o3o3: A Variant of Spectral Elements with a Regular Collocation Grid

J. STEPPELER

Climate Service Center, Hamburg, Germany

J. LI

International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China

F. FANG

Applied Modelling and Computation Group, Department of Earth Science and Engineering, Imperial College London, London, United Kingdom

J. ZHU

International Center for Climate and Environment Sciences, Institute of Atmospheric Physics, Chinese Academy of Sciences, and Department of Atmospheric Sciences, University of Chinese Academy of Sciences, Beijing, China

P. A. ULLRICH

Department of Land, Air and Water Resources, University of California, Davis, Davis, California

(Manuscript received 14 August 2018, in final form 9 March 2019)

ABSTRACT

In this study, an alternative local Galerkin method (LGM), the o3o3 scheme, is proposed. o3o3 is a variant or generalization of the third-order spectral element method (SEM3). It uses third-order piecewise polynomials for the representation of a field and piecewise third-degree polynomials for fluxes. For the discretization, SEM3 uses the irregular Legendre–Gauss–Lobatto grid while o3o3 uses a regular collocation grid. o3o3 can be regarded as an inhomogeneous finite-difference scheme on a uniform grid, which means that the finite-difference equations are different for each group with three points. A particular version of o3o3 is set as an example of many possibilities to construct LGM schemes on piecewise polynomial spaces in which the basis functions used are continuous at corner points and function spaces having continuous derivatives are shortly discussed. We propose a standard o3o3 scheme and a spectral o3o3 scheme as alternatives to the standard method of using the quadrature approximation. These two particular schemes selected were chosen for ease of implementation rather than optimal performance. In one dimension, compared to standard SEM3, o3o3 has a larger CFL condition benefiting from the use of a regular collocation grid. While SEM3 uses the irregular Legendre–Gauss–Lobatto collocation grid, o3o3 uses a regular grid. This is considered an advantage for physical parameterizations. The shortest resolved wave is marginally smaller than that with SEM3. In two dimensions, o3o3 is implemented on a sparse grid where only a part of the points on the underlying regular grid are used for forecasting.

1. Introduction

o3o3, an alternative local Galerkin method (LGM; Steppeler and Klemp 2017), is a generalization of the third-order spectral element method (SEM3; Taylor et al. 1997; Taylor and Fournier 2010). It uses third-order piecewise cubic polynomials for the representation of both the fields and the fluxes. Here, "*onom*" is an abbreviation of an LGM where the fields are represented by *n*th-degree

Denotes content that is immediately available upon publication as open access.

Corresponding author: J. Li, ljx2311@mail.iap.ac.cn

DOI: 10.1175/MWR-D-18-0288.1

© 2019 American Meteorological Society

This article is licensed under a Creative Commons Attribution 4.0 license (http://creativecommons.org/licenses/by/4.0/).

polynomials and the fluxes by such of degree m. The SEM3 and o3o3 methods use the same representation of fields. Steppeler and Klemp (2017) defined the schemes o3o3 and o2o3 without investigating the accuracy and stability. o3o3 is somewhat related to the family of multimoment finite-volume methods described by Chen and Xiao (2008), where the latter does not make a definite basis function assumption.

In this study, LGM is used in both methods to obtain the discretization. This is quadrature approximation for SEM3. The details of the LGM procedure for o3o3 will be described in section 3. Both o3o3 and SEM3 cover the computational area with cells, which are cells of square $3dx \times 3dz$ in this study. For SEM3, we use dx and dz to denote the average distance between the degrees of freedom of each field; however, for a uniformly spaced method such as o3o3, dx and dz also denote the exact distance. Within the cells, the fields are represented by third-order serendipity interpolation (Steppeler 1976; Ahlberg et al. 1967). This means that the corner points define a bilinear spline. A third-order field component is added to this, which is based on amplitudes on the edges of cells. In one dimension (1D), this means that fields and fluxes are represented as third-degree piecewise polynomials. For simplicity, we express the theory in 1D. In each cell, a field $h(x)$ is represented by a third-degree polynomial and the polynomials belonging to different cells fit together continuously. With a regular grid $x_i = i \times dx$ ($i = 0, 1, 2, 3, \ldots$), the cells for both SEM3 and o3o3 are the intervals (x_i, x_{i+3}) ($i = 0, 3, 6, 9, \ldots$). The coefficients of the third-degree polynomials for each cell can equivalently be described by gridpoint values at four points x'_i, x'_{i+1}, x'_{i+2}, x'_{i+3}, within the cell (x_i, x_{i+3}). The x'_i are called collocation points. Both SEM3 and o3o3 have the same collocation points at the cell boundary: $x'_i = x_i$, $x'_{i+3} = x_{i+3}$ ($i = 0, 3, 6, 9, \ldots$). For SEM3, the inner collocation points x'_{i+1}, x'_{i+2} need to be chosen such that the four collocation points of a cell interval are Legendre–Gauss–Lobatto points and thus form an irregular set of points even for a regular cell structure. For o3o3, the points are chosen regularly: $x'_{i+1} = x_{i+1}$, $x'_{i+2} = x_{i+2}$. The irregular spacing of collocation points with SEM3 is considered a disadvantage for realistic models with physical parameterization (i.e., Herrington et al. 2019).

SEM3 and the SEM schemes of higher polynomial degree than 3 have been frequently considered for application in realistic models. These two schemes modify the concept of the classical Galerkin scheme [see Steppeler (1987) for a review] to be local with explicit time schemes and therefore are suitable for massively multiprocessing computers (Taylor et al. 1997). The suitability of SEM for parallel computing with explicit time schemes comes from avoiding nonlocality in the computation of derivatives. For finite-element methods (FEMs) that use the classical Galerkin scheme on piecewise polynomial spaces, nonlocality is caused by a global mass matrix. SEM avoids nonlocality by using a lumped mass matrix obtained from inexact quadrature. It has other interesting properties, such as preservation of conservation laws (Giraldo 2001; Taylor and Fournier 2010). SEM3 provides a fourth-order uniform approximation, which means that the approximation order does not drop below 3 at any point. Tomita et al. (2001) gave an example how a low-order approximation at some points, such as poles, can lead to grid imprinting, meaning that the grid is seen in the numerical solution. Models with a uniform order of approximation weaken such grid imprinting. Baumgardner and Frederickson (1985) gave an example of a model on the sphere with a second-order uniform approximation and Steppeler et al. (2008) defined a shallow-water scheme on the sphere of uniform third-order. Both models weakened grid imprinting, but were nonconserving. One of the advantages of SEM of all polynomial degrees is that their order is uniform. They can weaken grid imprinting and are suitable for irregular grids (Taylor et al. 1997). SEM combines a high-order approximation with conservation. This is considered as a major advantage of SEM, as older high-order models, such as that of Kalnay et al. (1977) are nonconserving.

While o3o3 may share all of the mentioned advantages with SEM3, in the present paper we investigate only the conservation of first-order moments by o3o3 and the uniform approximation order of at least 3. A regular collocation grid is considered a major advantage for physical parameterization (Herrington et al. 2019). We describe the continuity properties and alternative LGM schemes in section 2. Section 3 outlines the o3o3 and SEM3 methods and presents the governing equations and their discretizations in flux form employed in all subsequent experiments. Section 4 describes the performance of a time step. Section 5 illustrates the results of the homogeneous advection test for accuracy and stability of the o3o3 method. The practical performance of o3o3 in 2D and 3D is discussed in section 6 and the study concludes in section 7.

2. Continuity properties and alternative local Galerkin schemes

This study investigates the o3o3 scheme (Steppeler and Klemp 2017) as an example of an alternative to the SEM scheme. The latter is currently the only LGM scheme near practical application in realistic models for forecast and climate simulation (Dennis et al. 2012). Analogous to SEM, o3o3 uses continuous basis functions but employs

an alternative discretization which results into stability on the uniform grid.

The approximation space, consisting of continuous piecewise polynomial functions, is a basis of a number of important numerical approaches, such as FEMs, SEMs, third-degree method (Giraldo 2001; Ahlberg et al. 1967; Steppeler 1976). However, this is not the only possible choice and this section aims to look at other possibilities. In the interior of cells, a polynomial representation is used. In this respect, the basis function space is characterized by the order of the polynomial representation. It is possible to approximate the flux in a higher polynomial space than the field. Therefore, there are two orders n and m to characterize the approximation spaces, named onom, with n being the polynomial order for the field and m that of the fluxes. All methods tested so far use $n = m$, but Steppeler and Klemp (2017) suggested the untested method o2o3, approximating the flux to third order, while the field is approximated by second-order polynomials. The resulting methods would then be third order. The regularity of fields at the cell boundaries is another distinguishing element of approximation spaces. We consider C^0 as the space of continuous functions, C^1 as the space of differentiable functions, C^{-1} as the space of discontinuous functions and so on (Evans 1998). In connection with polynomial representations in the interiors of cells, the C^{-1} space is used to obtain approximations (Cockburn and Shu 2001). Such methods are called discontinuous Galerkin method, which are not considered in this paper. For the regularity of fields and fluxes at cell boundaries we call the methods cn'cm'. The total characteristic of an approximation space becomes onomcn'cm'. The special case after o3o3 considered here could more precisely be called o3o3c0c0. The method o2o3 defined but not explored by Steppeler and Klemp (2017) would in more extensive notation be called o2o3c0c1. For the special case of o3o3, the definition of grids, function systems and related LGM approximations will be given in sections 3 and 4. The LGM to be used with a space onom is not uniquely determined. The o3o3 scheme to be described in the following is given in two versions and many more versions are possible. The algebraic appearance of a scheme based on a space onom may be different. SEM schemes normally define the polynomial in the interior of a cell by collocation grid points (Taylor et al. 1997; Giraldo 2001). This representation is also used with o3o3, but second and third derivatives at cell centers are used as alternatives in this paper and are called the spectral space. Alternative to second and third derivatives could be one-sided field derivatives at cell boundaries, an option not followed in this paper. Such alternative spectral spaces would give an equivalent alternative arithmetical form to the LGM procedure. Such alternative arithmetic forms can lead to different numerical efficiencies for the same scheme.

