Stability analysis of unstructured finite volume methods for linear shallow water flows using pseudospectra and singular value decomposition

Abdelaziz Beljadid\textsuperscript{a,}\textsuperscript{*}, Abdolmajid Mohammadian\textsuperscript{b}, Hazim Qiblawey\textsuperscript{b}

\textsuperscript{a}Department of Civil Engineering, University of Ottawa, 161 Louis Pasteur, Ottawa, Ontario, K1N 6N5, Canada
\textsuperscript{b}Department of Chemical Engineering, Qatar University, P.O. Box 2713, Doha, Qatar

\textbf{A R T I C L E   I N F O}

Article history:
Received 23 November 2015
Revised 23 May 2016
Accepted 18 July 2016
Available online 19 July 2016

\textbf{Keywords:}
Linear shallow water equations
Finite volume method
Unstructured mesh
Pseudospectra
Singular value decomposition

\textbf{A B S T R A C T}

The discretization of the shallow water system on unstructured grids can lead to spurious modes which usually can affect accuracy and/or cause stability problems. This paper introduces a new approach for stability analysis of unstructured linear finite volume schemes for linear shallow water equations with the Coriolis Effect using spectra, pseudospectra, and singular value decomposition. The discrete operator of the scheme is the principal parameter used in the analysis. It is shown that unstructured grids have a large influence on operator normality. In some cases the eigenvectors of the operator can be far from orthogonal, which leads to amplification of solutions and/or stability problems. Large amplifications of the solution can be observed, even for discrete operators which respect the condition of asymptotic stability, and in some cases even for Lax–Richtmyer stable methods. The pseudospectra are shown to be efficient for the verification of stability of finite volume methods for linear shallow water equations. In some cases, the singular value decomposition is employed for further analysis in order to provide more information about the existence of unstable modes. The results of the analysis can be helpful in choosing the type of mesh, the appropriate placements of the variables of the system on the grid, and the suitable discretization method which is stable for a wide range of modes.

© 2016 Elsevier Ltd. All rights reserved.

1. Introduction

Shallow water equations (SWEs) are the depth-averaged forms of Navier–Stokes equations under the assumptions that the vertical scales are much smaller than the horizontal ones, the velocity profile exhibits small changes throughout the depth, and the pressure is close to hydrostatic (e.g. Vreugdenhil, 1994). The system of SWEs is a hyperbolic system of conservation laws in the absence of viscous terms. Since these systems have a conservative form, finite volume (FV) methods could be a very convenient choice for their modeling. In the presence of source terms in the SWEs, FV schemes may lead to numerical oscillations. FV methods have been extensively examined in many previous studies by using upwind schemes, central schemes, and central-upwind schemes (e.g. Beljadid et al., 2012, 2013; Khouider and Majda, 2005a; Kurganov and Petrova, 2005, 2007; Kurganov and Tadmor, 2000; Lin et al., 2003; Mohammadian and Le Roux, 2006; Nessyahu and Tadmor, 1990; Russo, 2005).

In this study, we consider the linear shallow water equations (LSWEs) in the presence of the Coriolis term. The LSWEs appear in multi-scale models developed for some atmospheric flows (see Khouider and Majda, 2005a, 2005b) and their solution in the presence of the Coriolis effect is a challenging issue (Khouider and Majda, 2005b). They also appear in spectral methods based on vertical mode decomposition in the atmosphere (Majda, 2002). The development of stable FV methods of sufficient accuracy for LSWEs over unstructured grids is a challenging problem. Unstructured grids provide good flexibility for discretizing complex domains and local mesh refinement depending on the desired precision and boundary conditions. Several possibilities are available for the choice of control volumes, and different discrete formulations of the equations may be used, which can sometimes be a source of spurious modes leading to stability problems.

Pseudospectra have been used in several studies (e.g. Demmel et al., 1990; Reddy, 1994; Trefethen, 1997, 1999; Trefethen et al., 2001). Trefethen and Embree (2005) provide more references and details about pseudospectra and the behavior of non-normal matrices. The discrete operator of a numerical scheme which has a set of orthogonal eigenvectors is considered normal (Trefethen and Embree, 2005). For normal operators, each eigenvalue has condition

\footnote{\textsuperscript{*}Corresponding author. Present address: Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA.
E-mail address: beljadid@mit.edu (A. Beljadid).}

\url{http://dx.doi.org/10.1016/j.advwatres.2016.07.011}

\url{0309-1708/© 2016 Elsevier Ltd. All rights reserved.}
number equal to 1. However, most discrete operators of schemes developed for shallow water equations are non-normal since they have non-orthogonal eigenvectors. For systems which are governed by a non-normal operator, the solutions may have large amplifications for finite times, and the eigenvalues of the operator are sensitive to perturbations. Pseudospectra analysis and singular value decomposition are less employed in stability studies of numerical schemes compared to the commonly used methods such as asymptotic stability and Lax–Richtmyer stability. High-order finite-difference methods were analyzed by Zingg (1997) and Zingg and Lederle (2005) using spectra, pseudospectra, and singular value decomposition. The authors stated the importance of the use of these methods for detecting the instability of finite-difference schemes.

The objective of this paper is to introduce the pseudospectra and singular value decomposition method (SVD) as a new approach to study the stability of linear finite volume schemes. In our stability analysis, we used two-dimensional numerical schemes which are inspired from the one-dimensional numerical scheme used by Trefethen and Embree (2005). This scheme is an unstable implicit method which uses a second-order centered scheme in space and Crank–Nicolson method for temporal integration. Zingg and Lederle (2005) used this scheme in the stability analysis of high-order finite-difference methods.

The different temporal and spatial discretizations used to define the numerical schemes are employed in order to study and illustrate the importance of the proposed approach for stability analysis of schemes for LSWEs. This is the first study in which these stability analysis techniques are used in shallow water. It should be mentioned that the conditions of the asymptotic stability are related to the asymptotic behavior of the solutions. These conditions do not provide any information on the behavior of the solutions for finite times. In general, the discrete operators of most available schemes for SWEs are non-normal. In such cases, even if the spectrum corresponding to the fully-discrete form of the scheme has all eigenvalues within the unit circle, the scheme can have an unstable behavior or may lead to numerical amplifications of the solution for finite times. Lax–Richtmyer stability is the commonly used definition of stability. However, as will be shown in this paper for finite volume methods on unstructured grids, there are some cases in which a scheme is Lax–Richtmyer stable but one still faces some unstable modes. This may cause numerical oscillations if the scheme is applied to simulate physical phenomena, which includes different types of waves such as long and short waves in shallow water flows. Since there are various solutions for LSWEs that may have a very rich wave structure, stability analysis for a wide range of waves is required. In this paper, we will demonstrate the advantage of using pseudospectra analysis and singular value decomposition to detect instabilities of linear finite volume schemes over unstructured grids.

Several aspects are considered, such as the geometry of control volumes, the spatial methods used for the integration of the continuity and momentum equations, the boundary conditions, and the temporal integration schemes. Although the spatial schemes considered in this study are not comprehensive, they illustrate various strategies to analyze the stability of numerical methods and to choose the suitable unstructured scheme. Different mesh types are used in order to study their impact on nonmodal stability of the numerical schemes for LSWEs. As will be shown in this study, the results of the analysis can be helpful in choosing the type of mesh, the appropriate placements of the variables on the cells, and to test the discretization techniques for a wide range of modes. In our analysis, twenty numerical schemes are considered in which we use five spatial methods combined with four temporal integration schemes: the backward Euler method, the Crank–Nicolson method and the two- and three-step Adams–Bashforth schemes, which are used in many studies (e.g. Fringer et al., 2006; Marshallek et al., 1997; Walters et al., 2009a; 2009b).

This paper is organized as follows. In Section 2, LSWEs and the numerical schemes are presented. In Section 3, we present the discrete operators of the finite volume methods and the stability analysis techniques. In Section 4, the global analysis methodology used in the study is presented. Some numerical stability tests are performed in Section 5 using pseudospectra and singular value decomposition, and several numerical tests are performed using Kelvin, Poincaré and pure gravity waves in order to confirm the results of the analysis. In Section 6, a discussion and summary of the main findings of the paper are presented. Some concluding remarks complete the study.

2. Discretization of shallow water equations

2.1. Shallow water equations

The inviscid LSWEs are written in Cartesian coordinates as (Vreugdenhil, 1994):

\begin{align}
\eta_t + H \nabla \cdot \mathbf{u} &= 0 \\
\mathbf{u}_t + f \mathbf{k} \times \mathbf{u} + \nabla \eta &= 0,
\end{align}

(1)

where $H$ is the mean depth, $\eta$ represents the water surface elevation with respect to the reference plane, $\mathbf{u} = (u, v)^T$ is the vector composed of the depth-averaged velocity components in the $x$- and $y$-directions, respectively, $f$ is the Coriolis parameter, $g$ is the gravity acceleration, $\mathbf{k}$ is the unit vector in the vertical direction, and $(H + \eta)$ is the total water depth.

The LSWEs, where the conserved variable $\eta$ is used in the continuity equation and the primitive variables $(u, v)^T$ are used in the momentum equations, describe the dominant wave phenomena very well. Most of the relevant properties of the solutions for nonlinear SWEs are obtained by considering the linearized model (Vreugdenhil, 1994). The use of finite volume methods to solve this model is important since we will ensure the conservation of mass, where the conserved variable $\eta$ is used in the continuity equation and a good accuracy in the momentum conservation, where the primitive variables are considered in the momentum equations under some approximations.

In our stability analysis and some of the numerical tests, we use periodic boundary conditions in both $x$- and $y$-directions. In other numerical cases, we use the wall boundary condition in the $y$-direction and periodic boundary condition in the $x$-direction. We consider the approximation to the Coriolis parameter $f = f_0 + \beta y$, where $\beta$ is the linear coefficient of variation of $f$ with respect to $y$ and $f_0$ is a constant. The variable $y$ is the meridional distance from the equator (positive northward). The parameter $\beta$ is given as:

$$\beta = 2\bar{\omega}/R = 2.29 \times 10^{-11} \text{ m}^{-1} \text{s}^{-1},$$

(2)

where $\bar{\omega}$ and $R$ are the angular speed of the Earth’s rotation and the mean radius of the Earth respectively ($\bar{\omega} = 7.29 \times 10^{-5} \text{ rad s}^{-1}$, $R = 6371 \text{ km}$).

