} )</td>
<td>2.78 ( \times ) 10(^{-4} )</td>
<td>6.77 ( \times ) 10(^{-3} )</td>
<td>2.5</td>
</tr>
<tr>
<td>80 ( \times ) 40</td>
<td>450</td>
<td>4.69 ( \times ) 10(^{-5} )</td>
<td>1.00 ( \times ) 10(^{-5} )</td>
<td>2.03 ( \times ) 10(^{-4} )</td>
<td>4.8</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>( N_x \times N_y )</th>
<th>( \Delta t )</th>
<th>( l_1(v) )</th>
<th>( l_2(v) )</th>
<th>( l_\infty(v) )</th>
<th>( q^\text{emp} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>20 ( \times ) 10</td>
<td>1800</td>
<td>8.23 ( \times ) 10(^{-3} )</td>
<td>1.09 ( \times ) 10(^{-3} )</td>
<td>9.98 ( \times ) 10(^{-3} )</td>
<td>-</td>
</tr>
<tr>
<td>40 ( \times ) 20</td>
<td>900</td>
<td>7.87 ( \times ) 10(^{-5} )</td>
<td>1.58 ( \times ) 10(^{-4} )</td>
<td>2.66 ( \times ) 10(^{-3} )</td>
<td>2.8</td>
</tr>
<tr>
<td>80 ( \times ) 40</td>
<td>450</td>
<td>8.12 ( \times ) 10(^{-6} )</td>
<td>1.84 ( \times ) 10(^{-5} )</td>
<td>4.70 ( \times ) 10(^{-4} )</td>
<td>3.1</td>
</tr>
</tbody>
</table>

6.2. Unsteady flow with analytic solution

In a second, time dependent test, the analytic solution of the shallow water equations on the sphere derived in [37] has been employed to assess the performance of the proposed discretization. More specifically, the analytic solution defined in formula (23) of [37] was used. Since the exact
solution is periodic, the initial profiles also correspond to the exact solution an integer number of days later. The proposed SISLDG scheme has been integrated up to $t_f = 5$ days with $p^\eta = 4$ and $p^u = 5$ on a mesh of $20 \times 10$ elements, while the time step has been set to $\Delta t = 1800s$. In this case, the maximum Courant number based on the velocity is around 2. Error norms for $\eta, u, v$ of the above-mentioned integration have been computed at $t_f = 5$ days and displayed in tables 4 - 6 for the distorted and for the aligned mesh. This test shows that no significant effects are brought on the solution by the mesh distortion, even in consideration of the quite high polynomial degrees ($p^\eta = 4$ and $p^u = 5$) employed.

Table 4. Relative errors on $\eta$, distorted vs. aligned mesh, Läuter test case, TR-BDF2 time discretization. $\Delta t$ in seconds.

<table>
<thead>
<tr>
<th>mesh</th>
<th>$l_1(\eta)$</th>
<th>$l_2(\eta)$</th>
<th>$l_\infty(\eta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>distorted</td>
<td>$1.574439 \times 10^{-3}$</td>
<td>$2.015191 \times 10^{-3}$</td>
<td>$6.223918 \times 10^{-3}$</td>
</tr>
<tr>
<td>aligned</td>
<td>$1.574433 \times 10^{-3}$</td>
<td>$2.015189 \times 10^{-3}$</td>
<td>$6.220938 \times 10^{-3}$</td>
</tr>
</tbody>
</table>

Table 5. Relative errors on $u$, distorted vs. aligned mesh, Läuter test case, TR-BDF2 time discretization. $\Delta t$ in seconds.

<table>
<thead>
<tr>
<th>mesh</th>
<th>$l_1(u)$</th>
<th>$l_2(u)$</th>
<th>$l_\infty(u)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>distorted</td>
<td>$3.062825 \times 10^{-2}$</td>
<td>$3.816815 \times 10^{-2}$</td>
<td>$7.295160 \times 10^{-2}$</td>
</tr>
<tr>
<td>aligned</td>
<td>$3.062796 \times 10^{-2}$</td>
<td>$3.816801 \times 10^{-2}$</td>
<td>$7.293568 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Table 6. Relative errors on $v$, distorted vs. aligned mesh, Läuter test case, TR-BDF2 time discretization. $\Delta t$ in seconds.