The impact, that the different ways of organizing the calculations have on computational efficiency, is not investigated in this paper. However, it is clear that with the use of the second-order and third-order derivatives of the fields as amplitudes ($h_{xx,i}$ and $h_{xxx,i}$) and performance of the time step in gridpoint space, SEM3 involves the same transformation effort as o3o3. The spectral calculations may be marginally cheaper with o3o3. Therefore, it may be expected that the performance of one time step causes about the same cost as SEM3. This is marginally more expensive than the standard classical fourth-order (o4) scheme (see Steppeler et al. (2008) and Kalnay et al. (1977) for applications). A performance of o3o3 in spectral space for the edge amplitudes would bring the cost of o3o3 into the order of that of the o4 scheme. This latter option would require that the physics scheme be performed at corner points only, which would save computer time by itself. o2o3 is a method using the space c0c1 and its analysis is left for future work. With that said, a calculation by this study's authors suggested a CFL condition of 1.8. Although the time-stepping procedure of o2o3 is somewhat more complicated than that of o3o3, this work indicates there remains a large family of possible methods that are in need of further study. For example, it is possible to use C^1 as approximation space for $h(y)$. If the fluxes are also approximated in C^1, the result of differentiation would be in C^0, then a continuous function must be approximated by a C^1 function, being differentiable at corner points. This would lead to a new version of an o3o3 scheme. It is also possible to approximate the flux in a higher space, allowing for continuous derivatives, meaning the the second-derivative exists at corner points. Obviously such approximations for the flux can be done in C^2. Therefore, the resulting scheme would be classified as o3o5. There is no indication that such modified o3o3 or o3o5 would not work. Currently there exists no investigation into the accuracy and stability of such schemes.

Our investigations concentrate on rather low polynomial degrees. The LGM (Steppeler and Klemp 2017), SEM3 and o3o3 are to a large part related and these schemes will be described in the following section together.

3. The SEM3 and o3o3 schemes

Often SEM schemes are formulated in terms of Legendre–Gauss–Lobatto polynomials. Here we use a

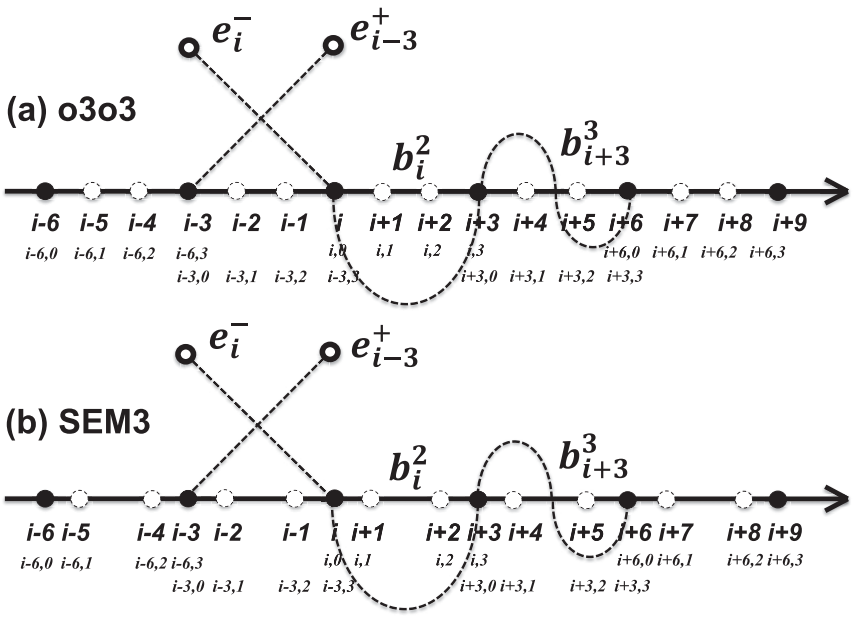


FIG. 1. (a) o3o3 and (b) SEM3 grids.

different formulation in terms of order-preserving polynomials (Steppeler 1976, 1987) which is arithmetically equivalent. We will formulate the representation of field functions $h(x)$ as piecewise cubic polynomials. This third-order polynomial representation of fields is the same for FEM3, SEM3, and o3o3. The cell interval (x_i, x_{i+3}) has a length of $3dx$. We define the center of this cell as x_i^m and introduce the local coordinate $x' = x - x_i^m$ ($i = 0, 3, 6, 9, \ldots$). Then, the cell is described as

$$x' \in (-dx', dx'), \tag{1}$$

with $dx' = 3/2dx$, $x_i^m = x_i + dx' = x_i + 3/2dx$.

Within each cell, the field $h(x)$ is defined as a third-order polynomial and this is determined by four gridpoint values. Therefore, the collocation grid x^c has 4 points $x_{i,i'}^c$ ($i = 0, 3, 6, 9, \ldots, i' = 0, 1, 2, 3$) in each cell interval. Here x_i and x_{i+3} are the cell boundaries:

$$\begin{cases} x_{i,0}^c = x_i = x_i^m - dx', \\ x_{i,3}^c = x_{i+3} = x_i^m + dx'. \end{cases} \tag{2}$$

The definition in Eq. (2) is valid both for SEM3 and o3o3. Obviously some of the collocation points are identical:

$$x_{i,3}^c = x_{i+3,0}^c. \tag{3}$$

The number of collocation points is identical to that of grid points x_i belonging to the interval (x_i, x_{i+3}). For the points $x_{i,1}^c$ and $x_{i,2}^c$ in the interior of the cell interval $(-dx', dx')$, the definitions differ for SEM3 and o3o3. For SEM3 the collocation points are defined as Legendre–Gauss–Lobatto points:

$$\begin{cases} x_{i,1}^c = x_{i+1} = x_i^m - \dfrac{1}{\sqrt{5}} dx', \\ x_{i,2}^c = x_{i+2} = x_i^m + \dfrac{1}{\sqrt{5}} dx'. \end{cases} \tag{4}$$

For o3o3 we use a regularly spaced collocation grid:

$$\begin{cases} x_{i,1}^c = x_{i+1} = x_i^m - \dfrac{1}{3} dx', \\ x_{i,2}^c = x_{i+2} = x_i^m + \dfrac{1}{3} dx'. \end{cases} \tag{5}$$

The grid and the collocation points are illustrated in Fig. 1. The gridpoint space $h_{i,i'} = h(x_{i,i'})$ ($i = 0, 3, 6, 9, \ldots, i' = 0, 1, 2, 3$) may be used to perform the time stepping.

There exists a one-to-one mapping between the gridpoint representation and spectral representation of h. The spectral representation is identical for SEM3 and o3o3. The corner amplitudes $h_{i,0}$ and $h_{i,3}$ are both grid amplitudes and spectral amplitudes and the corresponding basis functions are denoted as $e_{i,0}^+(x')$ and $e_{i+3,3}^-(x')$. Even though $x_{i+3,0} = x_{i,3}$, $h_{i+3,0}$ may be different from $h_{i,3}$ for discontinuous functions. For continuous functions we have $h_{i+3,0} = h_{i,3}$. As $i' = 0$ occurs only with upper index + and $i' = 3$ only with upper index −, the second lower index may be dropped and

the basis be denoted as $e_i^+(x')$ and $e_{i+3}^-(x')$. The two linear basis functions are defined for:

$$\begin{cases} e_i^+(x') = \dfrac{1}{2} + \dfrac{x'}{2dx'}, \\ e_i^-(x') = \dfrac{1}{2} - \dfrac{x'}{2dx'}, \end{cases} \tag{6}$$

where $x' \in (-dx', dx')$. For $x' \notin (-dx', dx')$, we define $e_i^+(x') = e_i^-(x') = 0$.

For the collocation grid Eq. (4), there are also weights $w_0, w_1, w_2, w_3 = (1/6)dx', (5/6)dx', (5/6)dx', (1/6)dx'$ defined. The weights can be used to form integrals of the fields, such as for the computation of mass. o3o3 does not use gridpoint formula such as Legendre–Gauss–Lobatto to compute integrals. Integrals are rather computed in spectral space by integrating the polynomials analytically.

For both SEM3 and o3o3, we define the basis function representation

$$h^1(x') = \sum_{i=0,3,6,\ldots} [h_i e_i^-(x') + h_{i+3} e_i^+(x')], \tag{7}$$

as the linear or first-order part of h where $e_i^-(x')$ and $e_i^+(x')$ may be called interval-related basis functions, as they are nonzero in one interval only. These can be used to describe functions with discontinuities at corner points, such as occur in this paper with derivatives. For continuous functions as in Eq. (7) we can introduce the full hat function $e_i(x')$:

$$e_i(x') = e_i^-(x') + e_{i-3}^+(x'), \tag{8}$$

and obtain the representation equivalent to Eq. (7):

$$h^1(x') = \sum_{i=0,3,6,\ldots} h_i e_i(x'). \tag{9}$$

The high-order part $h^h(x')$ is defined as

$$h^h(x') = h(x') - h^1(x'). \tag{10}$$

For a basis function representation of $h^h(x')$ a set of basis functions with support on one-grid interval only is defined as

$$\begin{cases} b_{i+(3/2)}^2(x') = (x'^2 - dx'^2)/2, \\ b_{i+(3/2)}^3(x') = (x'^3 - x' \times dx'^2)/6. \end{cases} \tag{11}$$

The amplitudes for the high-order part of $h(x')$ are second- and third-derivatives of h at the centers of grid intervals and we obtain the following basis function representation of $h^h(x)$:

$$h^h(x') = \sum_{i=0,3,6,\ldots} [h_{xx,i+(3/2)} b_{i+(3/2)}^2(x') + h_{xxx,i+(3/2)} b_{i+(3/2)}^3(x')]. \tag{12}$$

Equations (7)–(12) define a basis function representation of $h(x')$. The basis function representation is used in the same way for SEM3 and o3o3. The polynomial basis $e_i^-(x'), e_{i-3}^+(x'), b_{i+3/2}^2(x')$, and $b_{i+3/2}^3(x')$ for each grid interval, is also used in FEM. FEM uses the same spectral space as SEM3 and o3o3. The procedure described here is arithmetically equivalent to the use of Legendre polynomials which are often used with SEM3. The basis functions are different, but the function space is the same. The amplitudes h_i, h_{i+1}, and h_{i+2} form the gridpoint space for both SEM3 and o3o3. It is noted that h_{i+1}, h_{i+2} is defined differently for SEM3 and o3o3.