2.2. Finite volume schemes

2.2.1. Unstructured grid implementation

In this paper, we used a milder type of instability to test the ability of the new approach using pseudospectra and singular value decomposition method for stability analysis of linear finite volume methods for LSWEs. We used two-dimensional central schemes which are inspired from the one-dimensional example which can be considered in numerical analysis as the best known example of nonmodal behavior (Trefethen and Embree, 2005). The analysis is applied to LSWEs using some finite volume methods based on five grid configurations with respect to the location of the variables and
the control volume for each variable. The control volumes considered in this paper are shown in Fig. 1. The numerical tests are performed for the five grids using four temporal schemes: the backward Euler method, the Crank-Nicolson scheme and the two- and three-step Adams-Bashforth methods. For each grid of index i, the finite volume methods are denoted by i-BE, i-CN, i-AB2, and i-AB3 for the four temporal schemes respectively. The selected temporal schemes are used to test the ability of the proposed approach to detect the instability of finite volume methods for LSWEs. The analysis can be applied to finite volume methods using other types of temporal schemes.

The variables of the system are located at the geometric centers of triangles, at the vertices, or at the midpoints of the edges of triangles, as shown in Fig. 1. Numerical schemes using different control volumes for the velocity and the water height are also considered since they are applied in some cases to finite volume methods for hyperbolic systems (Chen et al., 2003; Danilov, 2012; Doyen and Gunawan, 2014). The mesh-vertex cell is defined by connecting the centers of mass of the surrounding triangles and the midpoints of the edges which have the center of the mesh-vertex as a common endpoint. The continuity equation and the momentum equations are integrated using different control volumes for the finite volume methods based on the grids 1, 2, and 5. For grid 1, the mesh vertex which is specified using dotted lines is used as control volume for the velocity and the triangular control volume is used for the water surface elevation. For grid 2, the mesh vertex and the triangular grid are used as control volumes for the water surface elevation and the velocity, respectively. For grid 5, the mesh vertex is used as control volume for the water surface elevation and the quadrilateral grid $N_1 G_1 N_2 G_2$ is used as control volume for the velocity. We denote by $\Omega_\eta$ and $\Omega_u$ the control volumes used for the continuity equation and the momentum equations respectively, and we consider the same control volume ($\Omega_\eta = \Omega_u$) for the finite volume methods using Grids 3 and 4.

2.2.2. Spatial discretization

The SWEs are integrated over all control volumes $\Omega_\eta$ and $\Omega_u$ as follows:

$$\int_{\Omega_\eta} (\eta_t + H \nabla \cdot \mathbf{u}) d\Omega_\eta = 0$$

$$\int_{\Omega_u} (\mathbf{u}_t + \nabla \cdot \mathbf{F} - \mathbf{S}) d\Omega_u = 0,$$

where $\mathbf{F} = (E, G)^T$ is the flux vector with $E = (g_\eta, 0)^T$ and $G = (0, g_\eta)^T$, and $\mathbf{S} = (f v, -f u)^T$ is the source term.

The Gauss divergence theorem is used to convert the surface integrals to the boundary integrals:

$$\int_{\Omega_\eta} H \nabla \cdot \mathbf{u} d\Omega_\eta = \int_{\Gamma_\eta} H \mathbf{u} \cdot \mathbf{n} d\Gamma_\eta$$

$$\int_{\Omega_u} \nabla \cdot \mathbf{F} d\Omega_u = \int_{\Gamma_u} \mathbf{F} \cdot \mathbf{n} d\Gamma_u,$$

where $\Gamma_\eta$ and $\Gamma_u$ are the boundaries of the control volumes, and $\mathbf{n}$ is the unit outward normal vector to the boundary of each control volume considered. Then, (3) and (4) lead to:

$$\frac{d}{dt} \int_{\Omega_\eta} \eta d\Omega_\eta = - \int_{\Gamma_\eta} H \mathbf{u} \cdot \mathbf{n} d\Gamma_\eta$$

$$\frac{d}{dt} \int_{\Omega_u} \mathbf{u} d\Omega_u = - \int_{\Gamma_u} \mathbf{F} \cdot \mathbf{n} d\Gamma_u + \int_{\Omega_u} \mathbf{S} d\Omega_u.$$

For all cases of the control volumes considered, the boundary integral $\int_{\Gamma_u} \mathbf{F} \cdot \mathbf{n} d\Gamma_u$ in (5) may be approximated by a summation over the cell edges using

$$\int_{\Gamma_u} \mathbf{F} \cdot \mathbf{n} d\Gamma_u = \sum_{k=1}^{M} \int_{\Gamma_k} \mathbf{F} \cdot \mathbf{n} d\Gamma_k = \sum_{k=1}^{M} (\mathbf{F}_k \cdot \mathbf{n}_k) l_k,$$

where $\Gamma_k$, $\mathbf{F}_k$, $\mathbf{n}_k$, and $l_k$, $k = 1, 2, ..., M$, are respectively the control volume edges, the outward fluxes, the unit outward normal vectors, and the lengths corresponding to the edges of the computational cells. For the mesh-vertex cell, we will use the notation $M = 2N$ for the number of its sides. The same form of Eq. (6) is obtained for the case of the surface elevation by replacing the flux $\mathbf{F}$ by the parameter $H \mathbf{u}$ and considering the geometrical characteristics of the computational cell used for the continuity equation.

In the following, we use the notations $(n_{kx}, n_{ky})^T$ for the components of the unit outward normal vector $n_k$ of the kth edge of the triangular grid. For grids 1 and 2, we denote by $n_{k1} = (n_{k1x}, n_{k1y})^T$ and $n_{k2} = (n_{k2x}, n_{k2y})^T$ the unit outward normal vectors to the two cell-vertex edges which are inside the kth neighboring triangle.
The approximation of the flux at these edges will be denoted by \( F_k \) and its components by \( F_{ik} = (g_{i}, n_{k})^{T} \) and \( G_{k} = (0, g_{y})^{T} \). The same notations are used in the case of grids 4 and 5 for the unit normal vectors to the two cell-vertex edges located on either side of the edge of the neighboring triangle (see Fig. 1) and for the approximation of the flux.

In the finite volume method based on Grid 1, the momentum equations are integrated over the mesh-vertex cell. The continuity equation is integrated over the triangles and the required interface values on each edge are calculated as averages of the values of the variables at the two ends of the edge. The summation over the cell edges given by Eq. (6) can be written explicitly as

\[
\sum_{k=1}^{M} (F_{k} \cdot n_{k}) l_{k} = \sum_{k=1}^{N} \left( F_{ik} (n_{k1}l_{k1} + n_{k2}l_{k2}) + G_{k} (n_{k1}l_{k1} + n_{k2}l_{k2}) \right).
\]

For surface elevation one obtains

\[
\int_{\Gamma_{k}} H \mathbf{u} \cdot n d\Gamma_{k} = \sum_{k=1}^{N} \left[ H (u_{i} + u_{y}) (n_{kx}l_{k1} + n_{ky}l_{k2}) + \frac{v_{i} + v_{y}}{2} n_{kk} l_{k} \right],
\]

where \((u_{k1}, v_{k1})^{T}\) and \((u_{k2}, v_{k2})^{T}\) are the velocity vectors at the two ends of the kth edge of the triangular control volume.

A similar method is used for Grid 2 by considering the Voronoi cells as control volumes for the surface elevation and the triangles as control volumes for the velocity. The summation over the cell interfaces given by Eq. (6) can be written explicitly in this case as

\[
\sum_{k=1}^{N} (F_{k} \cdot n_{k}) l_{k} = \sum_{k=1}^{3} g \frac{\eta_{k1} + \eta_{k2}}{2} \left[ n_{kx} E_{k} + n_{ky} G_{k} \right],
\]

where \(\eta_{k1}\) and \(\eta_{k2}\) are the water surface elevations at the endpoints of the kth edge of the triangular computational cell.

For surface elevation one obtains

\[
\int_{\Gamma_{k}} H \mathbf{u} \cdot n d\Gamma_{k} = \sum_{k=1}^{N} \left[ H (u_{k1} l_{k1} + n_{k2} l_{k2} + H v_{k} (n_{k1} l_{k1} + n_{k2} l_{k2}) \right],
\]

where \((u_{k}, v_{k})^{T}\) is the vector velocity at the node of the mesh vertex, which is the center of the kth neighboring triangular cell.

The finite volume methods based on Grids 3 and 4 use the average values of the variables to obtain the values at the interfaces of the triangles and mesh-vertex cells respectively. The summation over the cell interfaces given by Eq. (6) can be written explicitly for grid 3 as

\[
\sum_{k=1}^{3} (F_{k} \cdot n_{k}) l_{k} = \sum_{k=1}^{3} g \frac{\eta_{k1} + \eta_{k2}}{2} \left[ n_{kx} E_{k} + n_{ky} G_{k} \right],
\]

and

\[
\int_{\Gamma_{k}} H \mathbf{u} \cdot n d\Gamma_{k} = \sum_{k=1}^{3} \left[ H \frac{u_{i} + u_{y}}{2} n_{kx} l_{k1} + \frac{v_{i} + v_{y}}{2} n_{ky} l_{k1} \right],
\]

where \((\eta_{k1}, \eta_{k2})^{T}\) are the vectors of variables at the centers of the triangular cell \(\Omega_{k}\) and its neighboring cell \(\Omega_{ik}\) respectively.

For grid 4 one obtains

\[
\sum_{k=1}^{M} (F_{k} \cdot n_{k}) l_{k} = \sum_{k=1}^{N} g \frac{\eta_{k1} + \eta_{k2}}{2} \left[ n_{k1x} l_{k1} + n_{k2x} l_{k2} + n_{k1y} l_{k1} + n_{k2y} l_{k2} \right],
\]

\[
\int_{\Gamma_{k}} H \mathbf{u} \cdot n d\Gamma_{k} = \sum_{k=1}^{N} \left[ H \frac{u_{i} + u_{y}}{2} (n_{k1x} l_{k1} + n_{k2x} l_{k2}) + \frac{v_{i} + v_{y}}{2} (n_{k1y} l_{k1} + n_{k2y} l_{k2}) \right].
\]

where \((\eta_{k1}, \eta_{k2})^{T}\) and \((\eta_{ik}, \eta_{ik}, \eta_{ik})^{T}\) are the vectors of variables at the centers of the cell-vertex \(\Omega_{k}\) and its neighboring cell \(\Omega_{ik}\) respectively.