<table>
<thead>
<tr>
<th>mesh</th>
<th>$l_1(v)$</th>
<th>$l_2(v)$</th>
<th>$l_\infty(v)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>distorted</td>
<td>$2.105254 \times 10^{-2}$</td>
<td>$2.195037 \times 10^{-2}$</td>
<td>$3.832416 \times 10^{-2}$</td>
</tr>
<tr>
<td>aligned</td>
<td>$2.105328 \times 10^{-2}$</td>
<td>$2.195039 \times 10^{-2}$</td>
<td>$3.833373 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

For comparison, analogous errors have been computed with the same discretization parameters but employing the off-centered Crank Nicolson method with $\theta = 0.6$. The results are displayed in tables 7 - 9 both for distorted and aligned mesh. Again the difference in error norms in the two mesh configurations is negligible, confirming the flexibility of the proposed formulation with respect to the mesh choice itself. By comparing tables 4 - 6 and 7 - 9, it can also be recognized that a substantial improvement in the errors between the TR-BDF2 scheme and the off-centered Crank Nicolson is attained.

In conclusion, these preliminary results confirm that the extension of the formulation proposed in [1] to generic structured meshes of (in general distorted) quadrilaterals and to the flux form of the continuity equa-
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes

Table 7. Relative errors on $\eta$, distorted vs. aligned mesh, Läuter test case, off-centered Crank Nicolson time discretization. $\Delta t$ in seconds.

<table>
<thead>
<tr>
<th>mesh</th>
<th>$l_1(\eta)$</th>
<th>$l_2(\eta)$</th>
<th>$l_\infty(\eta)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>distorted</td>
<td>$8.416259 \times 10^{-3}$</td>
<td>$1.019254 \times 10^{-2}$</td>
<td>$1.875049 \times 10^{-2}$</td>
</tr>
<tr>
<td>aligned</td>
<td>$8.414481 \times 10^{-3}$</td>
<td>$1.019255 \times 10^{-2}$</td>
<td>$1.875039 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Table 8. Relative errors on $u$, distorted vs. aligned mesh, Läuter test case, off-centered Crank Nicolson time discretization. $\Delta t$ in seconds.

<table>
<thead>
<tr>
<th>mesh</th>
<th>$l_1(u)$</th>
<th>$l_2(u)$</th>
<th>$l_\infty(u)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>distorted</td>
<td>$1.462592 \times 10^{-1}$</td>
<td>$1.737328 \times 10^{-1}$</td>
<td>$2.968882 \times 10^{-1}$</td>
</tr>
<tr>
<td>aligned</td>
<td>$1.462585 \times 10^{-1}$</td>
<td>$1.737323 \times 10^{-1}$</td>
<td>$2.969513 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

Table 9. Relative errors on $v$, distorted vs. aligned mesh, Läuter test case, off-centered Crank Nicolson time discretization. $\Delta t$ in seconds.

<table>
<thead>
<tr>
<th>mesh</th>
<th>$l_1(v)$</th>
<th>$l_2(v)$</th>
<th>$l_\infty(v)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>distorted</td>
<td>$2.834871 \times 10^{-1}$</td>
<td>$2.8788142 \times 10^{-1}$</td>
<td>$3.741655 \times 10^{-1}$</td>
</tr>
<tr>
<td>aligned</td>
<td>$2.834917 \times 10^{-1}$</td>
<td>$2.8788145 \times 10^{-1}$</td>
<td>$3.741587 \times 10^{-1}$</td>
</tr>
</tbody>
</table>

7. Conclusions and future perspectives

We have proposed a numerical formulation that extends in several respects the results of [1] for the shallow water equations in spherical geometry. The novelty presented here is the application to general structured meshes of quadrilaterals (not aligned with the coordinate directions) and the conservation of mass. Numerical results in section 6 show that no substantial accuracy loss is caused by the mesh distortion, even for relatively high polynomial degrees, while both mass preservation and an higher level of accuracy than a standard off-centered Crank Nicolson approach are provided.

As the method presented [1] (and its previous version in [10]) can be applied in principle on arbitrarily unstructured and even nonconforming meshes, this extension represents only a first step towards more general implementations on fully nonstructured meshes.

The conservation of mass itself is attained here by using a mixed Eulerian (for the continuity equation in flux form) semi-Lagrangian (for the momentum equation) formulation, and a more ambitious extension to a fully conservative and semi-Lagrangian formulation on the lines of [11] is
left for further studies.

Acknowledgements

This research work has been supported financially by the Abdus Salam International Center for Theoretical Physics, Earth System Physics Section. I am extremely grateful to Filippo Giorgi of ICTP for his strong interest in this work and his continuous support. Financial support has also been provided by the INDAM-GNCS project *Metodi numerici semi-impliciti e semi-Lagrangiani per sistemi iperbolic di leggi di bilancio*. I would also like to acknowledge useful conversations on the topics of this paper with L. Bonaventura.

REFERENCES


9. T. Davies, M. Cullen, A. Malcolm, M. Mawson, A. Staniforth, A. White,
A mass conservative TR-BDF2 DG discretization of the SWE on general meshes


22. Y. Saad and M. Schultz, GMRES: A generalized minimal residual al-
G. Tumolo


A mass conservative TR-BDF2 DG discretization of the SWE on general meshes