For SEM3 the definition is

$$\begin{cases} h_{i+1} = h(x_{i,1}^c), \\ h_{i+2} = h(x_{i,2}^c), \end{cases} \tag{13}$$

with $x_{i,i'}^c$ defined in Eq. (4) ($i = 0, 3, 6, \ldots, i' = 0, 1, 2, 3$). For o3o3 the definition of the corner modes is also given in Eq. (13), but $x_{i,i'}^c$ defined in Eq. (5) ($i = 0, 3, 6, \ldots, i' = 0, 1, 2, 3$).

The spectral space is the same for FEM3, SEM3, and o3o3. It is formed by the amplitudes h_i, $h_{xx,i+(3/2)}$ and $h_{xxx,i+(3/2)}$ ($i = 0, 3, 6, \ldots$). Because collocation points are different between these methods, the transformation formula between gridpoint space and spectral space are different for SEM3 and o3o3. The formulas have the same form but use different coefficients. As the values h_i ($i = 0, 3, 6, \ldots$) are gridpoint values and spectral coefficients at the same time, we need transformation formula for $i + 1$ and $i + 2$ only.

We define the linear part of h as

$$h_{i+i'}^{\text{lin}} = h_i e_i^-(x_{i,i'}^c) + h_{i+3} e_i^+(x_{i,i'}^c), \quad i' = 1, 2, \quad i = 0, 3, 6, \ldots, \tag{14}$$

with $x_{i,i'}^c$ defined in Eq. (4) for SEM3 and $x_{i,i'}^c$ defined in Eq. (5) for o3o3 ($i' = 1, 2, i = 0, 3, 6, \ldots$).

Define the high-order difference field $h_{i+i'}^{\text{diff}}$ as

$$h_{i+i'}^{\text{diff}} = h_{i+i'} - h_{i+i'}^{\text{lin}}. \tag{15}$$

Then, considering the definition of the high-order basis functions $b_{i+(3/2)}^2(x')$, $b_{i+(3/2)}^3(x')$ in Eq. (11), we obtain for the amplitudes $h_{xx,i+(3/2)}$ and $h_{xxx,i+(3/2)}$ used in Eq. (12):

$$\begin{cases} h_{xx,i+(3/2)} = (h_{i+1}^{\text{diff}} + h_{i+2}^{\text{diff}})/[2b_{i+(3/2)}^2(x_{i,1}^c)], \\ h_{xxx,i+(3/2)} = -(h_{i+2}^{\text{diff}} - h_{i+1}^{\text{diff}})/[2b_{i+(3/2)}^3(x_{i,1}^c)], \end{cases} \tag{16}$$

where $x_{i,i'}^c$ is defined in Eq. (4) for SEM3 and in Eq. (5) for o3o3. The transformation formula can be used in the same way for SEM3 and o3o3, but the collocation points $x_{i,1}^c$ and $x_{i,2}^c$ are defined differently for these two methods.

While Eqs. (14)–(16) give the transformation from gridpoint space to spectral space, the transformation from spectral space to gridpoint space can be done using Eqs. (7)–(12), when the spectral coefficients h_i, $h_{xx,i+(3/2)}$, and $h_{xxx,i+(3/2)}$ are given.

In the test problems below, we use homogeneous advection of the field with velocity $u_0 = 1$:

$$h_t(x) = -u_0 h_x(x) = \mathrm{fl}_x \tag{17}$$

with $\mathrm{fl} = -u_0 h$ and $u_0 = 1$, where in Eq. (17) the lower indices t or x mean differetiation to time or space.

The time stepping with a time step dt is done using the fourth-order Runge–Kutta (RK4) time scheme, as described by Steppeler et al. (2008). The Runge–Kutta scheme can be performed if the spatial derivative of the flux $\mathrm{fl}(x)$ is approximated according to Eqs. (7)–(12) in the same space as $h(x)$.

In this study, we have $\mathrm{fl} = -u_0 h(x)$. The approximated form of $\mathrm{fl}(x)$ in the grid is obtained analogously to Eqs. (7)–(12) and the transformation from gridpoint space to spectral space and vice verse are done in the same way as for $h(x)$. The differentiation of $\mathrm{fl}(x)$ can be done directly in the basis function representation Eqs. (7)–(12):

$$\begin{aligned}\mathrm{fl}(x') = -u_0 \sum_{i=0,3,6,\ldots} [&h_i e_i^-(x') + h_{i+3} e_i^+(x') \\ &+ h_{xx,i+(3/2)} b^2_{i+(3/2)}(x') \\ &+ h_{xxx,i+(3/2)} b^3_{i+(3/2)}(x')],\end{aligned} \tag{18}$$

and

$$\begin{aligned}h_t(x') = \mathrm{fl}_x(x') = -u_0 \sum_{i=0,3,6,\ldots} [&h_i e^-_{x,i}(x') + h_{i+3} e^+_{x,i}(x') \\ &+ h_{xx,i+(3/2)} b^2_{x,i+(3/2)}(x') \\ &+ h_{xxx,i+(3/2)} b^3_{x,i+(3/2)}(x')].\end{aligned} \tag{19}$$

For the RK4 time integration method, we need to obtain the x-derivative at the collocation points $x^c_{i,i'}$, ($i' = 1, 2$). This can be done by inserting $x^c_{i,i'}$ into Eq. (19). For the inner points $x^c_{i,1}$ and $x^c_{i,2}$ of a cell interval (x_i, x_{i+3}), this gives an unique value, as the representation Eq. (18) of $\mathrm{fl}(x)$ is differentiable at these points.

$$\begin{aligned}\mathrm{fl}_x(x^c_{i,i'}) = -\sum_{i=0,3,6,\ldots} [&h_i e^-_{x,i}(x^c_{i,i'}) + h_{i+3} e^+_{x,i}(x^c_{i,i'}) \\ &+ h_{xx,i+(3/2)} b^2_{x,i+(3/2)}(x^c_{i,i'}) \\ &+ h_{xxx,i+(3/2)} b^3_{x,i+(3/2)}(x^c_{i,i'})]\end{aligned} \tag{20}$$

for $i' = 1, 2$. As h_i, h_{i+3}, $h_{xx,i+(3/2)}$, $h_{xxx,i+(3/2)}$ are polynomials in $x^c_{i,i'}$, Eq. (20) is a scheme of finite-difference method (FDM) for $h_t(x^c_{i,i'})$.

For the corner grid points $i' = 0$, $\mathrm{fl}(x)$ is not differentiable and we get at these points double values $\mathrm{fl}^+_x(x^c_{i,0})$ and $\mathrm{fl}^-_x(x^c_{i,0})$, representing differentiation from right and left, given by

$$\begin{cases}\mathrm{fl}^-_x(x^c_{i,0}) = -\sum_{i=0,3,6,\ldots} [h_{i-3} e^-_{x,i-3}(+dx') + h_i e^+_{x,i-3}(+dx') + h_{xx,i-(3/2)} b^2_{x,i-(3/2)}(+dx') + h_{xxx,i-(3/2)} b^3_{x,i-(3/2)}(+dx')], \\ \mathrm{fl}^+_x(x^c_{i,0}) = -\sum_{i=0,3,6,\ldots} [h_i e^-_{x,i}(-dx') + h_{i+3} e^+_{x,i}(-dx') + h_{xx,i+(3/2)} b^2_{x,i+(3/2)}(-dx') + h_{xxx,i+(3/2)} b^3_{x,i+(3/2)}(-dx')].\end{cases} \tag{21}$$

With SEM3, the derivative at $x^c_{i,0}$ is approximated as

$$\mathrm{fl}_x(x^c_{i,0}) = \frac{1}{2}[\mathrm{fl}^-_x(x^c_{i,0}) + \mathrm{fl}^+_x(x^c_{i,0})]. \tag{22}$$

The computation of the spatial derivative of $\mathrm{fl}(x)$ given in Eqs. (20)–(22) is used for SEM3. The discontinuous representation Eq. (21) is mass conserving by construction. From the Legendre–Gauss–Lobatto integration formula, it follows that Eqs. (20)–(22) is also mass conserving. Equation (21) means that the construction of the high-order part of h_t is done in spectral space and transformed in gridpoint space. Other choices with SEM3 for the collocation points can violate mass conservation. Instability is possible with explicit time integration methods such as leapfrog or RK4 schemes (however, see Ullrich (2014b) for stable alternatives).

Equations (20)–(22) give the spatial derivative $\mathrm{fl}_x(x)$ of $\mathrm{fl}(x)$ at all collocation points and the RK4 time step for SEM3 can therefore be performed in the collocation gridpoint space. Equations (20)–(22) are FDM equations in gridpoint space with the unusual feature that the FDM equations given in Eqs. (20)–(22) are different for each of the grid points i, $i + 1$, $i + 2$. Such difference schemes using more than one difference equation for different grid points are called inhomogeneous FDM schemes (Chen and Xiao 2008; Ullrich 2014b). Therefore, SEM3 can be written as an inhomogeneous FDM schemes, while schemes such as centered differences or classical fourth-order spatial differences (Steppeler et al. 2008) are homogeneous FDM schemes.

o3o3 in the version discussed here also computes the time derivatives of $h(x)$ at the collocation points,

which are equally spaced [see Eq. (5)] in the case of o3o3. The RK4 time step can then be performed in gridpoint space. Equations (20)–(22) will not be used for computing spatial derivatives. The time derivatives for the high-order part $h^h(x)$ will rather be computed in spectral space and then transformed to gridpoint space. The discontinuous form Eq. (22) of $h_t(x)$ will not be used with o3o3 to compute derivatives.

As with Eq. (19) we assume that for the time derivative $h_t(x)$ of $h(x)$, a third-order basis function representation with spectral amplitudes for the time derivatives $h_{t,i}$, $h_{t,xx,i+(3/2)}$, $h_{t,xxx,i+(3/2)}$ is:

$$h_t(x') = \sum_{i=0,3,6,\ldots} [h_{t,i} e_i^-(x') + h_{t,i+3} e_i^+(x') + h_{t,xx,i+(3/2)} b^2_{i+(3/2)}(x') + h_{t,xxx,i+(3/2)} b^3_{i+(3/2)}(x')]. \quad (23)$$

From Eq. (23) an expression for the time derivative of mass follows:

$$M_t = \int dx h_t(x) = \sum_{i=0,3,6,\ldots} \left\{ \frac{3}{2} h_{t,i} dx + \frac{3}{2} h_{t,i+3} dx + \frac{1}{2} h_{t,xx,i+(3/2)} \left[\frac{2}{3} (dx')^3 - 2(dx')^3 \right] \right\} = \sum_{i=0,3,6,\ldots} \left(h_{t,i} dx' + h_{t,i+3} dx' - \frac{2}{3} h_{t,xx,i+(3/2)} dx'^3 \right). \quad (24)$$

Equation (24) follows from Eq. (23) using

$$\begin{cases} \int dx e^+(x) = \int dx e^-(x) = dx', \\ \int dx b^2(x) = -\frac{2}{3} dx'^3, \\ \int dx b^3(x) = 0. \end{cases} \quad (25)$$

Equation (25) follows from Eqs. (6) and (11). The last of the equations in Eq. (25) follows from the symmetry of $b^3(x)$, defined in Eq. (11).