For grid 5, the Gauss divergence theorem is applied to the continuity equation using the values of the velocity vectors \((u_{ik}, v_{ik})^{T}\) which are known at the midpoints of the edges of triangles. These points are located on the boundary of the cell-vertex used as control volume. We obtain

\[
\int_{\Gamma_{k}} H \mathbf{u} \cdot n d\Gamma_{k} = \sum_{k=1}^{N} \left[ H u_{ik} (n_{k1x} l_{k1} + n_{k2x} l_{k2}) + H v_{ik} (n_{k1y} l_{k1} + n_{k2y} l_{k2}) \right]
\]

For the momentum equations, the control volume for grid 5 is the quadrilateral \(N_{1} G_{1} N_{2} G_{2}\) in which the surface elevations are known at the vertices \(N_{1}\) and \(N_{2}\). To obtain the values of the surface elevation at the vertices \(G_{1}\) and \(G_{2}\) a two-dimensional linear reconstruction on each side of the edge \(N_{1} N_{2}\) is applied. The Green-Gauss theorem is applied over the two triangles with the centers \(G_{1}\) and \(G_{2}\) to obtain the gradients \((\nabla \eta)_{i}, i = 1, 2,\) of the surface elevation which are used in the two linear reconstructions. For example in order to obtain the value of the surface elevation at \(G_{1}\) we use the following gradient for the reconstruction:

\[
(\nabla \eta)_{1} = (\eta_{k1} + \eta_{k2}) l_{k1} l_{k2},
\]

where \(\eta_{k1}\) and \(\eta_{k2}\) are the surface elevations at the endpoints of the kth triangle edge. Linear reconstruction is used to obtain the values of the surface elevation at the midpoints of the four sides of the quadrilateral \(N_{1} G_{1} N_{2} G_{2}\). These values are used to apply the Gauss divergence theorem to obtain the discretization of the flux term in the same way applied to the continuity equation.

To treat boundary conditions, for simplicity in the case of periodic boundaries we ensure that there is one-to-one correspondence between the primary triangular cells of the two periodic boundaries. In this way, the boundary conditions are formulated using the ghost cells for all types of the different computational grids used for numerical schemes. The ghost computational cells are additional cells outside the physical domain which are added in order to ensure the periodicity of the system. The values of the variables at the "outlet" periodic boundary are set equal to the corresponding ones at the "inlet" periodic boundary.

In Section 5.2.1, the numerical test for the scheme 1—CN is performed using the wall boundary condition in the y—direction. In this numerical test, the velocity variables are located directly on the boundary which are set equal to zero at each time step. The water surface elevation is located at the geometric centers of triangles, where the values of the velocity are known at the vertices of the triangles including those located on boundaries. These values are used in Eq. (8) to compute the water surface elevation at each time step.
Eqs. (5) and (6) and the boundary conditions are used to obtain the following semi-discrete form for each computational cell

\[
\begin{align*}
\frac{\partial \eta_i}{\partial t} & = \sum_{i=1}^{n} \delta_{ij} u_i + \sum_{i=1}^{n} \gamma_i v_i \\
\frac{\partial \eta_j}{\partial t} & = \sum_{i=1}^{n} \mu_{ij} \eta_i + \beta y_j v_j \\
\frac{\partial \eta_j}{\partial t} & = \sum_{i=1}^{n} \nu_i \eta_i - \beta y_j u_j,
\end{align*}
\]  

(17)

where \((\eta_i, u_i, v_i)^T\) is the approximation of the cell average of the solution, and \(i_j\) and \(j_j\) are the set of indices of the points used in the explicit formulation of the right-hand sides of Eqs. (5) and (6).

Note that the continuity equation and the momentum equations have different indices \(i\) for the schemes based on Grids 1, 2, and 5. The parameters \(\delta_{ij}, \gamma_i, \mu_{ij}, \) and \(v_j\) depend on the finite volume method and the geometry of the computational cell.

To characterize the mesh structure of the unstructured grids, the following definition of the skewness parameter is used

\[
\text{Skewness} = \frac{\text{Optimal Cell Size} - \text{Cell Size}}{\text{Optimal Cell Size}},
\]  

(18)

where for a triangular grid, the optimal cell size is the size of an equilateral triangle with the same circumscribed circle. In our study, we use the primary triangular grids with a skewness parameter of value less than 0.50.

3. Discrete operators of the schemes and a skewness analysis

3.1. Temporal discretization

The backward Euler method, the Crank–Nicolson method and the two- and three-step Adams–Bashforth methods are considered as temporal schemes (see Appendix A). The notations \(\Delta t\) and \(t^n := n \Delta t\) are used respectively for the time-step and the time at step \(n\). The approximation at time \(t^n\) of the cell averages of the water surface elevation and the components of the velocity are denoted respectively by \(\eta^n, u^n, \) and \(v^n\).

The four temporal schemes are applied to discretize the system (17). The time discretization for these schemes is given by Eqs. (44), (46) and (48) in Appendix A. For the backward Euler method and the Crank–Nicolson method, the global system can be written in the following form of dimension \(M \times M\), using the variable \(U^n := (\eta^n, u^n, v^n)^T\) with \(\eta^n := (\eta^n_1, \eta^n_2, \ldots, \eta^n_n), u^n := (u^n_1, u^n_2, \ldots, u^n_n)\) and \(v^n := (v^n_1, v^n_2, \ldots, v^n_n)\)

\[
AU^{n+1} = BU^n,
\]  

(19)

where \(A\) and \(B\) are the matrices of dimension \(M \times M\) deduced from Eqs. (17) and the boundary conditions. The dimensions \(p\) and \(q\) represent respectively the numbers of control volumes used to compute the water surface elevation and the velocity, and one can verify that \(p = q\) for the finite volume methods using the grids 3 and 4.

The time discretization using the two-step Adams–Bashforth method and the boundary conditions are used together to obtain the global system of equations in the matrix form by introducing the new global variable \(V^n := (U^n, U^n)^T\):

\[
V^{n+1} = CV^n,
\]  

(20)

where the discrete operator \(C\) is of dimension \(2M \times 2M\).

Similar to methods \(i\)-ABZ2, the time discretization using the three-step Adams–Bashforth method and the boundary conditions can be rewritten in a global matrix form by introducing a new global variable \(V^n := (U^n, U^{n+1}, U^n)^T\). Then the same form of Eq. (20) is obtained, where the operator \(C\) is of dimension \(3M \times 3M\). The coefficients of the matrices \(A, B,\) and \(C\) are obtained by using Eqs. (44), (46), and (48) (see Appendix A) and the boundary conditions. Finally, when backward Euler method or Crank–Nicolson method is used, we obtain a generalized eigenvalue problem with the parameter-dependent matrix \(B - \lambda A\), where \(\lambda\) denotes the variable of the eigenvalues. The same form of Eq. (20) can be obtained by taking \(C = A^{-1}B\), which has the same pseudospectra as that of \(B - \lambda A\) (Trefethen and Embree, 2005).

3.2. Stability analysis techniques

3.2.1. Stability

The spectrum and the spectral radius for any arbitrary matrix \(A\) are denoted here by \(\sigma(A)\) and \(\rho(A)\), respectively. The suitable norm to apply in the proposed approach using pseudospectra and singular value decomposition method is the \(L^2\)-norm, denoted by \(\|A\| = \max_{\|x\|=1} \|Ax\|\). The use of \(L^2\)-norm leads to strong links between pseudospectra and singular value decomposition since interesting properties are obtained as explained in Appendix B. If other norms are used, they will be specified in the notation. Following the previous section, if we consider a uniform time-step \(\Delta t\), the operator \(C\) will be independent of time and will depend only on the structure and size of the different grids and on the time-step size \(\Delta t\). A theory was developed by Lax and Richtmyer (1956) for stability analysis, which leads to the equivalence theorem. They proved that for a finite-difference approximation with an operator \(C\) which satisfies the consistency, the stability is a necessary and sufficient condition to ensure the convergence of the method. Additionally, they gave a definition of stability: the approximation based on the operator \(C\) is Lax–Richtmyer stable if for any fixed value of time \(T\), there is a constant \(\mathcal{C} \geq 1\) such that

\[
\|V^n\| \leq \mathcal{C} \|V^0\|,
\]  

(21)

for all \(n \geq 0\), where \(n \Delta t \in [0, T]\), the constant \(\mathcal{C}\) is independent of step \(n\) and \(V^0\) is the initial condition. This definition is extended by Lax and Richtmyer (1956) where they consider that the bound of the set \(\{C(\Delta t)\}^n\) is generally a continuous function of \(\Delta t\) in some interval \([0, \tau]\). Then, the new stability condition (Lax and Richtmyer, 1956) can be stated as: the approximation based on the operator \(C\) is Lax–Richtmyer stable if the set \(\{C(\Delta t)\}^n\) is uniformly bounded for \(0 < \Delta t < \tau\) and \(0 < n \Delta t \leq T\). Other definitions of stability are introduced in the literature, such as in Carpenter et al. (1994), Beam and Warming (1993), and Gustafsson et al. (1972).

Another concept is asymptotic stability, in which the behavior of the solution for large times is considered. It requires that the power of the operator \(C\) (i.e. \(C^n\)) is bounded for infinite time. For each positive integer \(n\), the spectral radius of the operator verifies the following inequality (Trefethen and Embree, 2005):

\[
\rho(C) \leq \|C^n\|^{1/n},
\]  

(22)

Then, for the case of asymptotic stability, it is necessary to have \(\rho(C) \leq 1\) since \(\|C^n\|\) is bounded. The asymptotic stability is based on the spectral radius condition in which it is necessary to have the eigenvalues of the operator \(C\) inside the unit circle. The strict inequality \(\rho(C) < 1\) leads to a zero limit when \(n\) tends to infinity for the parameter \(\|C^n\|\). However, even if a scheme is asymptotically stable, the solution for finite times may have excessive amplifications, especially in the case of non-normal matrices, and therefore the spectral radius is not necessarily a good indicator of the behavior of the scheme for finite times.

