

Multi-Level Spectral Deferred Corrections Scheme for the Shallow Water Equations on the Rotating Sphere

François P. Hamon^{a,*}, Martin Schreiber^{b,c}, Michael L. Minion^d

^aCenter for Computational Sciences and Engineering, Lawrence Berkeley National Laboratory, Berkeley, USA

^bDepartment of Mathematics/Computer Science, University of Exeter, Exeter, United Kingdom

^cChair of Computer Architecture and Parallel Systems, Technical University of Munich, Germany

^dDepartment of Applied Mathematics, Lawrence Berkeley National Laboratory, Berkeley, USA

Abstract

Efficient time integration schemes are necessary to capture the complex processes involved in atmospheric flows over long periods of time. In this work, we propose a high-order, implicit-explicit numerical scheme that combines Multi-Level Spectral Deferred Corrections (MLSDC) and the Spherical Harmonics (SH) transform to solve the wave-propagation problems arising from the shallow-water equations on the rotating sphere.

The iterative temporal integration is based on a sequence of corrections distributed on coupled space-time levels to perform a significant portion of the calculations on a coarse representation of the problem and hence to reduce the time-to-solution while preserving accuracy. In our scheme, referred to as MLSDC-SH, the spatial discretization plays a key role in the efficiency of MLSDC, since the SH basis allows for consistent transfer functions between space-time levels that preserve important physical properties of the solution.

We study the performance of the MLSDC-SH scheme with shallow-water test cases commonly used in numerical atmospheric modeling. We use this suite of test cases, which gradually adds more complexity to the nonlinear system of governing partial differential equations, to perform a detailed analysis of the accuracy of MLSDC-SH upon refinement in time. We illustrate the stability properties of MLSDC-SH and show that the proposed scheme achieves up to eighth-order convergence in time. Finally, we study the conditions in which MLSDC-SH achieves its theoretical speedup, and we show that it can significantly reduce the computational cost compared to single-level Spectral Deferred Corrections (SDC).

Keywords: high-order time integration, multi-level spectral deferred corrections, implicit-explicit splitting, atmospheric flows, shallow-water equations on the rotating sphere, spherical harmonics

1. Introduction

The numerical modeling of global atmospheric processes presents a challenging application area requiring accurate time integration methods for the discretized governing partial differential equations. These complex processes operate on a wide range of time scales but often have to be simulated over long periods of time – up to a hundred years for long-term paleoclimate studies – which constitutes a challenge for the design of stable and efficient integration schemes. One strategy for creating more efficient temporal integration schemes for such systems is to employ a semi-implicit scheme that allows larger time steps to be taken than with explicit methods at a cost that is less than that of fully implicit methods (Giraldo, 2005). A second strategy is to use a parallel-in-time strategy to solve multiple time steps concurrently on multiple processors. Examples of parallel-in-time methods include Parareal (Lions et al., 2001), the Parallel Full Approximation Scheme in Space and Time, (PFASST, Emmett and Minion (2012)), and

*Corresponding author

MultiGrid Reduction in Time (MGRIT, Falgout et al. (2014)). In this work, we consider semi-implicit, iterative, multi-level temporal integration methods based on Spectral Deferred Corrections (SDC) that are easily extended to high-order and also serve as a first step toward constructing parallel-in-time integration methods for the atmospheric dynamics based on PFASST.

SDC methods are first presented in Dutt et al. (2000) and consist in applying a sequence of low-order corrections – referred to as sweeps – to a provisional solution in order to achieve high-order accuracy. Single-level SDC schemes have been applied to a wide range of problems, including reacting flow simulation (Bourlioux et al., 2003; Layton and Minion, 2004), atmospheric modeling (Jia et al., 2013), particle motion in magnetic fields (Winkel et al., 2015), and radiative transport modeling (Crockatt et al., 2017). In Jia et al. (2013), a fully implicit SDC scheme is combined with the Spectral Element Method (SEM) to solve the shallow-water equations on the rotating sphere. The authors demonstrate that the SDC method can take larger stable time steps than competing explicit schemes such as leapfrog, second-order Runge-Kutta methods, and implicit second-order Backward Differentiation Formula (BDF) method without loss of accuracy.