It is possible to define $h_t(x)$ by defining the grid amplitudes $h_{t,i}$, $h_{t,xx,i+(3/2)}$, and $h_{t,xxx,i+(3/2)}$ ($i = 0, 3\ 6, 9, \ldots$) as an inhomogeneous FDM scheme: for the corner amplitude we define

$$h_{t,i} = -u_0 h_{x,i} = -u_0 \left[-\frac{1}{3} \frac{(h_{i+2} - h_{i-2})}{4dx} + \frac{4}{3} \frac{(h_{i+1} - h_{i-1})}{2dx} \right], \quad i = 0, 3, 6, \ldots, \quad (26)$$

which is the classical fourth-order FDM scheme for regular resolution. It exists also for irregular resolution, which was applied by Steppeler et al. (2008).

For the derivation of the amplitude $h_{t,xx,i+(3/2)}$, the time derivative $dM_{t,i}$ of the mass contained in the cell interval (x_i, x_{i+3}) ($i = 0, 3, 6, 9$) is considered. According to Eq. (26), it is determined by the fluxes at the cell boundaries x_i and x_{i+3}:

$$dM_{t,i} = \int_{x_i}^{x_{i+3}} dx h_t(x) = -u_0 (h_{i+3} - h_i). \quad (27)$$

From Eq. (24) we obtain:

$$dM_{t,i} = h_{t,i} dx' + h_{t,i+3} dx' - \frac{2}{3} h_{t,xx,i+(3/2)} dx'^3, \quad i = 0, 3, 6, \ldots \quad (28)$$

From Eqs. (27) and (28) we obtain for the amplitude $h_{t,xx,i+(3/2)}$:

$$h_{t,xx,i+(3/2)} = u_0 \frac{3(h_{i+3} - h_i)}{2dx'^3} + \frac{3(h_{t,i} + h_{t,i+3})}{2dx'^2}. \quad (29)$$

In Eq. (29), h_t can be any third-order approximation for the time derivative of h at the corner points. Here we use Eq. (26) for the standard o3o3 scheme and Eq. (22) for o3o3 with spectral derivative at corner points. Using Eq. (29), we obtain $h_{t,xxx,i+(3/2)}$ by FDM:

$$h_{t,xxx,i+(3/2)} = \frac{h_{t,xx,i+(9/2)} - h_{t,xx,i-(3/2)}}{6dx}. \quad (30)$$

Equation (30) uses a rather large stencil, including five grid intervals. The particular approximations used in this study are made for easy implementation and not for optimal performance. Optimal performance is not explored here, as the intention is only to give an example for alternative continuous Galerkin schemes. Inside a grid interval, the collocation points in this formulation are always regular. However, the grid sizes are allowed to be irregular to accommodate to the irregular grid structures which may occur with approximations on the sphere. Equation (26) would be valid for irregular grids as it is. The FDM schemes Eqs. (22) and (29) are formulated for regular grids. These formulas for irregular meshes must be replaced by the differentiation in the irregular grid.

4. Performance of a time step

The time step procedure for the field $h(x)$ is done in gridpoint space, using the formulas for flux divergence derived in section 3. The gridpoint space to be used is not necessarily the regular grid x_i ($i = 0, 1, 2, 3, \ldots$) but the $x^c_{i,i'}$ ($i = 0, 3, 6, \ldots, i' = 0, 1, 2, 3$). It is defined in Eqs. (2) and (4) for SEM3 and in Eqs. (2) and (5) for o3o3.

RK4 requires the computation of the spatial derivatives $h_x(x)$. The collocation grid is redundant for the description of continuous fields [see Eq. (3)], as corner points occur twice. In the collocation grid, discontinuous fields can be represented, which in this paper occur as derivatives. In the redundant collocation grid, according to field values in Eq. (20), $h_i^+ = h(x^c_{i+3,0})$ and $h_i^- = h(x^c_{i,3})$ left and right of a corner point may be assigned differently. For continuous fields such as $h(x)$, the basis function representation of $h(x)$ is obtained by assigning the same amplitude $h(x_i)$ ($i = 0, 3, 6, 9, \ldots$) to $h_i^+ = h(x^c_{i+3,0})$ and $h_i^- = h(x^c_{i,3})$.

There is a nonredundant version of the collocation grid, called the finite-difference grid, defined as

$$x^{\mathrm{fd}}_{i+i'} = x^c_{i,i'} \tag{31}$$

for $i = 0, 3, 6, 9, \ldots, i' = 0, 1, 2, i + i' = 0, 1, 2, 3, 4, \ldots$

The finite-difference grids for SEM3 and o3o3 are shown in Fig. 1 together with the basis functions $e_i^+(x)$, $e_i^-(x)$, $b^2_{i+(3/2)}(x)$, and $b^3_{i+(3/2)}(x)$ in Fig. 1. For o3o3 the finite-difference grid is the regular grid:

$$x^{\mathrm{fd}}_i = x_i. \tag{32}$$

And gridpoint values are assigned:

$$h^{\mathrm{fd}}_i = h(x^{\mathrm{fd}}_i), \qquad i = 0, 1, 2, 3, 4, \ldots \tag{33}$$

To perform an RK4 time step, a transformation into spectral space is done obtaining the spectral amplitudes $h_i, h_{xx,i+(3/2)}, h_{xxx,i+(3/2)}$ ($i = 0, 3, 6, 9, \ldots$) from Eqs. (13)–(16) for both SEM3 and o3o3. These coefficients are used to compute the values $h_{x,i} = h_x(x^{\mathrm{fd}}_i)$ of $h_x(x)$ in the finite-difference grid. For SEM3 this is done using Eqs. (20) and (22). For o3o3, $h_t(x)$ is first computed in spectral space and then transformed to the finite-difference grid. In spectral space, the amplitudes for $h_t(x)$ are $h^{\mathrm{fd}}_{t,i}$, $h^{\mathrm{fd}}_{t,xx,i+(3/2)}$, $h^{\mathrm{fd}}_{t,xxx,i+(3/2)}$ ($i = 0, 3, 6, \ldots$), which are computed from Eqs. (26), (29), and (30). Equation (23) is then used to compute the gridpoint values $h^{\mathrm{fd}}_{t,i} = h^{\mathrm{fd}}_t(x^{\mathrm{fd}}_i)$ ($i = 0, 1, 2, 3, 4, \ldots$). With SEM3 and o3o3, the RK4 time step is then performed using the gridpoint values of $h_t(x)$ in the x^{fd}_i grid.

5. Results

The test example is the homogeneous 1D advection equation [Eq. (17)]. We use an interval of 600 points with $dx = 1$ and periodic boundary conditions. The initial conditions are given by

$$h_i(x^{\mathrm{fd}}_i) = 4 \exp\left[-\frac{(x^{\mathrm{fd}}_i - 150)^2}{f^{\mathrm{length}}}\right], \tag{34}$$

which is a 1D version of the test used by Steppeler and Klemp (2017). We also use the peak solution defined as $h_{150} = 4$, $h_{149} = h_{151} = 8/3$, $h_{148} = h_{152} = 4/3$, and $h_i = 0$ for all other i in the experiments.

To get an indication of stability, the solution is shown for an extended time, here for the advection over $30\,000dx$. For a given time step dt, the number of steps performed is

$$n^{\mathrm{time}} = \frac{le^{\mathrm{space}}}{dt} \tag{35}$$

where $le^{\mathrm{space}} = 300$ or $le^{\mathrm{space}} = 30\,000dx$. For the experiments, dt was chosen to be 1.

Figure 2 shows the initial values representing the peak solution, as defined above over a distance of $300dx$ and $30\,000dx$ for standard nonconserving o4 spatial differences (Figs. 2b,c). The transport over the larger distance of $30\,000dx$ shows an increased dispersion error. Figures 2d–i are the same result for standard o3o3, o3o3 with spectral differences at corner points and SEM3. Figure 3 is the same as Fig. 2 for the smooth solution with $f^{\mathrm{length}} = 4$.

For the transport over $30\,000dx$ with $dt = 1$, the maximum of $h(x)$ is given in Table 1 for the four schemes investigated. SEM3 is the best by a small margin and the two versions of o3o3 are rather similar in accuracy to the classical o4 scheme. As a quantitative measure of accuracy for the transport over the distance of $30\,000dx$ the forecasted maximum of $h(x)$ divided by four with solutions using $dt = 1$, $dx = 1$ is shown in Table 1. The higher the maximum is, the more accurate the scheme is. The accuracies for the four schemes are rather similar. For o4 and o3o3, the results are also shown for the larger time steps $dt = 2$ and $dt = 2.5$. The rather strong dependence of the results on the time step is remarkable. The spread of the results with dt is such that we cannot conclude more, than that the fourth-order schemes are similar in accuracy.

SEM3 and the two versions of o3o3 presented here have the advantage over standard o4 differences that they are mass-conserving schemes. The standard o4 scheme was used in a realistic model by Kalnay et al. (1977) and the lack of conservation was considered to be a disadvantage of that model.