3.2.2. Pseudospectra and singular value decomposition

The concept of singular value decomposition (SVD) is defined in general for complex matrices. In our case we review the necessary
parameters for our study, where we consider real matrices. Any matrix $Q \in \mathbb{R}^M$ can be written in the form (Susanto et al., 1998):

$$Q = W_s W_q^T,$$

where the matrix $S$ is given as:

$$S = \begin{pmatrix} D & Z_1 \\ Z_2 & Z_3 \end{pmatrix}. \quad (24)$$

with $W_s W_s = I$, $W_q W_q = I$. $I$ is the identity matrix, and $Z_i$, $i = 1, 2, 3$ are zero matrices. The matrix $QQ^T$ is real and symmetric, so it is diagonalizable and its form leads to positive eigenvalues. The strictly positive eigenvalues are denoted by $s_j^2$. The diagonal matrix used in the SVD method is $D = \text{Diag}(s_1, s_2, \ldots, s_p)$, and the parameters $s_j$ are the singular values arranged in descending order $s_1 \geq s_2 \geq \cdots \geq s_p > 0$. The smallest singular value of the matrix $Q$ is denoted by $s_{\text{min}}(Q)$.

There are several definitions for the pseudospectra of a matrix (see e.g. Trefethen and Embree, 2005). In this paper we employ the definition based on the singular values and the $L^2$-norm. For this case, the $\varepsilon$-pseudospectrum of the operator $C$ is the set $\sigma_{\varepsilon}(C)$, defined as:

$$\sigma_{\varepsilon}(C) = \{ z \in \mathbb{C} : s_{\text{min}}(zI - C) < \varepsilon \}. \quad (25)$$

If the matrix $C$ is normal, the $\varepsilon$-pseudospectrum $\sigma_{\varepsilon}(C)$ is the union of the open disks of radius $\varepsilon$ which have the points of the spectrum as their centers. The $\varepsilon$-pseudospectrum is given in this case by

$$\sigma_{\varepsilon}(C) = \sigma(C) + \Delta_{\varepsilon}, \quad (26)$$

where $\Delta_{\varepsilon}$ is the open disk of radius $\varepsilon$ and center $z = 0$ in $\mathbb{C}$, and each complex number in $\sigma(C) + \Delta_{\varepsilon}$ is the sum of two complex numbers from $\sigma(C)$ and $\Delta_{\varepsilon}$. In this case all eigenvalues have a condition number equal to 1 and for any uniform perturbation, one will observe uniform evolution without any bulge in the $\varepsilon$-pseudospectrum. For a non-normal matrix, the $\varepsilon$-pseudospectrum can be large. For this case, even if the eigenvalues of the operator $C$ are inside the unit circle, one may observe large deformation of $\varepsilon$-pseudospectrum near the largest eigenvalue outside the unit circle when the operator is perturbed. Therefore, the use of pseudospectra is important to complete the stability analysis. Zingg (1997) and Zingg and Lederle (2005) studied the high-order finite-difference methods using spectras, pseudospectra, and singular value decomposition and showed the importance of the use of these methods for detecting the instability. As observed by Zingg and Lederle (2005) for the finite-difference methods, our numerical tests have shown that for some finite volume methods the $\varepsilon$-pseudospectrum presents negligible bulge but one observes significant amplification for some cases. In such cases, the use of the SVD method can be helpful to predict instabilities of numerical schemes.

In the following, we consider general explicit numerical methods ($Q^{n+1} = M(Q^n)$) where $M(J)$ is the numerical operator of schemes, $Q^n$ is the approximation of the solution with error $E^n$ and $q^n$ is the exact solution at time level $n$ ($Q^n = q^n + E^n$). The global error can be decomposed into two terms as follows

$$E^{n+1} = Q^{n+1} - q^{n+1} = M(q^n + E^n) - M(q^n) + \Delta t n^n,$$

where the local truncation error is defined as

$$\tau^n = \frac{M(q^n) - q^{n+1}}{\Delta t}. \quad (28)$$

The study of the local truncation error allows to bound the one-step error $\Delta t \tau^n$. The second error term $M(q^n + E^n) - M(q^n)$ is related to the effect of the numerical scheme on the previous global error $E^n$. This second term is very important for the stability of schemes and in general stability theories (LeVeque, 2002) are developed to bound this term. The behavior of the discrete operator of the scheme can enormously influence this term even if the norm of the previous global error $\|E^n\|$ is small. The discrete operator can be non-normal which leads to amplifications of the solutions in finite time and may leads to problem of convergence. The proposed approach is relevant for the analysis of this term since as will be shown in numerical stability tests, it leads to more insight on the behavior of the discrete operator of schemes and the numerical solutions in finite time compared to asymptotic stability and Lax–Richtmyer stability.

4. Global analysis methodology

The unstructured grid has a large impact on the structure of the discrete operators $C$ of numerical schemes. In general, one can obtain non-normal discrete operators with a high condition number. Some finite volume methods have almost-normal operators if uniform geometry of the computational cells is used. This case will be observed for the scheme 3-CN with $f_i = 0$ and $\beta_i = 1$, which uses the finite volume method based on the Grid 3 and the Crank–Nicolson time discretization method. There are many possible metrics to characterize the non-normality of matrices (e.g. Chaitin-Chatelin and Frayssé (1996); Henrici (1962)). The following scalar parameter, which goes back to Henrici, is used to quantify the distance from normality of the operator $C$

$$\tau(C) = \|C - C^T C\|/\|C\|^2. \quad (29)$$

First, we start by analyzing the spectrum of the finite volume method and its $\varepsilon$-pseudospectrum which provides more information about the behavior of the discrete operator of the scheme. As mentioned before, the definition of pseudospectra based on Eq. (25) is applied. The Lax–Richtmyer stability of the scheme is studied by analyzing the behavior of the power of the discrete operators of different sizes of the scheme. For an arbitrary large number $n_{\text{max}}$, the behavior of the parameter $\|C^n\|$ is studied in order to find the value of the number $n$, $n < n_{\text{max}}$, which leads to the maximum value of this parameter. This number is used to apply the SVD method for the matrix $C = W_s W_q^T$, where the matrices $W_s$, $S$, and $W_q$ are deduced from the matrix $C$, as explained in Section 3.2.2. The use of large number $n_{\text{max}}$ will further improve the analysis using SVD method. The maximum amplification is obtained for the singular vector $v_1$ corresponding to the maximum singular value of the matrix $C$, and we have $\|C^n\| = s_{\text{max}}(C^n)$. This maximum is reached by using this vector (see Appendix B) and if we use the matrices of the form $W_s \{u_1, u_2, \ldots, u_n\}$ and $W_q \{v_1, v_2, \ldots, v_n\}$, where $u_i$ and $v_i$ are the columns of these matrices, we obtain:

$$C^n v_1 = s_{\text{max}} u_1. \quad (30)$$

In the right-hand side of (30), some components of the vector $s_{\text{max}} u_1$ can be much larger than those of the vector $v_1$, hence some local amplifications can be observed. This is a good indicator which will be used to check if there are any unstable modes. In this paper, we use the terminology first singular vector for the singular vector corresponding to the maximum singular value of the operator $C$.

The proposed approach can be extended to the case where the time step is variable. The solution at each time level $t_{n+1}$ can be computed using

$$u_{n+1} = C_n u_n. \quad (31)$$

where the discrete operator $C_n$ at time level $n$ depends on the time step $\Delta t$. 

132 A. Beljadid et al. / Advances in Water Resources 96 (2016) 127–144
One obtains

$$U^{n+1} = \prod_{k=0}^{n} C_n U^0$$

Pseudospectra can be applied to the operator $C_n$. For an arbitrary large number $n_{\text{max}}$, we study the behavior of the parameter $\|C_n\|$, where $C_n = \prod_{k=0}^{n} C_n$ in order to find the unfavorable value of $n$, $n \leq n_{\text{max}}$, which leads to the maximum value of this parameter. The unfavorable value of $n$ is used to apply SVD method to the operator $C_n$. Pseudospectra can be applied to $C(\Delta t)$ for several values of $\Delta t$ in order to find the maximum value of time step $\Delta t_{\text{max}}$ which ensure the stability of the scheme. To use an adaptive time step, CFL condition can be applied with the condition that the time step should be less than $\Delta t_{\text{max}}$ which is obtained using the proposed approach.

The proposed techniques can be extended to study the stability of high-order finite volume methods and in the case of nonlinear shallow water equations, where the discrete operator is variable in time. We will assume that the discrete operator $C_n$ is constant between the time levels $t_n$ and $t_{n+1}$ and the solution at each time level $t_{n+1}$ satisfies Eq. (31). The discrete operator $C_n$ can be computed numerically and at each time level we obtain Eq. (32). Similar to the methodology described above, pseudospectra can be applied to the operator $C_n$ and SVD method can be applied to the operator $C_n$ using the unfavorable value of $n$, $n \leq n_{\text{max}}$ for an arbitrary large number $n_{\text{max}}$.

5. Numerical stability tests

In this section, numerical stability tests are performed and are presented in four parts. In each part we present the tests related to one of the temporal schemes, which is combined with the five spatial numerical methods explained in Section 2.2. We consider the backward Euler method, the Crank–Nicolson method and the two- and three-step Adams– Bashforth methods as temporal schemes. The spectra and pseudospectra of schemes are used in the analysis, and depending on the results, we complete our analysis by using the singular value decomposition if needed. For each finite volume method, we first use periodic boundary conditions in both $x$- and $y$-directions. In some cases, the tests are repeated by using the wall boundary condition in the $y$-direction and keeping the periodic boundary condition in the $x$-direction.