The approach considered here for atmospheric simulations builds on the work of Speck et al. (2015), in which a Multi-Level Spectral Deferred Corrections (MLSDC) scheme is proposed to improve the efficiency of the SDC time integration process while preserving its high-order accuracy. MLSDC relies on the construction of coarse space-time representations – referred to as levels – of the problem under consideration. The calculations are then performed on this hierarchy of levels in a way that shifts a significant portion of the computational burden to the coarse levels. As in nonlinear multigrid methods, the space-time levels are coupled by the introduction of a Full Approximation Scheme (FAS) term in the collocation problems solved on coarse levels. With this multi-level approach, the iterative correction process requires fewer fine sweeps than the standard SDC scheme but still achieves fast convergence to the fixed point solution. Synthetic numerical examples demonstrate the efficiency and accuracy of the MLSDC approach.

The MLSDC approach is combined here with a spatial discretization based on the global Spherical Harmonics (SH) transform to solve the shallow-water equations on the rotating sphere. This study is relevant for practical applications since the SH transform is implemented in major forecasting systems such as the Integrated Forecast System (IFS) at the European Centre for Medium-Range Weather Forecasts (ECMWF, Wedi et al. (2013)) and the Global Spectral Model (GSM) at the Japan Meteorological Agency (JMA, Kanamitsu et al. (1983)). Using a highly accurate method in space significantly reduces the spatial discretization errors and allows us to focus on the temporal integration. Our approach, referred to as MLSDC-SH, uses a temporal splitting in which only the stiff linear terms in the governing equations are treated implicitly, whereas less stiff terms are evaluated explicitly. Here, the word *stiff* is used to denote the terms that limit the time step size of fully explicit schemes. The temporal integration scheme retains the main features of the multi-level algorithm presented in Speck et al. (2015), and takes full advantage of the structure of the spatial discretization to achieve efficiency. Specifically, we construct accurate interpolation and restriction functions between space-time levels by padding or truncating the spectral representation of the variables in the SH transform. In addition, the spherical harmonics combined with the implicit-explicit temporal splitting considered in this work circumvent the need for a global linear solver and rely on an efficient local solver for the implicit systems.

We illustrate the properties of MLSDC-SH using a widely used suite of shallow-water test cases (Williamson et al., 1992; Galewsky et al., 2004). We start the numerical study with a steady-state benchmark that highlights the connection between the magnitude of the spectral coefficients truncated during coarsening and the convergence rate of MLSDC-SH upon refinement in time. Then we proceed to more challenging unsteady test cases to show that MLSDC-SH is stable for large time steps and achieves up to eighth-order temporal convergence. Finally, we investigate the conditions in which the proposed scheme achieves its theoretical speedup and we demonstrate that MLSDC-SH can reduce the computational cost compared to single-level SDC schemes.

In the remainder of the paper, we first introduce the system of governing equations in Section 2. Then, we briefly review the fundamentals of the spatial discretization based on the global SH transform in Section 3. In Section 4, we describe the implicit-explicit temporal integration scheme, with an emphasis on the Multi-Level Spectral Deferred Corrections (MLSDC) scheme. Finally, in Section 5, we present numerical examples on the sphere demonstrating the efficiency and accuracy of our approach.

2. Governing equations

We consider the Shallow-Water Equations (SWE) on the rotating sphere. These equations capture the main horizontal effects present in the full atmospheric equations. Well-defined test cases are available – such as those considered in this work – that relate the SWE to some key features of the full atmospheric equations. Hence, they provide a simplified assessment of the properties of temporal and spatial discretizations for atmospheric simulations on the rotating sphere. We use the vorticity-divergence formulation (Bourke, 1972; Hack and Jakob, 1992) in which the prognostic variables $\mathbf{U} = [\Phi, \zeta, \delta]^T$ are respectively the potential, Φ , the vorticity, ζ , and the divergence, δ . Here, the vorticity and divergence state variables are used to overcome the singularities in the velocity field at the poles.