To see the dependence of the accuracy on the resolution, the solution of Figs. 2d–g (o3o3 with $f^{\mathrm{length}} = 4$,

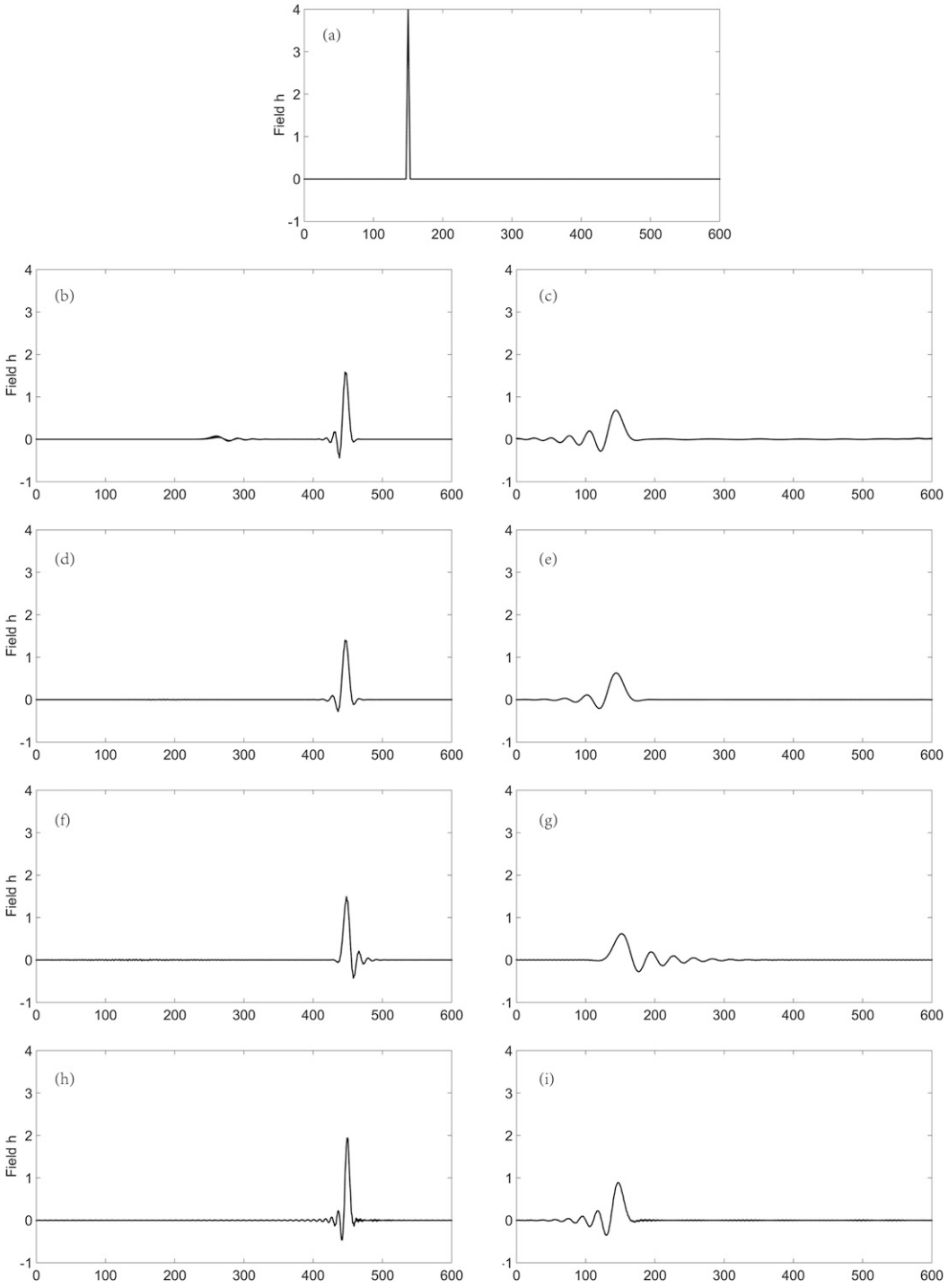


FIG. 2. Runs with the peak initial condition and transport over 300 points with different spacing for $dx = 1$ and $dt = 1$: (a) initial values, (b) standard o4 spatial difference, (d) o3o3 standard difference, (f) o3o3 spectral difference, and (h) SEM3; transport over 30 000 points: (c) standard o4 spatial difference, (e) o3o3 standard difference, (g) o3o3 spectral difference, and (i) SEM3 at all points.

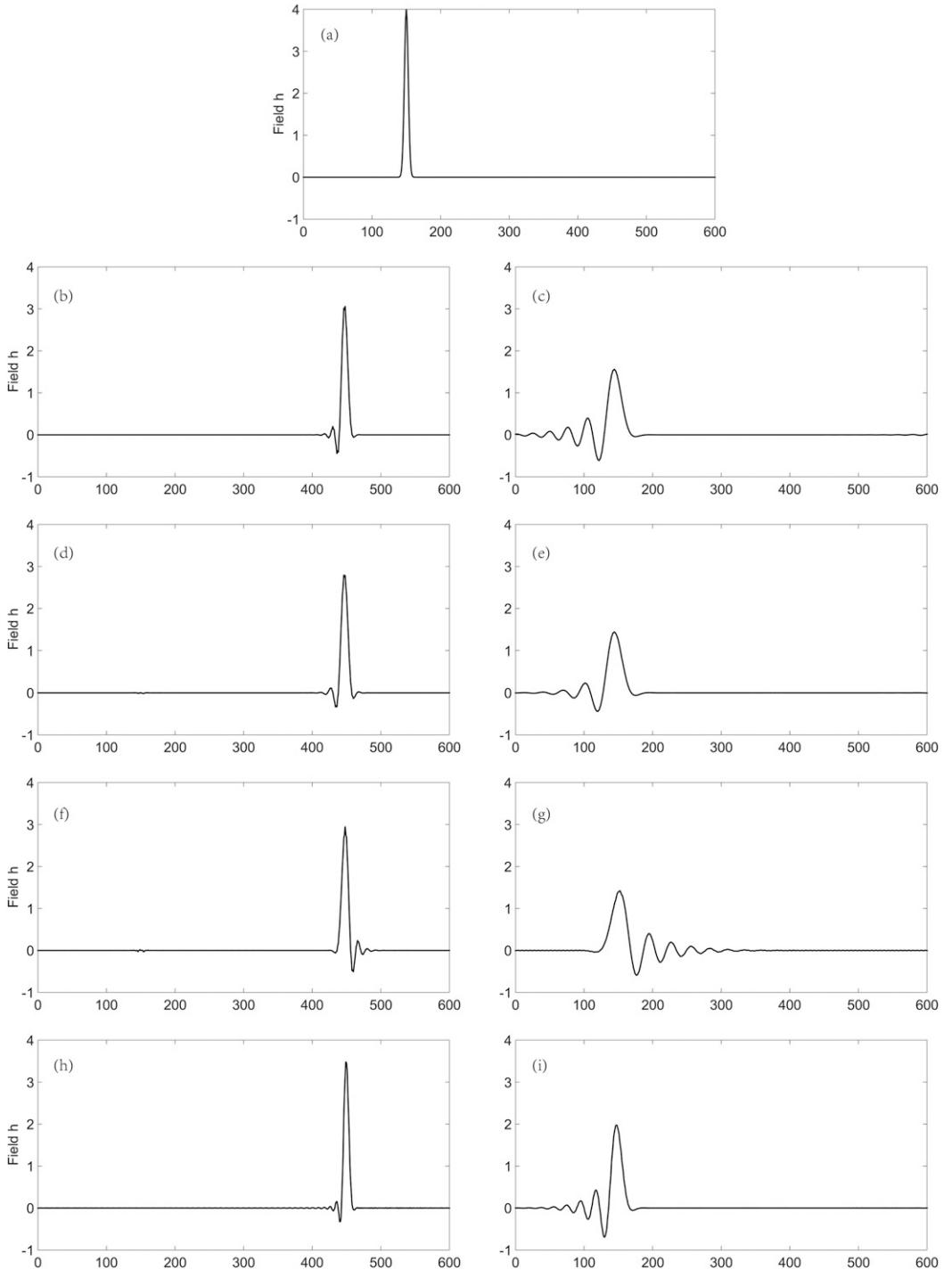


FIG. 3. As in Fig. 2, but for a smooth solution.

$dt = 1, dx = 1$) is repeated for a forecast time of 600 with $dx = 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 1$, $dt = 1/16$ and the results are shown in Fig. 4 for both the o3o3 method and the o3o3 method with spectral differentiation. The accuracy of both schemes does not go beyond 5×10^{-3}. This is caused by the presence of a computational mode within this bell shaped solution. By plotting in a smaller scale (not shown here), this can be shown to be a small-scale wave of this amplitude and this is connected to the presence of the dispersion properties which will be

TABLE 1. Maximum of field $h(x)$ after a transport over 30 000dx. The higher the value, the more accurate is a scheme.

Schemes	Standard o4		o3o3		o3o3 with spectral difference	SEM3
Time step (s)	1.0	2.0	1.0	2.5	1.0	1.0
Peak solution	0.135	0.146	0.100	0.139	0.079	0.209
Smooth solution with 4	0.279	0.313	0.214	0.309	0.149	0.402
Smooth solution with 8	0.364	0.488	0.311	0.501	0.208	0.419

investigated later. The convergence experiment was repeated using one sine wave over 192 points. Such a pure wave has no component of a computational mode. It was shown, that for the prediction of one wave cycle the convergence was fourth-order and the accuracy went down to 2×10^{-5} for the resolution $dx = 1$ and wavelength 192 (not shown here).

The time step and dispersion are investigated by using spectral solutions. For $j = 0, 1, 2, \ldots$, we assume

$$h_j = h_0 e^{ik(j \times dx - ct)}, \tag{36}$$

where c is the phase velocity. Then, we assume amplitudes $\mathbf{A} = (h_j, h_{xx,j}, h_{xxx,j})$ in spectral space. For each wavenumber k, we have a linear relation between $\mathbf{A}_k$ and $\mathbf{A}_{t,k}$, where $\mathbf{A}_{t,k}$ is the time derivative of $\mathbf{A}_k$:

$$\mathbf{A}_{t,k} = \mathbf{M}^k \mathbf{A}_k. \tag{37}$$

The matrix $\mathbf{M}^k$ depends on the dimensionless wavenumber k (see the definition in appendix). The eigenvalues of $\mathbf{M}^k$ are defined as e_1^k, e_2^k, e_3^k. The imaginary parts of $e_j^k (j = 1, 2, 3)$ indicate the phase velocity whereas a negative real part is indicative of implicit diffusion. The exact solution should have a linear dependence on k. A negative derivative to k of the phase velocity means negative group velocities which is inherently unphysical and should likely be filtered. The phase velocities in dependence of kdx are shown in Fig. 5 for the two versions of o3o3. This plot is generated by appending the eigenvalues e_1^k, e_2^k, e_3^k as a function of k in sequence (see e.g., Ullrich et al. 2018).