The discrete operators $C$ of the schemes are obtained by using the dimensionless form of the LSWEs. Eqs. (1) are converted into a dimensionless form on the equatorial $\beta$-plane using the variables $\tilde{x} = x/L, \tilde{y} = y/L, \tilde{H} = H/L, \tilde{u} = u/U$ and $\tilde{v} = v/U$. The reference values of the depth ($\bar{H}$), time ($\bar{t}$), length ($\bar{L}$) and velocity ($\bar{U}$) scales are expressed (Beljadid et al., 2012; 2013) as:

$$\bar{H} = H$$

$$\bar{t} = \beta^{-1/2} (gH)^{-1/4}$$

$$\bar{L} = \frac{1}{\beta \bar{t}}$$

$$\bar{U} = \frac{\bar{l}}{\bar{t}}$$

Using the above reference scales, the dimensionless system is obtained by setting $H = g = 1$ in the original system. A non-dimensional domain $[0, L] \times [0, L]$ is used with $L = 3$. In our analysis we use the ratios $\Delta t/d_m$, where $d_m$ is the smallest distance between the locations of the variables. This ratio corresponds to the Courant–Friedrichs–Lewy number for the dimensionless system. To confirm the results of the stability analysis, several numerical tests are performed using Kelvin, Poincaré and pure gravity waves.

5.1. Test cases 1 using the backward Euler method

5.1.1. Numerical tests with 1-BE and 2-BE

The finite volume methods 1-BE and 2-BE use the backward Euler method as temporal scheme and the spatial schemes based on Grid 1 and Grid 2, respectively. In the numerical tests we consider $f_c = 0$ and $\beta = 1$, and discrete operators $C$ of dimension $1200 \times 1200$ by using 400 triangles to discretize the computational domain. These numerical tests are performed using periodic boundary conditions and the time-step $\Delta t = 0.5d_m$. The spectrum of these schemes are inside the unit circle, which confirms that the schemes are asymptotically stable. The pseudospectra applied to the operator $C$ of the schemes are shown in Fig. 2, where the bulges are observed near the largest eigenvalues for the two numerical schemes.

5.1.2. Numerical tests with 3-BE

The finite volume method 3-BE uses the spatial scheme based on Grid 3 and the backward Euler method as temporal scheme. The operators $C$ with dimension $1200 \times 1200$ and the time-step $\Delta t = 0.5d_m$ are used. The numerical examples are performed using periodic boundary conditions. The pseudospectra of the discrete operator of the scheme in the case of $f_c = 0$ and $\beta = 1$ is shown in Fig. 3 and no bulge is observed near the largest eigenvalue. The singular value decomposition is used with $n_{\text{max}} = 300$ to study the amplification of the first singular vector in the case of $f_c = 1$ and $\beta = 0$. Amplification of the solution is observed in
5.2.2. Numerical tests with 2-CN

The method 2-CN uses the spatial scheme based on Grid 2 and the Crank–Nicolson method as temporal scheme. The Coriolis parameter is considered with \( f_c = 0 \) and \( \beta = 1 \), and we choose \( \Delta t = 0.5d_m \). The spectrum of this scheme is inside the unit circle, which confirms that the scheme is asymptotically stable. Based on numerical tests, we conclude that the power of the discrete operator of the scheme \( C \) is bounded. This method is an example of schemes in which the spectrum is inside the unit circle and the power of the operator is bounded but its behavior, as shown in Fig. 5 (left), influences the stability of the solution. For \( n_{\text{max}} = 1000 \), the maximum of \( C \) is reached at \( n = 50 \) (\( \|C^{50}\| = 3.0664 \)). The pseudospectra of the scheme are shown in Fig. 5 (right) in which a bulge is observed near the largest eigenvalue. Fig. 6 (left) shows the first singular vector \( \nu_1 \) of \( C^{50} \) and the vector \( C^{50}\nu_1 \), which is the solution of the scheme when this singular vector is taken as the initial condition. Following the tendency of the solution, we observe an amplification of the solution and the ratios between the peak of this solution, and the peak of \( \nu_1 \) is 3.82. Instability clearly appears in the case of the wall boundary conditions in the \( y \)-direction while keeping the periodicity along the \( x \)-axis (test b). Fig. 6 (right) shows the pseudospectra for this case using a time-step \( \Delta t = 0.3d_m \), where a large bulge is observed. A large amplification of the solution is observed at time \( t = 50\Delta t \) when the first singular vector is taken as initial condition and the factor of amplification of the peak for this case is 11.15.

5.13. Numerical tests with 4-BE and 5-BE

The numerical schemes 4-BE and 5-BE use the backward Euler method as temporal scheme and the spatial schemes based on Grid 4 and Grid 5, respectively. Numerical tests are performed using \( f_c = 0 \) and \( \beta = 0 \), and discrete operators \( C \) of dimension 1200 \( \times \) 1200 where 400 triangles are used to discretize the domain. We consider periodic boundary conditions in both \( x \)- and \( y \)-directions, and the time-step \( \Delta t = 0.5d_m \). The spectrum is inside the unit circle, and these methods are asymptotically stable. The pseudospectra show signs of instability for these schemes, where we observe small bulges near the largest eigenvalues as shown in Fig. 4.

5.2. Test cases 2 using the Crank–Nicolson method

5.2.1. Numerical tests with 1-CN

The finite volume method 1-CN uses the spatial scheme based on Grid 1 and the Crank–Nicolson method as temporal scheme. In the numerical tests we consider \( f_c = 0 \) and \( \beta = 1 \), and an operator \( C \) of dimension 1200 \( \times \) 1200 by using 400 triangles for the computational domain. First (test a), we consider periodic boundary conditions in both \( x \)- and \( y \)-directions and the time-step \( \Delta t = 0.5d_m \). The spectrum of this scheme is inside the unit circle, which confirms that the scheme is asymptotically stable. Based on numerical tests, we conclude that the power of the discrete operator of the scheme \( C \) is bounded. This method is an example of schemes in which the spectrum is inside the unit circle and the power of the operator is bounded but its behavior, as shown in Fig. 5 (left), influences the stability of the solution. For \( n_{\text{max}} = 1000 \), the maximum of \( C \) is reached at \( n = 50 \) (\( \|C^{50}\| = 3.0664 \)). The pseudospectra of the scheme are shown in Fig. 5 (right) in which a bulge is observed near the largest eigenvalue. Fig. 6 (left) shows the first singular vector \( \nu_1 \) of \( C^{50} \) and the vector \( C^{50}\nu_1 \), which is the solution of the scheme when this singular vector is taken as the initial condition. Following the tendency of the solution, we observe an amplification of the solution and the ratios between the peak of this solution, and the peak of \( \nu_1 \) is 3.82. Instability clearly appears in the case of the wall boundary conditions in the \( y \)-direction while keeping the periodicity along the \( x \)-axis (test b). Fig. 6 (right) shows the pseudospectra for this case using a time-step \( \Delta t = 0.3d_m \), where a large bulge is observed. A large amplification of the solution is observed at time \( t = 50\Delta t \) when the first singular vector is taken as initial condition and the factor of amplification of the peak for this case is 11.15. This case as shown in Fig. 3. The peak of the parameter \( \|C^n\| \) is reached at \( n = 82 \) and we obtain \( \|C^{82}\| = 1.175 \). The ratio between the peak of the solution at \( t = 82\Delta t \) which is obtained using the scheme when the first singular vector is taken as the initial condition and the peak of this first singular vector has a value of 2.179.

Fig. 3. Left: Eigenvalue spectrum and pseudospectra of the numerical method 3-BE in the case \( f_c = 0 \) and \( \beta = 1 \) with periodic boundary conditions. Right: The first singular vector of \( C^{50} \) (____), and the solution at time \( t = 82\Delta t \) (____) using the numerical method 3-BE in the case \( f_c = 1 \) and \( \beta = 0 \) with periodic boundary conditions.

Fig. 4. Eigenvalue spectrum and pseudospectra of the numerical methods 4-BE (left) and 5-BE (right) in the case \( f_c = 0 \) and \( \beta = 0 \) using periodic boundary conditions.
an operator \( \mathbf{C} \) of dimension 1200 \( \times \) 1200 where 400 triangles are used to discretize the domain. We consider periodic boundary conditions in both \( x \) - and \( y \) -directions, and we used the time-step \( \Delta t = 0.5d_m \). The spectrum is inside the unit circle, and this method is asymptotically stable. The pseudospectra of the scheme are shown in Fig. 7 (left), where we observe a small bulge near the largest eigenvalue compared to the pseudospectra of the scheme 1-CN.

5.2.3. Numerical tests with 3-CN

The finite volume method 3-CN uses the spatial scheme based on Grid 3 and the Crank–Nicolson method as the temporal scheme. The operator \( \mathbf{C} \) with dimension 1200 \( \times \) 1200 and the time-step \( \Delta t = 0.5d_m \) are used. This method and the method 4-CN presented below perform better, in the case \( f_c = 0 \) and \( \beta = 1 \), than the other numerical schemes considered in this paper. When an unstructured grid is considered with a uniform geometry of triangles and under periodic boundary conditions we obtain an operator \( \mathbf{C} \) with negligible distance from normality. For this case (test a), the discrete operator of the scheme is Lax–Richtmyer stable (\( \| \mathbf{C} \| \leq 1 \)). As shown in Fig. 7 (right), the entire spectrum of this method lies inside the unit circle. Its pseudospectra is shown in the same figure and no bulge is observed near the largest eigenvalue. We use singular value decomposition with \( n_{\text{max}} = 300 \) to study the amplification of the first singular vector. It can be observed that the maximum of the ratio \( \| \mathbf{C}^n v_1 \|_\infty / \| v_1 \|_\infty \) is reached at \( n = 80 \). Fig. 8 (left) shows the first singular vector \( v_1 \) of \( \mathbf{C}^{80} \) and the vector \( \mathbf{C}^{80} v_1 \), which is the solution obtained using the scheme when this singular vector is taken as the initial condition. Based on this figure, we do not observe any amplification, and we obtain a ratio of reduction of the peak of 0.9886.

For the second test, a general geometry of triangles is used and the behavior of \( \| \mathbf{C} \| \) for several sizes of \( \mathbf{C} \) is studied. Based on the numerical tests, this parameter is uniformly bounded. Therefore, the approximation based on this operator of the finite volume scheme is Lax–Richtmyer stable. The pseudospectra of the scheme show good results in Fig. 8 (right), since we observe uniform evolution near the largest eigenvalue outside the unit circle. The peak of the parameter \( \| \mathbf{C} \| \) is reached at \( n = 10 \), and its value is \( \| \mathbf{C}^{10} \| = 1.3670 \). The singular value decomposition is used at step \( n = 10 \), which is the most unfavorable case here. Fig. 10 (left) shows the first singular vector of \( \mathbf{C}^{10} \) and the solution at time \( t = 10\Delta t \) when this singular vector is taken as the initial condition. No sign of instability is observed in this case; the ratio between the peak of the solution at \( t = 10\Delta t \) and the peak of the first singular vector has a value of 1.365.