For o3o3, a large part of the spectrum is stationary (phase velocity is 0). These stationary modes arise when the discrete fourth derivative at corner points evaluates to zero and $h_{i+3} = h_i$, in which Eqs. (29) and (30) evaluate to zero. For another part of the spectrum, the phase velocity is negative. While it clearly indicates that there is room for improvements, the comparison with the two control runs o4 and SEM3 shows that o3o3 is competitive in accuracy.

Ullrich (2014a) defines a measure of the effective resolution, indicating the smallest scale of a method to obtain useful forecasts. This is the smallest wave for which the absolute error of phase velocity is less than 1%. The effective resolution comes out to be 7.5dx for o3o3, 7.9dx for o4, and 8.4dx for SEM3. The large value for SEM3 is caused by the spectral gap (Ullrich et al. 2018). Notably, o3o3 and o4 do not have a spectral gap. For SEM3 the solution can be improved by hyperviscosity. The impact of hyper-viscosity is not investigated for o3o3. All four schemes investigated are neutral, in having no implicit diffusion.

The largest imaginary eigenvalue for o3o3 is 3.29 (dx = 1/3), thus resulting in a CFL condition of 2.53 under the RK4 scheme. A table of CFL condition for the methods investigates here is given in Table 2. For comparison, it is mentioned that RK4 with centered differences in space has a CFL condition of 2.8. The CFL condition of o3o3 with spectral differences at corner points turned out to be higher than standard o4 differences by 0.2. For standard o3o3, the CFL condition was higher than standard o4 by 0.5, while for SEM3 it was lower by 0.5. Standard o3o3 had a CFL condition higher than SEM3 by a factor of 1.67, which resulted in shorter runtime of the program.

Using the measure of effective resolution, the accuracy of o3o3 is similar to the other third- or fourth-order

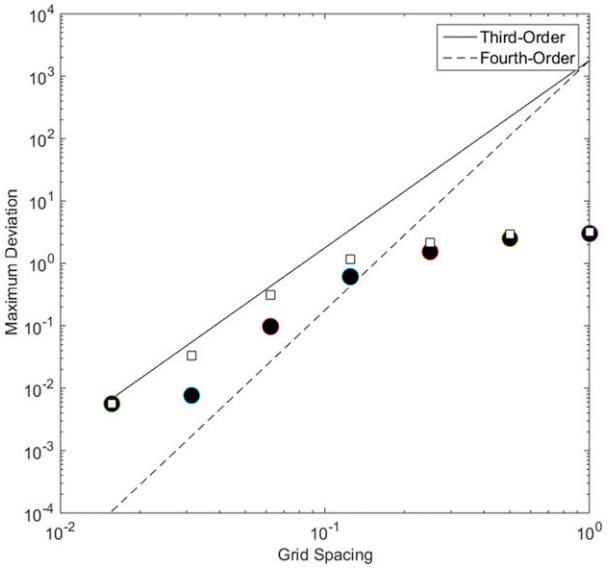


FIG. 4. Maximum error of o3o3 standard (dotted scatters) and o3o3 spectral differences (square scatters) for different grid spacings: dx = 1/64, 1/32, 1/16, 1/8, 1/4, 1/2, 1.

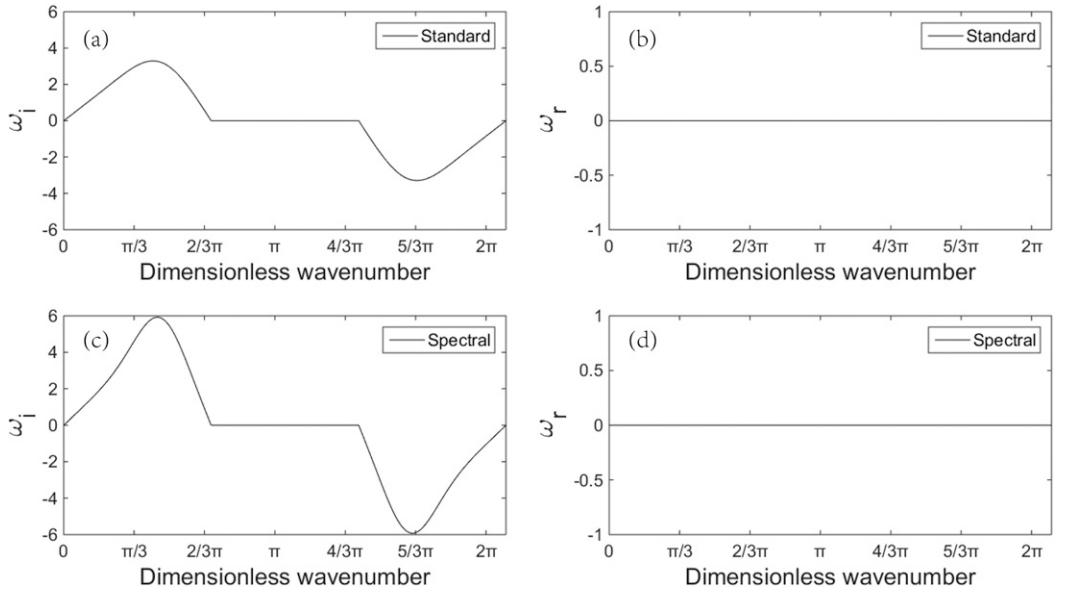


FIG. 5. Spectral diagrams of (left) imaginary and (right) real parts of eigenvalues dependent on k, where o3o3 standard is in the first row and o3o3 with spectral differences at corner points is in the second row.

schemes used for the control run. SEM3 and o3o3 are mass conserving. An advantage of o3o3 over SEM3 is the regular collocation grid and the larger time step. For 1D there are small differences of the cost to produce a time step. In 2D, o3o3 has a considerable advantage, as the sparse grid used in higher-dimensional spaces is potentially much more economical than the full grid (see details in section 6).

6. Practical considerations for 2D and 3D

The practical usefulness of o3o3 or any other numerical scheme depends very much on factors like the number of operations necessary per grid point, suitability for scaling well on multiprocessing computers and the data volume per grid point which needs to be communicated to other points. In this section, we only give a preliminary answer, as the mentioned factors depend very much on the implementation of a scheme and can realistically be investigated only by doing systematic experiments on different computer architectures. Such investigations in the field of informatics are beyond the scope of this study. For 1D calculations and a rather coarse estimate, we do not consider the small differences in performance between the schemes SEM3, o3o3, and o4. These methods should be rather similar in performance per grid point. Though SEM3 and o3o3 are simpler in the performance of a time step than o4, SEM3 and o3o3 need an extra (small) spectral transformation compared to o4, which is not necessary with o4. As a rough estimate, we can expect a similar performance per grid point. In this paper, we do not investigate computational advantages by using larger time steps, which can be chosen somewhat higher with o3o3. However, in 2D, there may be computational advantages for o3o3 by using sparse grids. The use of sparse grids with o3o3 is similar to the situation with FEMs (Ahlberg et al. 1967). With sparse grids, some of the points of a regular 2D or 3D grid are not predicted in time and a corresponding saving of computer time is achieved. Therefore, considerations of numerical efficiency are incomplete without using more than one dimension.

We use tensor function spaces to expand to more than 1D and investigate an example for gaining computational efficiency from sparse grids. Lagrange elements and simplices are not considered in this paper. SEM3 implementations use the full grid $(x_i, z_k) = (i \times dx, k \times dz)$, where $i, k = 1, 2, 3, \ldots$Third-order FEMs use a reduced or sparse grid obtained from the full grid by omitting points when neither i or k are multiples of 3 (Ahlberg et al. 1967). For the 2D case, the full and reduced grids are shown in Fig. 6. All points together shown in Fig. 6a form the full grid. The points shown in

TABLE 2. The CFL condition of RK4 with spatial centered difference is 2.8.

Schemes	Standard o4	o3o3	o3o3 with spectral difference	SEM3
CFL condition with RK4 time stepping	2.0	2.5	2.2	1.5

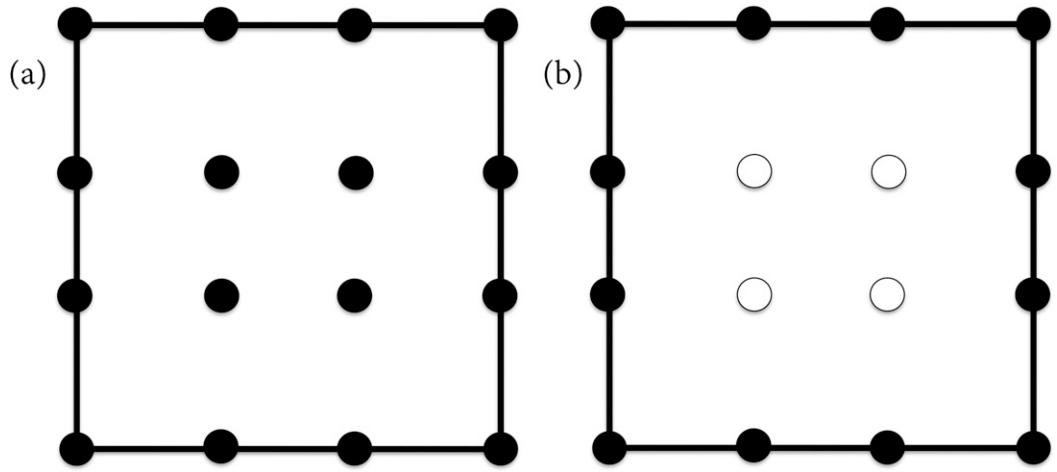


FIG. 6. Computational grid in 2D. (a) full grid and (b) sparse grid where unused points are indicated in white.

Fig. 6b as white are unused for dynamics. The points in black form the sparse grid and are called dynamic points. The fields in the interior of the cell are obtained by bilinear interpolation when plotting (see details in Ahlberg et al. 1967; Steppeler 1987). The sparseness factor, being the ratio of the number of points of the sparse and the full grid is $S = 5/9$ (about 1/2). With the o3o3 scheme most of the terms used for the calculations are done on coordinate lines. Operations defined in 1D can be used on the coordinate line.

The 2D advection equation is defined as:

$$h_t(x,z) = \mathrm{fl}^x_x(x,z) + \mathrm{fl}^z_z(x,z), \tag{38}$$

where $\mathrm{fl}^x_x = -u_0 h_x(x,\, z)$ and $\mathrm{fl}^x_z(x,\, z) = -w_0 h_z(x,\, z)$.

The grid point amplitudes of $h(x,\, z)$ and the fluxes have two indices i, k. As in the 1D case, the corner amplitudes $h_{i,k}$ ($i,\, k = 0,\, 3,\, 6,\, \ldots$) are both gridpoint values and spectral amplitudes and are not transformed. The spectral amplitudes for the 2D case are $h_{xx,i+(3/2),k}$, $h_{xxx,i+(3/2),k}$ and $h_{zz,i,k+(3/2)}$, $h_{zzz,i,k+(3/2)}$. As in the sparse grids all points are on coordinate lines $x = x_\mathrm{i}$ and $z = z_\mathrm{k}$, the spectral amplitudes can be computed by 1D operations along these coordinate lines, as described in sections 3 and 4.

As an example considered, the flux $\mathrm{fl}^x(x,\, z)$ in x direction is divided into two parts:

$$\mathrm{fl}^x(x,z) = \mathrm{fl}^{x,1}(x,z) + \mathrm{fl}^{x,2}(x,z), \tag{39}$$

with $\mathrm{fl}^{x,1}(x,\, z) = \sum_k \mathrm{fl}^x(x,\, z_k)$ and $\mathrm{fl}^{x,2}(x,\, z) = \mathrm{fl}^x(x,\, z) - \mathrm{fl}^{x,1}(x,\, z)$.

According to the serendipity interpolation $\mathrm{fl}^{x,2}(x,\, z)$ is a cubic spline in z direction and linear in x direction. The $\mathrm{fl}^{x,2}(x,\, z)$ is also a small field. When approximating analytic functions, the maximum of $\mathrm{fl}^{x,2}(x,\, z)$ goes to 0 in fourth order. This can be used to design approximations and the motivation of the sparse grid also comes from sparseness. The basis functions to be used with the unused points would be squares of the basis functions used with $\mathrm{fl}^{x,2}(x,\, z)$. This means that the neglections leading to the sparse grid concern small terms.

To compute the time derivative $\mathrm{fl}^{x,1}_t(x,\, z)$ of $\mathrm{fl}^{x,1}(x,\, z)$, we can just solve the 1D problem along the line $z = z_k$. For $\mathrm{fl}^{x,2}_t(x,\, z)$ we have centered differences, which are conserving because the serendipity interpolation is linear in x direction:

$$\mathrm{fl}^{x,2}_t(x_i, z_{k+(3/2)}) = \frac{1}{6dx}[\mathrm{fl}^{x,2}(x_{i+3}, z_{k+(3/2)}) - \mathrm{fl}^{x,2}(x_{i-3}, z_{k+(3/2)})] \tag{40}$$

The treatment of $\mathrm{fl}^{z,2}(x,\, z)$ is done in an analogous way. The unused points are not invoked for forecasting. Their values can be interpolated for plotting.

Figure 7 gives the result of advection for a square grid of 300×300 point grid with $dx = dz = 1.0$, $dt = 1.0$ s. The velocity components u_0 and w_0 are (1, 1) for the first 100 time steps, $(-1, 0)$ for the second 100 time steps and $(0, -1)$ for the last 100 time steps. This velocity field is changed in time to obtain a solution where the initial field is reproduced after 300 time steps. In Fig. 7, 50 time steps are done between plots and the plots are done to show the sparse grid structure. The unused points are assigned a time derivative of 0. Therefore, the initial values at such points remain there at all times. In the area of the forecasted field, the unused points have the value 0. To see the field structure better, blowups at 50th time steps are given at the areas of the initial values and the forecasted fields (Fig. 7c). The forecasted field has holes at places, where

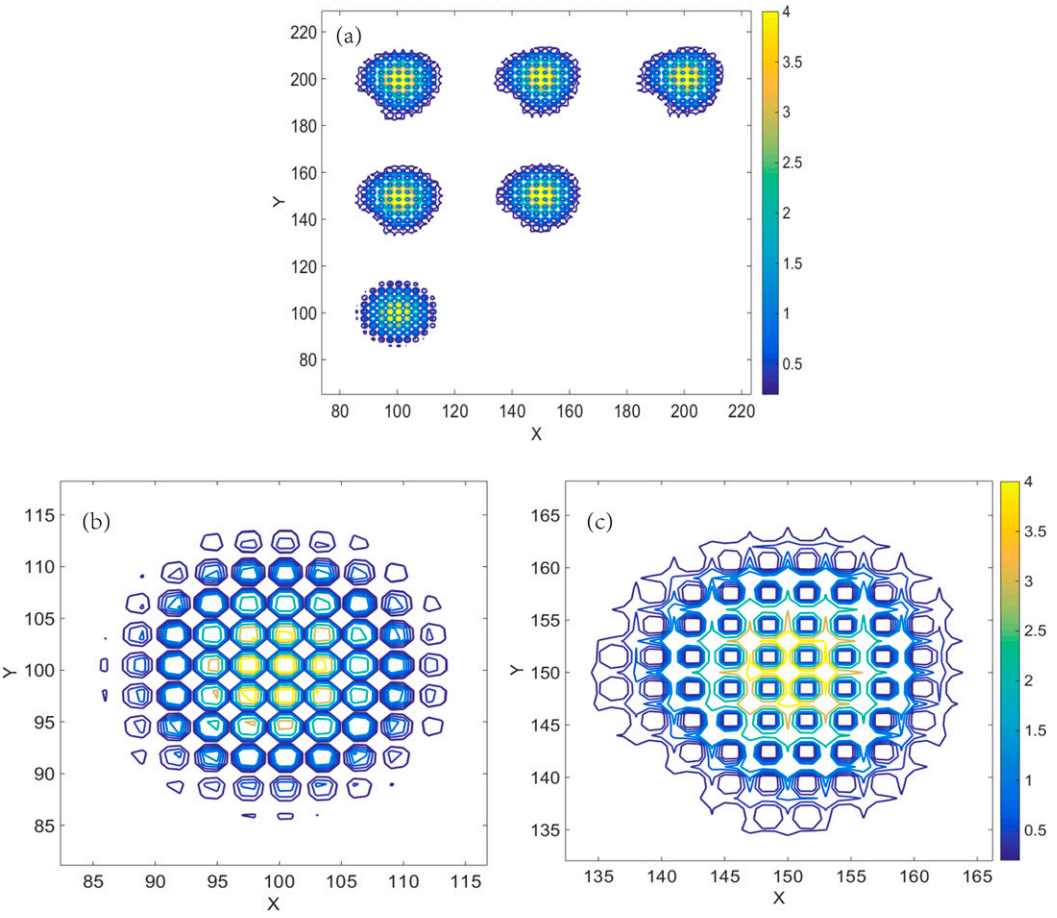


FIG. 7. 2D results: advection with a homogeneous velocity field changing in time. (a) The plots are done to indicate the sparse grid. The unused points are plotted with amplitude 0 for the forecasted fields and at the position of the initial values the unused amplitudes have their original values. The plot gives a blowup of the fields belonging to (b) the initial time and (c) the 50th time step to show that these diagrams are like the positive and negative pictures of the same structure.

there are maxima in the area of the initial conditions. When the forecast arrives at the initial position, the two figures combine to give a smooth field again. This way of plotting is chosen to give a graphical illustration of sparseness. When interpolating to the unused points before plotting, ordinary smooth fields are obtained.

The example included here is to discuss the potential of computer efficiency associated with sparse grids. A substantial gain of computer efficiency may be obtained by the sparse grid, which comes naturally with o3o3. The sparseness factor in 2D is the relation of the number of dynamic points to all points: $S = 5/9$. If the computational cost of o3o3 per grid point is the same as that the same as for a FDM scheme on the full grid, this promises a saving of computer time in 2D by the factor of S. SEM schemes, such as Giraldo (2001) often use SEM only for the horizontal. However, doing the sparse grid in 3D creates a larger potential for saving from sparseness. With o3o3 in 3D, we have $S = 7/27$ (about 1/4) for cubic cells.

For time stepping, we use RK4 and this scheme is not totally uncompetitive (Durran 2012). In practical modeling RK3 is often preferred to RK4 (Klemp et al. 2007). Also, the different arithmetic forms of o3o3 can have an impact on the numerical performance of a model. While for a realistic model in practical use such small differences in performance may be worth investigating, in the present paper we do not investigate this.

It should be mentioned that the two versions of o3o3 presented here are not the only possibilities, but rather examples for alternative LGMs. It is not likely that the schemes presented are optimal. The procedures in this paper were chosen for simplicity of programming and not for being optimal. In Eq. (30), a rather wide stencil for the computation of $h_{t,xxx,i+(3/2)}$, spreading over ten points was used. A more narrow stencil for the computation of this quantity is an option which could be investigated. A large number of options exist, which could be used for optimization.

7. Conclusions

The two versions of the o3o3 LGM scheme share with SEM3 the properties of being at least uniform fourth order and conserving first-order moments. The accuracy of the simulation was rather similar for the schemes tested. In this study, it was not investigated if other interesting properties of SEM3 are shared by o3o3. o3o3 differs from SEM3 by having a regular collocation grid. One of the o3o3 schemes had a CFL condition about 25% higher than standard o4 differences according to Table 1. SEM3 had a more restrictive CFL condition than either of the o3o3 schemes.

Two versions of o3o3 were investigated. These are examples for LGM schemes being alternatives to the currently popular quadrature approximation with SEM schemes. There are more options, such as using other FDM schemes at the corner points and employing polynomial spaces being more regular, such as differentiable at the corner points. The investigation of this large family of LGM schemes is above the scope of this paper and could potentially lead to a further increase of efficiency.

For 2D, o3o3 was implemented on the sparse serendipity grid, where not all points of the corresponding regular grid are used as dynamic points. In 2D, the sparseness factor, which is the proportion of the dynamic points to all points is 5:9 on square grids, which promises a corresponding increase of computational efficiency. The sparseness factor in 3D is 7:27.

On the negative side, our particular implementation of o3o3 has a rather large stencil of 16 points and a null space. Such features might be improved by investigating other options for o3o3. The dynamic equations are not solved on such unused points, which promises a considerable saving of computer time.

Acknowledgments. No author reported any potential conflicts of interest. This work is jointly supported by the National Key Research and Development Program of China (Grant 2017YFC0209800), the China Postdoctoral Science Foundation (Grant 2016M601101), and the National Key Research and Development Program of China (Grant 2016YFC1401705). The authors thank the cities of Erquy, Brittany, France and Bad Orb, Germany for providing office space for this cooperation. Dr. Fang is grateful for the support of the EPSRC grant: Managing Air for Green Inner Cities (MAGIC) (EP/N010221/1).