In the third test, the stability of the finite volume method 3-CN is analyzed using pure gravity waves (\( f_c = \beta = 0 \)). Fig. 9 shows...
the pseudospectra for this case, where we observe a bulge near the largest eigenvalue. The instability is more important using the pseudospectra for the case $f_c = 1$ and $\beta = 0$ as shown in the same figure. This case includes the solutions of Poincaré waves. In Section 5.5.2, a numerical test is performed using these waves where amplifications of the solution are observed.

5.2.4. Numerical tests with 4-CN

The finite volume method 4-CN uses the spatial scheme based on Grid 4 and the Crank–Nicolson method. First, we consider the case $f_c = 0$ and $\beta = 1$ using an operator $C$ of dimension $1200 \times 1200$ and a time-step $\Delta t = 0.5d_m$. The behavior of the parameter $\|C\|$ is studied for various cases by using discrete operators of
can not use the commonly used methods such as Fourier analysis. For these cases, the proposed approach using pseudospectra can be applied easily which is an important advantage of the proposed stability analysis method. For LSWEs, the discrete operator of the scheme is a constant matrix which depends on the time-step and the method used to discretize the equations. As explained in Section 4, the proposed approach can be extended to study the stability of numerical schemes for nonlinear shallow water equations where the discrete operator of the scheme is variable in time.

Note that if small values of the parameter $\beta$ are considered, signs of instability of the scheme 4-CN are observed using pseudospectra. Fig. 11 shows the pseudospectra of the scheme for the case $f_r = 0$ and $\beta = 0.1$, where we observe a bulge near the eigenvalue of largest magnitude. In Fig. 11, the pseudospectra of the scheme 4-CN are shown for the case $f_r = \beta = 0$, which corresponds to the so-called numerical scheme P1-P1 studied in Le Roux et al. (2007) using Fourier analysis. The proposed approach is able to detect the instabilities of the scheme and it leads to compatible results to those obtained using Fourier analysis. In Section 5.5.3, a numerical test is performed using pure gravity waves where the shallow water equations with nonlinear terms in the continuity equation are considered.

5.2.5. Numerical tests with 5-CN

This method uses the spatial scheme based on the unstructured Grid 5 and the Crank–Nicolson method. An operator $C$ of dimension $1200 \times 1200$ and a time-step $\Delta t = 0.5d_m$ are
considered. The pseudospectra of this scheme for the case \( f_c = 0 \) and \( \beta = 1 \) are shown in Fig. 12 (right), where a slight bulge is observed. To clearly see the behavior of the scheme for the most unfavorable modes, we use the singular value decomposition method. Fig. 12 (left) shows the first singular vector of \( \mathbf{C}^0 \) and the solution at time \( t = 50\Delta t \) when this singular vector is taken as the initial condition. For this case, \( \| \mathbf{C}^0 \| = 3.6933 \) but the ratio between the peak of the solution at \( t = 50\Delta t \) and the peak of the first singular vector has a value of 5.2402. The discrete operator of this scheme is very far from normality, and amplifications of the solutions can be observed in finite times.

The pseudospectra of the scheme 5-CN are shown in Fig. 13 for the case of pure gravity waves (\( f_c = \beta = 0 \)) and a bulge near the largest eigenvalue is observed. This scheme suffers from the presence of spurious modes more than the scheme 4-CN for the case of pure gravity waves. The first singular vector and the solution using this vector as initial condition are shown in Fig. 13, which show large amplifications of the solution.

5.3. Test cases 3 using the two-step Adams–Bashforth method

Here the numerical tests are performed by using the two-step Adams–Bashforth method as a temporal scheme with \( f_c = 0 \) and \( \beta = 1 \) for the five spatial methods. We consider a discretization using 400 triangles, which leads to the operator \( \mathbf{C} \) of dimension \( 2400 \times 2400 \). Following the numerical experiments, we observe that those methods suffer from severe stability problems. Fig. 14 (left) shows the pseudospectra of the method 5-AD2 for \( CFL = 0.5 \), in which we observe a sudden bulge near the largest eigenvalue.

For \( CFL = 0.1 \), the singular value decomposition is better-suited than the pseudospectra for visualizing the instability of the other schemes which are impractical for unstructured grids. The operators of the schemes using the two-step Adams–Bashforth method are far from normality. For \( CFL = 0.1 \), the coefficient of normality is \( \tau (\mathbf{C}) = 0.7 \) for the scheme 3-AD2, and it is about 0.9 for the other schemes.

5.4. Test cases 4 using the three-step Adams–Bashforth method

5.4.1. Numerical tests with 1-AD3

This finite volume method uses spatial scheme 1 combined with the three-step Adams–Bashforth method. The stability of this scheme is studied for the case \( f_c = 0 \) and \( \beta = 1 \). In the numerical test, we consider a discretization using 400 triangles, which leads to the operator \( \mathbf{C} \) of dimension \( 3600 \times 3600 \) and a time-step \( \Delta t = 0.5t_{\text{dp}} \) is used. The parameter \( \| \mathbf{C} \| \) is bounded, but the pseudospectra shown in Fig. 14 (right) show that the scheme is unstable since a local bulge is observed near the largest eigenvalue.

5.4.2. Numerical tests with 2-AD3 and 5-AD3

The numerical tests for the finite volume methods 2-AD3 and 5-AD3 were performed by using the three-step Adams–Bashforth
method with \( f_c = 0 \) and \( \beta = 1 \). We used a computational grid which leads to the discrete operators of the schemes of dimension \( 3600 \times 3600 \) and we used a time-step \( \Delta t = 0.5d_m \). Based on the numerical experiments, the power of the discrete operator of each scheme is not bounded, and thus these schemes are not asymptotically stable.

5.4.3. Numerical tests with 3-AD3

This finite volume method uses the spatial scheme 3 combined with the three-step Adams–Bashforth scheme. This test is performed using the Coriolis parameter with \( f_c = 0 \) and \( \beta = 1 \), and an operator of dimension \( 3600 \times 3600 \) and a time-step \( \Delta t = 0.5d_m \). We numerically studied the behavior of \( \|C^n\| \) for several sizes of \( C \), as shown in Fig. 15 (left) for the three dimensions \( 1296 \times 1296, 2304 \times 2304, \) and \( 3600 \times 3600 \). Based on the numerical tests, this parameter is uniformly bounded, and the approximation based on the operator \( C \) of the finite volume scheme is Lax–Richtmyer stable. Following Fig. 15 (right), the spectrum of the scheme is inside the unit circle and one obtains good result for the stability according to the pseudospectra, since a uniform evolution is observed near the largest eigenvalue outside the unit circle. For the operator \( C \) of dimension \( 3600 \times 3600 \), the maximum of \( \|C^n\| \) is reached at \( n = 4 \), and we have \( \|C^n\| = 4.2120 \). Following the SVD method we obtain a ratio of the peak of the solution to the peak of the initial condition equal to 4.1172, which is less than the value obtained by using \( L^2 \) norm. Similar to the scheme 3-CN, if general Coriolis parameter is used, the pseudospectra show signs of instability for scheme 3-AD3.

5.4.4. Numerical tests with 4-AD3

This finite volume method uses the spatial scheme 4 combined with the three-step Adams–Bashforth method. We consider an operator \( C \) of dimension \( 3600 \times 3600 \) and a time-step \( \Delta t = 0.5d_m \). The numerical test is performed using the Coriolis parameter with \( f_c = 0 \) and \( \beta = 1 \). We numerically studied the behavior of \( \|C^n\| \) for several sizes of \( C \) by considering the dimensions \( 1296 \times 1296, 2304 \times 2304, \) and \( 3600 \times 3600 \) of the operator. According to the numerical tests, the power of the operator is uniformly bounded (\( \|C^n\| < 3.6 \)) and the peak of all sizes of the operators is reached at \( n = 3 \). Therefore the approximation based on the operator \( C \) of the finite volume scheme is Lax–Richtmyer stable. As can be observed in Fig. 16 (right), the spectrum of the scheme is inside the unit circle and we obtain a good result for the stability according to the pseudospectra, since we observe uniform evolution near the largest eigenvalue outside the unit circle. The peak of \( \|C^n\| \) for an operator \( C \) of dimension \( 3600 \times 3600 \) is reached at \( n = 3 \), and we have \( \|C^n\| = 3.5762 \). Based on the singular value decomposition method, see Fig. 16 (left), we obtain a ratio of the peak of the

Fig. 14. Left: Eigenvalue spectrum and pseudospectra of method 5-AD2 using unstructured grid and periodic boundary conditions. Right: Eigenvalue spectrum and pseudospectra of method 1-AD3 using unstructured grid and periodic boundary conditions with \( \Delta t = 0.5d_m \).

Fig. 15. Left: Evolution of parameter \( \|C^n\| \) using 3-AD3 method for different sizes of the matrix operator of the numerical scheme: \( C_1 : 1296 \times 1296, C_2 : 2304 \times 2304, \) and \( C_3 : 3600 \times 3600 \). Right: Eigenvalue spectrum and pseudospectra of method 3-AD3 using unstructured grid and periodic boundary conditions with \( \Delta t = 0.5d_m \).
solution to the peak of the initial condition equal to 1.1253, which is less than the value obtained by using $L^2$ norm.

The same numerical test is performed using different computational cells in order to study the impact of the mesh density on the results. The values of the ratio $\|C^2v_1\|_\infty/\|v_1\|_\infty$ is computed for different cell sizes using the value of $n$ which leads to the maximum of amplification. The results are presented in Table 2, where we do not observe any amplification for the different mesh sizes.

Similar to the scheme 4-CN, when geometric Coriolis parameter is used, pseudospectra show signs of instability for scheme 4-AD3.

### 5.5. Numerical experiments

#### 5.5.1. Numerical tests for kelvin waves

In this section, numerical tests are performed for linear $\beta$-plane SWEs using Kelvin waves, which are in perfect geostrophic balance in the meridional direction since the meridional flow, $v$, is zero, and they are propagated in the eastward direction. These waves are exact solutions of Eqs. (1)

$$
\begin{align*}
&u(x, y, t) = -\cos(2\pi \frac{X - t}{L}) e^{-(y-L/2)^2/2}, \\
&v(x, y, t) = 0, \\
&\theta(x, y, t) = u(x, y, t),
\end{align*}
$$

where we consider a domain $[0, L] \times [0, L]$ with $L = 6$ (nondimensional) and $X = L/4$.

These solutions are in a steady state in a moving frame. When they are considered as an initial condition, they must be preserved, and in particular, their total energy over one spatial period must remain constant.

For the dimensionless form of linear SWEs, one obtains the total energy

$$
E_\mathrm{n}(t) = 1/2 \int_\Omega \rho(H^2 + v^2 + g\eta^2) d\Omega = 1/2 \int_\Omega \|\mathbf{U}\|_2^2 d\Omega,
$$

where $\mathbf{U} = (\eta, u, v)$.

The evolution of the total energy over time will be analyzed, since it gives an idea of the instability of the numerical schemes. Fig. 17 shows the evolution of the Kelvin total energy of the finite volume methods 1-CN, 2-CN, 3-CN, 4-CN, and 5-CN for two periods in space and time. For $f_\theta = 0$ and $\beta = 1$, the methods 3-CN and 4-CN are stable and they perform very well in the conservation of energy for Kelvin waves. The methods 1-CN and 5-CN have a bounded total energy but it increases in finite times, which is in agreement with our analysis using the pseudospectra and SVD, where it was demonstrated that there are some particular modes which present amplification for finite times. The numerical method 2-CN has some signs of instability when the pseudospectra are used, which is not visible in the behavior of the Kelvin total energy. The pseudospectra and SVD were able to detect the instability of the two schemes, which confirms the efficiency of those methods for stability analysis. The three dimensional view of the water surface elevation for the scheme 4-CN at one time period and the isolines of the solution for the scheme 3-CN are shown in Fig. 18 for the case $f_\theta = 0$ and $\beta = 1$. For this case, the obtained solutions are oscillation-free, which confirm the results of our analysis using pseudospectra and singular value decomposition.

#### 5.5.2. Numerical test for poincaré waves

Poincaré waves, also known as inertia-gravity waves, behave as gravity waves affected by the rotation. They are obtained from Eq. (1) when solutions of the form $(\theta, u, v) = (\tilde{\theta}, \tilde{u}, \tilde{v}) e^{ik(x - l/2-y)}$ are considered, where $\tilde{\theta}, \tilde{u}, \tilde{v}$ are the amplitudes. The parameters $k$ and $l$ are the wave numbers in the $x-$ and $y-$directions, respectively. After substituting the periodic solutions in (1), the following dispersion relation is obtained.

$$
\omega^2 = f^2 + \frac{gH(k^2 + l^2)}{kH}
$$

The Coriolis parameter $f$ is constant here. The special case $l = 0$ leads to the following solution which is used in our numerical experiments

$$
\begin{align*}
\eta(x, y, 0) &= A \cos(kx - \omega t), \\
u(x, y, 0) &= \frac{\omega A}{kH} \cos(kx - \omega t), \\
v(x, y, 0) &= \frac{f A}{kH} \sin(kx - \omega t)
\end{align*}
$$

We consider a domain $[0, 2L] \times [-L, L]$, the amplitude $A = 1$, the mean water depth $H = 1$, and the dimensionless wavelength $L = 1$. This numerical test is performed using the parameter values
Fig. 17. Change in total energy for Kelvin waves using the schemes i-CN, i = 1, 2, 3, 4, and 5 for two time periods with CFL = 0.5.

Fig. 18. The isolines of the scheme 3-CN (left) and the three dimensional view of the water surface elevation for the scheme 4-CN (right) at one time period.

Fig. 19. Left: water surface elevation at time t = 5 for Poincaré waves using the scheme 3-CN. Right: evolution of the maximum of the water surface elevation until time t = 5.

$g = 1$ and $f = 1$. The simulations are performed using the numerical scheme 3-CN with CFL=0.5 and the size of the computational cells $\Delta x = \Delta y = 0.1$. Fig. 19 (left) shows the solution at time $t = 5$, where the solution of Poincaré waves is considered as initial condition. The same figure shows the evolution of the maximum of the water surface elevation $\eta_{\text{max}}$ until time $t = 5$. As expected in our analysis using pseudospectra for the case $f_c = 1$ and $\beta = 0$, we observe small amplifications of the solution.

5.5.3. Numerical test for pure gravity waves

This numerical example is performed using pure gravity waves to test the scheme 4-CN. The shallow water equations are considered with the nonlinear term in the continuity equation

$$\eta_t + \nabla \cdot (hu) = 0,$$

$$u_t + (Hu + g \nabla \eta) = 0,$$

(37)

where $h = H + \eta$ is the total water depth. For momentum equations, the same discretization given by Eq. (13) is used. For the continuity equation, we apply the discretization given by (14) using the variables $h u_i$ and $h v_i$ instead of $H u_i$ and $H v_i$. The Crank-Nicolson scheme is used for temporal integration and a semi-implicit discrete system with a nonlinear term is obtained. This system can be written, using the variables $\eta_i$, as follows

$$U_{n+1} = M_1 U^n + M_2 N^n + M_3 N^{n+1},$$

(38)
where $M_i$ are matrices of dimension $M \times M$, $M = 3p$ and the vector $N$ of nonlinear terms is of dimension $3p$ and it is given by

$$N = (0, 0, \ldots, 0, \eta_1 u_1, \eta_2 u_2, \ldots, \eta_p u_p, \eta_1 v_1, \eta_2 v_2, \ldots, \eta_p v_p)^T$$

The system (38) is solved using an iterative method, where we consider the following equation

$$U^{n+1, m+1} = M_1 U^n + M_2 N^n + M_3 U^{n+1, m},$$

where $m$ identifies iteration level. The initial value for the iterative method is $U^{n+1, 0} = U^n$. Upon convergence in iteration, the difference in iteration $U^{n+1, m+1} - U^{n+1, m}$ approaches zero.

The numerical test is performed using the computational domain $[-L, L] \times [-L, L]$, where $L = 2.4$, and the following initial condition

$$h(x, y, 0) = \begin{cases} 1 + \delta (\rho_0^2 - \rho_1^2), & \rho_1 < \rho < \rho_0, \\ 1 + \delta (\rho_0^2 - \rho_1^2), & \rho_0 < \rho, \\ 1, & \text{otherwise}, \end{cases}$$

where $\rho_0 = 0.60L$, $\rho_1 = 0.55L$ and $\delta$ is an arbitrary constant parameter which is introduced to obtain non-smooth solutions, using large values of this parameter, to test the stability of the scheme. Fig. 20 shows the initial condition with $\delta = 5$ and the solution at time $t = 1.4$ of Eq. (37) using the numerical scheme 4-CN with $f_e = \beta = 0$, CFL = 0.5 and grid size $|\Omega| = 0.0407$. Large deformations of the solution are observed which confirms the results of the stability analysis presented in Section 5.2.4 using pseudospectra for pure gravity waves. The same conclusions are obtained in Le Roux et al. (2007), where the authors observed the presence of classical spurious elevation modes.

6. Summary of the results of stability analysis

In this section, we summarize the results of the stability analysis using the proposed approach. The finite volume methods 1-BE and 2-BE with $f = \beta y$ are unstable using pseudospectra. The finite volume method 3-BE is stable for $f = \beta y$ using pseudospectra and in the case $f = 1$, instabilities are detected using SVD method. Pseudospectra show signs of instability for the nonlinear schemes 4-BE and 5-BE using $f = 0$. The finite volume methods 1-CN and 2-CN are asymptotically stable, but the pseudospectra show that these methods are unstable. When the wall boundary conditions are considered, the pseudospectra show large bulges and the instabilities are amplified.

The finite volume methods 3-CN and 4-CN, with variable Coriolis parameter of the form $f = \beta y$, perform better than the other schemes considered in this paper. For this case when an unstructured grid is considered with uniform geometry of triangles and under periodic boundary conditions, we obtain a discrete operator with a negligible distance from normality. These schemes are Lax–Richtmyer stable and the pseudospectra confirm that there is no oscillation. However, in the case of pure gravity waves ($f = 0$), the pseudospectra show signs of instability of these schemes. The pseudospectra are able to detect the instabilities and we obtain consistent results to those obtained using Fourier analysis for the scheme 4-CN which corresponds to the so-called numerical scheme P1-P1. For the finite volume method 5-CN, the pseudospectra show a slight bulge, and the SVD method was very effective in detecting unstable modes for this scheme.

The finite volume methods using the two-step Adams–Bashforth method as a temporal scheme are unstable for LSWEs on unstructured grid. The finite volume method 1-AD3 is asymptotically stable. However, the pseudospectra are effective in detecting the instability for this scheme. The finite volume methods 2-AD3 and 5-AD3 are not asymptotically stable. The finite volume methods 3-AD3 and 4-AD3 lead to good results for stability, according to pseudospectra and SVD method in the case of the Coriolis parameter of the form $f = \beta y$. For pure gravity waves ($f = 0$), pseudospectra show signs of instability for these schemes. For the finite volume method 5-CN, the pseudospectra show a slight bulge, and the SVD method was very effective for detecting unstable modes for this scheme.

The asymptotic stability and the Lax–Richtmyer stability are not sufficient to obtain stable finite volume schemes for linear shallow water equations. In some cases the numerical schemes can be asymptotically stable and even stable in the sense of Lax–Richtmyer stability but pseudospectra can detect the instability of these schemes. In some cases the use of singular value decomposition method leads to more insight into the existence of particular unstable modes.

If the Coriolis parameter depends on the space $(f = f_e + \beta y)$ by using non-zero values of $\beta$, Fourier analysis can not be used. For this case, the proposed approach using Pseudospectra and singular value decomposition can be applied. The proposed approach could be extended to study the stability of high-order finite volume methods and in the case of nonlinear shallow water equations, where the discrete operator of the scheme is variable in time. Finally, the reader is reminded that it is not generally possible to measure the non-normality of the operator of schemes by using only one parameter given by Eq. (29). The use of this
operator is efficient only if we compare the finite volume methods which use the same temporal scheme.