APPENDIX

Computation of the Evolution Matrix M

The evolution matrix $\mathbf{M}$ is given by

$$\mathbf{M} = \mathbf{M}^1 \cdot e^{-2i\delta} + \mathbf{M}^2 \cdot e^{-i\delta} + \mathbf{M}^3 + \mathbf{M}^4 \cdot e^{i\delta} + \mathbf{M}^5 \cdot e^{2i\delta} + \mathbf{M}^6 \cdot e^{3i\delta}, \tag{A1}$$

where $\mathbf{M}$ is applied to the field h_j ($j = 0, 3, 6, \ldots$), the field h_{xx} ($j = 1, 4, 7, \ldots$), and the field h_{xxx} ($j = 2, 5, 8, \ldots$), $i = \sqrt{-1}$, $\delta = k/1000 \cdot 2\pi$, $k = 0, 1, 2, \ldots, 1000$. The matrices $\mathbf{M}^1, \mathbf{M}^2, \ldots, \mathbf{M}^6$ are 3×3 matrices. They are given by

$$\mathbf{M}^1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ -\frac{1}{6dx}\mathbf{M}^2_{2,1} & -\frac{1}{6dx}\mathbf{M}^2_{2,2} & -\frac{1}{6dx}\mathbf{M}^2_{2,3} \end{pmatrix} \tag{A2}$$

$$\mathbf{M}^2 = \begin{pmatrix} \mathbf{M}^2_{1,1} & \mathbf{M}^2_{1,2} & \mathbf{M}^2_{1,3} \\ \frac{2}{3dx^2}\mathbf{M}^2_{1,1} & \frac{2}{3dx^2}\mathbf{M}^2_{1,2} & \frac{2}{3dx^2}\mathbf{M}^2_{1,3} \\ -\frac{1}{6dx}\mathbf{M}^3_{2,1} & -\frac{1}{6dx}\mathbf{M}^3_{2,2} & -\frac{1}{6dx}\mathbf{M}^3_{2,3} \end{pmatrix} \tag{A3}$$

$$\mathbf{M}^3 = \begin{bmatrix} \mathbf{M}^3_{1,1} & \mathbf{M}^3_{1,2} & \mathbf{M}^3_{1,3} \\ \frac{4}{9dx^3} + \frac{2}{3dx^2}(\mathbf{M}^3_{1,1} + \mathbf{M}^2_{1,1}) & \frac{2}{3dx^2}(\mathbf{M}^3_{1,2} + \mathbf{M}^2_{1,2}) & \frac{2}{3dx^2}(\mathbf{M}^3_{1,3} + \mathbf{M}^2_{1,3}) \\ \frac{1}{6dx}(\mathbf{M}^2_{2,1} - \mathbf{M}^4_{2,1}) & \frac{1}{6dx}(\mathbf{M}^2_{2,2} - \mathbf{M}^4_{2,2}) & \frac{1}{6dx}(\mathbf{M}^2_{2,3} - \mathbf{M}^4_{2,3}) \end{bmatrix} \tag{A4}$$

$$\mathbf{M}^4 = \begin{bmatrix} \mathbf{M}^4_{1,1} & 0 & 0 \\ -\frac{4}{9dx^3} + \frac{2}{3dx^2}(\mathbf{M}^4_{1,1} + \mathbf{M}^3_{1,1}) & \frac{2}{3dx^2}(\mathbf{M}^4_{1,2} + \mathbf{M}^3_{1,2}) & \frac{2}{3dx^2}(\mathbf{M}^4_{1,3} + \mathbf{M}^3_{1,3}) \\ \frac{1}{6dx}(\mathbf{M}^3_{2,1} - \mathbf{M}^5_{2,1}) & \frac{1}{6dx}(\mathbf{M}^3_{2,2} - \mathbf{M}^5_{2,2}) & \frac{1}{6dx}(\mathbf{M}^3_{2,3} - \mathbf{M}^5_{2,3}) \end{bmatrix} \tag{A5}$$

$$\mathbf{M}^5 = \begin{pmatrix} 0 & 0 & 0 \\ \frac{2}{3dx^2}\mathbf{M}^4_{1,1} & \frac{2}{3dx^2}\mathbf{M}^4_{1,2} & \frac{2}{3dx^2}\mathbf{M}^4_{1,3} \\ \frac{1}{6dx}\mathbf{M}^4_{2,1} & \frac{1}{6dx}\mathbf{M}^4_{2,2} & \frac{1}{6dx}\mathbf{M}^4_{2,3} \end{pmatrix} \quad \text{(A6)}$$

$$\mathbf{M}^6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{1}{6dx}\mathbf{M}^5_{2,1} & \frac{1}{6dx}\mathbf{M}^5_{2,2} & \frac{1}{6dx}\mathbf{M}^5_{2,3} \end{pmatrix}, \quad \text{(A7)}$$

where $\mathbf{M}^k_{j_1,j_2}$ is the element of matrix $\mathbf{M}^k$ in row j_1 and column j_2, $dx = x_{i+1} - x_i$, where x_{i+1} and x_i are grid points (we assume $dx = 1/3$ for simplification), $\mathbf{M}^2_{1,:} = [-1/2, 7/36, 1/72]$, $\mathbf{M}^3_{1,:} = [0, -7/36, 1/72]$, and $\mathbf{M}^4_{1,1} = 1/2$ for standard o3o3, while $\mathbf{M}^2_{1,:} = [-1, 1/2, 1/12]$, $\mathbf{M}^3_{1,:} = [0, -1/2, 1/12]$, and $\mathbf{M}^4_{1,1} = 1$ for spectral o3o3.

REFERENCES

Ahlberg, J. H., E. N. Nilson, and J. L. Walsh, 1967: *The Theory of Splines and their Application*. Academic Press, 296 pp.

Baumgardner, J. R., and P. O. Frederickson, 1985: Icosahedral discretization of the two-sphere. *SIAM J. Numer. Anal.*, **22**, 1107–1115, https://doi.org/10.1137/0722066.

Chen, C., and F. Xiao, 2008: Shallow water model on cubed-sphere by multi-moment finite volume method. *J. Comput. Phys.*, **227**, 5019–5044, https://doi.org/10.1016/j.jcp.2008.01.033.

Cockburn, B., and C. W. Shu, 2001: Runge– Kutta discontinuous Galerkin methods for convection-dominated problems. *J. Sci. Comput.*, **16**, 173–261, https://doi.org/10.1023/A:1012873910884.

Dennis, J. M., and Coauthors, 2012: CAM-SE: A scalable spectral element dynamical core for the community atmosphere model. *Int. J. High Perform. Comput. Appl.*, **26**, 74–89, https://doi.org/10.1177/1094342011428142.

Durran, D., 2012: *Numerical Methods for Fluid Dynamics with Applications to Geophysics*. Springer, 465 pp.

Evans, L. C., 1998: *Partial Differential Equation*. Graduate Studies in Mathematics, Vol. 19, American Mathematical Society, 749 pp.

Giraldo, F. X., 2001: A spectral element shallow water model on spherical geodesic grids. *Int. J. Numer. Methods Fluids*, **35**, 869–901, https://doi.org/10.1002/1097-0363(20010430)35:8<869::AID-FLD116>3.0.CO;2-S.

Herrington, A. R., P. H. Lauritzen, M. A. Taylor, S. Goldhaber, B. E. Eaton, J.T. Bacmeister, K. A. Reed, and P. A. Ullrich, 2019: Physics–dynamics coupling with element-based high-order Galerkin methods: Quasi equal-area physics grid. *Mon. Wea. Rev.*, **147**, 69–84,https://doi.org/10.1175/MWR-D-18-0136.1.

Kalnay-Rivas, E., A. Bayliss, and J. Storch, 1977: The 4th order GISS model of the global atmosphere. *Beitr. Phys. Atmos.*, **50**, 299–311.

Klemp, J. B., W. C. Skamarock, and J. Dudhia, 2007: Conservative split-explicit time integration methods for the compressible nonhydrostatic equations. *Mon. Wea. Rev.*, **135**, 2897–2913, https://doi.org/10.1175/MWR3440.1.

Steppeler, J., 1976: The application of the second and third degree methods. *J. Comput. Phys.*, **22**, 295–318, https://doi.org/10.1016/0021-9991(76)90051-6.

——, 1987: *Galerkin and Finite Element Methods in Numerical Weather Prediction.* Duemmler, 120 pp.

——, and J. B. Klemp, 2017: Advection on cut-cell grids for an idealized mountain of constant slope. *Mon. Wea. Rev.*, **145**, 1765–1777, https://doi.org/10.1175/MWR-D-16-0308.1.

——, P. Ripodas, and S. Thomas, 2008: Third-order finite-difference schemes on isocahedral-type grids on the sphere. *Mon. Wea. Rev.*, **136**, 2683–2698, https://doi.org/10.1175/2007MWR2182.1.

Taylor, M. A., and A. Fournier, 2010: A compatible and conservative spectral element method on unstructured grids. *J. Comput. Phys.*, **229**, 5879–5895, https://doi.org/10.1016/j.jcp.2010.04.008.

Taylor, M., J. Tribbia, and M. Iskandarani, 1997: The spectral element method for the shallow water equations on the sphere. *J. Comput. Phys.*, **130**, 92–108, https://doi.org/10.1006/jcph.1996.5554.

Tomita, H., M. Tsugawa, M. Satoh, and K. Goto, 2001: Shallow water model on a modified isosahedral geodesic grid by using spring dynamics. *J. Comput. Phys.*, **174**, 579–613, https://doi.org/10.1006/jcph.2001.6897.

Ullrich, P. A., 2014a: A global finite-element shallow-water model supporting continuous and discontinuous elements. *Geosci. Model Dev.*, **7**, 3017–3035, https://doi.org/10.5194/gmd-7-3017-2014.

——, 2014b: Understanding the treatment of waves in atmospheric models. Part I: The shortest resolved waves of the 1d linearized shallow-water equations. *Quart. J. Roy. Meteor. Soc.*, **140**, 1426–1440, https://doi.org/10.1002/qj.2226.

——, D. Reynolds, J. E. Guerra, and M. A. Taylor, 2018: Impacts and importance of diffusion on the spectral element method: A linear analysis. *J. Comput. Phys.*, **375**, 427–446, https://doi.org/10.1016/j.jcp.2018.06.035.