7. Conclusions

The discrete form of finite volume methods on unstructured grids for shallow water equations can lead to spurious modes which usually can affect accuracy and/or cause stability problems. The appearance of oscillations is mainly due to the structure of the mesh, the placement of the variables on the system on the grid, and/or the employed finite volume method. Unstructured grids have large impact on the structure of the discrete operators of finite volume methods, which leads to non-normal matrices. For these matrices, the asymptotic stability and the Lax–Richtmyer stability are not sufficient to obtain the stability for all modes. In this study, pseudospectra and singular value decomposition were employed for detecting instabilities of linear finite volume methods for LSWEs over unstructured grids. The analysis has shown that the pseudospectra are very effective when studying the stability of finite volume methods. Moreover, in some cases the use of the singular value decomposition method can lead to more insight about the existence of some unstable modes. For the Crank–Nicolson method and the three-step Adams–Bashforth method, it is shown that it is important to consider the placement of the variables of the system together at the center of the computational cell.

The results of the analysis can be helpful in choosing the type of computational grid, the appropriate placements of the variables of the system on the grid, and the suitable discretization techniques for a wide range of modes. For linear shallow water equations with a source term which includes a variable Coriolis parameter, the commonly used methods for stability analysis such as Fourier analysis cannot be used. In this case, Pseudospectra and singular value decomposition can be applied. The proposed techniques for stability analysis can be extended to study the stability of high-order finite volume methods and in the case of nonlinear shallow water equations, where the discrete operator of the scheme is variable in time.

Acknowledgments

The authors thank the anonymous reviewers for their valuable comments and suggestions to improve the quality of the paper. This publication was made possible by NPRP Grant 4-935-2-354 from the Qatar National Research Fund (a member of the Qatar Foundation). The statements made herein are solely the responsibility of the authors.

Appendix A. Time integration methods

We consider the ODE defined by

\[ \frac{d\mathbf{U}}{dt} = \mathbf{F}(t, \mathbf{U}), \quad \mathbf{U}(0) = \mathbf{U}^0 \]  

(42)

A1. Crank–Nicolson method and backward Euler method

The Crank–Nicolson method and the backward Euler method are used to solve Eq. (42)

\[ \frac{\mathbf{U}^{n+1} - \mathbf{U}^n}{\Delta t} = \alpha \mathbf{F}(t_n, \mathbf{U}_n) + (1 - \alpha) \mathbf{F}(t_{n+1}, \mathbf{U}_{n+1}). \]  

(43)

where we used \( \alpha = 0.5 \) for Crank–Nicolson method and \( \alpha = 0 \) for backward Euler method, and \( \mathbf{U}^n \) is the numerical solution at time \( t_n = n\Delta t \).

The time discretization of Eq. (17) is written as:

\[ \frac{\eta_j^{n+1} - \eta_j^n}{\Delta t} = \alpha \sum_{i \notin j} \delta_i u_i^n + (1 - \alpha) \sum_{i \notin j} \delta_i u_i^{n+1} + \alpha \sum_{i \notin j} \gamma_i v_i^n + (1 - \alpha) \sum_{i \notin j} \gamma_i v_i^{n+1} \]

\[ u_j^{n+1} - u_j^n = \alpha \sum_{i \notin j} \mu_i \eta_i^n + (1 - \alpha) \sum_{i \notin j} \mu_i \eta_i^{n+1} + \alpha f_j v_j^n + (1 - \alpha) f_j v_j^{n+1} \]

\[ v_j^{n+1} - v_j^n = \alpha \sum_{i \notin j} v_i \eta_i^n + (1 - \alpha) \sum_{i \notin j} v_i \eta_i^{n+1} - \alpha f_j u_j^n - (1 - \alpha) f_j u_j^{n+1}, \]  

(44)

where \( f_j = f_c + \beta y_j \).

A2. Two-step Adam–Bashforth method

The two-step Adam–Bashforth method explicitly solves Eq. (42) by using the following expression:

\[ \mathbf{U}^{n+2} = \mathbf{U}^{n+1} + \frac{3}{2} \Delta t \mathbf{F}(t_{n+1}, \mathbf{U}_{n+1}) - \frac{1}{2} \Delta t \mathbf{F}(t_n, \mathbf{U}_n) \]  

(45)

The time discretization of Eq. (17) using the two-step Adam–Bashforth method is written as:

\[ \eta_j^{n+2} = \eta_j^{n+1} + 3/2 \Delta t \sum_{i \notin j} \delta_i u_i^{n+1} - 1/2 \Delta t \sum_{i \notin j} \delta_i u_i^n \]

\[ + 3/2 \Delta t \sum_{i \notin j} \gamma_i v_i^{n+1} - 1/2 \Delta t \sum_{i \notin j} \gamma_i v_i^n, \]

\[ u_j^{n+2} = u_j^{n+1} + 3/2 \Delta t \sum_{i \notin j} \mu_i \eta_i^{n+1} - 1/2 \Delta t \sum_{i \notin j} \mu_i \eta_i^n \]

\[ + 3/2 \Delta t f_j v_j^{n+1} - 1/2 \Delta t f_j v_j^n, \]

\[ v_j^{n+2} = v_j^{n+1} + 3/2 \Delta t \sum_{i \notin j} v_i \eta_i^{n+1} - 1/2 \Delta t \sum_{i \notin j} v_i \eta_i^n \]

\[ - 3/2 \Delta t f_j u_j^{n+1} + 1/2 \Delta t f_j u_j^n \]  

(46)

A3. Three-step Adam–Bashforth method

The three-step Adam–Bashforth method explicitly solves Eq. (42) by using the following expression:

\[ \mathbf{U}^{n+2} = \mathbf{U}^{n+1} + \frac{5}{12} \Delta t \mathbf{F}(t_{n+1}, \mathbf{U}_{n+1}) - \frac{16}{12} \Delta t \mathbf{F}(t_n, \mathbf{U}_n) \]

\[ + \frac{23}{12} \Delta t \mathbf{F}(t_{n+1}, \mathbf{U}_{n+1}) \]  

(47)

The time discretization of Eq. (17) using the three-step Adam–Bashforth method is written as:

\[ \eta_j^{n+3} = \eta_j^{n+2} + 23/12 \Delta t \sum_{i \notin j} \delta_i u_i^{n+2} - 4/3 \Delta t \sum_{i \notin j} \delta_i u_i^{n+1} \]

\[ + 5/12 \Delta t \sum_{i \notin j} \delta_i u_i^n + 23/12 \Delta t \sum_{i \notin j} \gamma_i v_i^{n+2} \]

\[ - 4/3 \Delta t \sum_{i \notin j} \gamma_i v_i^{n+1} + 5/12 \Delta t \sum_{i \notin j} \gamma_i v_i^n, \]

\[ u_j^{n+3} = u_j^{n+2} + 23/12 \Delta t \sum_{i \notin j} \mu_i \eta_i^{n+2} - 4/3 \Delta t \sum_{i \notin j} \mu_i \eta_i^{n+1} \]

\[ + 5/12 \Delta t \sum_{i \notin j} \mu_i \eta_i^n + 23/12 \Delta t f_j v_j^{n+2} \]

\[ - 4/3 \Delta t f_j v_j^{n+1} + 5/12 \Delta t f_j v_j^n \]  

(48)
\[ p_j^{n+3} = v_j^{n+2} + 23/12 \Delta t \sum_{i \neq j} v_i n_j^{n+2} - 4/3 \Delta t \sum_{i \neq j} v_i p_j^{n+1} + 5/12 \Delta t \sum_{i \neq j} v_i n_j^{n+2} - 23/12 \Delta t f_j u_j^{n+2} + 4/3 \Delta t f_j u_j^{n+1} - 5/12 \Delta t f_j p_j^n \]  
(48)

Appendix B. SVD method and \( L^2 \)-norm

In the SVD method, we use matrices of the form \( W_1 = [u_1, u_2, \ldots, u_N] \) and \( W_2 = [v_1, v_2, \ldots, v_N] \), where \( u_i \) and \( v_i \) are the columns of these matrices. These vectors are of norm 1 (this using the \( L^2 \)-norm: \( \| u_i \|_2 = \| v_i \|_2 = 1 \)).

We obtain

\[ C^i v_i = s_{\text{max}} u_i \]  
(49)

then

\[ \| C^i v_i \| = s_{\text{max}} \| u_i \| \]

The SVD method is based on the result that the \( L^2 \)-norm of any matrix \( A \) is the largest singular value of \( A \) which leads to:

\[ C^i = s_{\text{max}} \]

This can be justified using \( L^2 \)-norm as follows

\[ C^i = \sup_{x \neq 0} \frac{C^i x}{\| x \|} = \frac{W_1 S W_2^T x}{\| x \|} \]

The singular value decomposition of the matrix \( C^i \) leads to

\[ C^i = W_1 S W_2^T \]  
(50)

where the matrices \( W_1 \) and \( W_2 \) satisfy the conditions: \( W_1^T W_1 = I \) and \( W_2^T W_2 = I \). Using these conditions we obtain the following equalities

\[ \sup_{x \neq 0} \frac{C^i x}{\| x \|} = \sup_{x \neq 0} \frac{W_1 S W_2^T x}{\| x \|} = \sup_{x \neq 0} \frac{W_1^T S x}{\| x \|} \]

If we consider the variable \( y = W_1^T x \) which is equivalent to \( x = W_2 y \), then one obtains the following equalities

\[ \sup_{x \neq 0} \frac{C^i x}{\| x \|} = \sup_{y \neq 0} \frac{S y}{\| y \|} = \sup_{y \neq 0} \frac{S y}{\| y \|} \]

This leads to

\[ \sup_{x \neq 0} \frac{C^i x}{\| x \|} = \sup_{y \neq 0} \frac{\left( \sum_{i=1}^{m} s_i y_i^2 \right)^{1/2}}{\| y \|} \leq s_{\text{max}} \]  
(51)

The equality is reached for the above inequality for the case \( y = (1, 0, 0, \ldots, 0)^T \) and since \( x = W_2 y \) we conclude that the maximum in (51) is satisfied for \( v_i \) and we obtain

\[ C^i v_i \| = s_{\text{max}} \]  
(52)

References