The system of governing partial differential equations is

$$\frac{\partial \Phi'}{\partial t} = -\nabla \cdot (\Phi' \mathbf{V}) - \bar{\Phi} \delta + \nu \nabla^2 \Phi', \quad (1)$$

$$\frac{\partial \zeta}{\partial t} = -\nabla \cdot (\zeta + f) \mathbf{V} + \nu \nabla^2 \zeta, \quad (2)$$

$$\frac{\partial \delta}{\partial t} = \mathbf{k} \cdot \nabla \times (\zeta + f) \mathbf{V} - \nabla^2 \left(\Phi + \frac{\mathbf{V} \cdot \mathbf{V}}{2} \right) + \nu \nabla^2 \delta, \quad (3)$$

where \mathbf{k} is the outward radial unit vector. The average geopotential, $\bar{\Phi} = g\bar{h}$, is written as the product of the gravitational acceleration by the average height, and Φ' is defined as $\Phi' = \Phi - \bar{\Phi}$. The horizontal velocity vector is $\mathbf{V} \equiv \mathbf{i}u + \mathbf{j}v$, where \mathbf{i} and \mathbf{j} are the unit vectors in the eastward and northward directions, respectively. The Coriolis force is represented by $f = 2\Omega \sin \phi$, where Ω is the angular rate of rotation, and ϕ is the latitude. The diffusion coefficient is denoted by ν . Including a diffusion term in the governing equations is used in practice in atmospheric simulations to stabilize the flow dynamics and reduce the errors caused by nonlinearly interacting modes. Using the inviscid equations is not a viable option due to the extremely fast generation of small-scale features (Galewsky et al., 2004), in particular for global spectral methods using a collocated grid. For simplicity and reproducibility, we employ a second-order diffusion term with a diffusion coefficient set to $\nu = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$ for all spatial resolutions as in Galewsky et al. (2004). To express the velocities as a function of the prognostic variables, ζ and δ , we first use the Helmholtz theorem which relates \mathbf{V} to a scalar stream function, ψ , and a scalar velocity potential, χ ,

$$\mathbf{V} = \mathbf{k} \times \nabla \psi + \nabla \chi. \quad (4)$$

Using the identities

$$\zeta \equiv \mathbf{k} \cdot (\nabla \times \mathbf{V}), \quad (5)$$

$$\delta \equiv \nabla \cdot \mathbf{V}, \quad (6)$$

the application of the curl and divergence operators to (4) yields $\zeta = \nabla^2 \psi$ and $\delta = \nabla^2 \chi$. The Laplacian operators can be efficiently inverted using the SH transform to compute the stream function, ψ , and the velocity potential, χ , as a function of ζ and δ , as explained in Section 3. Equations (1), (2), and (3), form the system that we would like to solve.

Next, we use the identities (5) and (6) to split the right-hand side of (1), (2), and (3) into linear and nonlinear parts as follows

$$\frac{\partial \mathbf{U}}{\partial t} = \mathcal{L}_G(\mathbf{U}) + \mathcal{L}_F(\mathbf{U}) + \mathcal{N}(\mathbf{U}). \quad (7)$$

The first term in the right-hand side of (7) represents the linear wave motion induced by gravitational forces and also includes the diffusion term

$$\mathcal{L}_G(\mathbf{U}) \equiv [-\bar{\Phi} \delta + \nu \nabla^2 \Phi', \nu \nabla^2 \zeta, -\nabla^2 \Phi + \nu \nabla^2 \delta]^T. \quad (8)$$

The second term in the right-hand side of (7) contains a linear harmonic oscillator on the velocity components that includes the Coriolis term

$$\mathcal{L}_F(\mathbf{U}) \equiv [0, -f\delta - \mathbf{V} \cdot \nabla f, f\zeta + \mathbf{k} \cdot (\nabla f) \times \mathbf{V}]^T. \quad (9)$$

The third term in the right-hand side of (7) represents the nonlinear operators

$$\mathcal{N}(\mathbf{U}) \equiv \left[-\nabla \cdot (\Phi' \mathbf{V}), -\nabla \cdot (\zeta \mathbf{V}), \mathbf{k} \cdot \nabla \times (\zeta \mathbf{V}) - \nabla^2 \frac{\mathbf{V} \cdot \mathbf{V}}{2} \right]^T. \quad (10)$$

In Section 4, this decomposition is used to define the temporal implicit-explicit splitting chosen based on the stiffness of the different terms. Next, the details of the spatial discretization of \mathcal{L}_F , \mathcal{L}_G , and \mathcal{N} are presented.

3. Spatial discretization

This section presents an overview of the spatial discretization based on the global SH transform applied to the system of governing equations. The global SH transform is a key feature of the multi-level scheme presented here since it allows for simple and accurate data transfer between different spatial levels. We will show with numerical examples in Section 5 that this is critical for the design of efficient MLSDC schemes. In the SH scheme, the representation of a function of longitude λ and Gaussian latitude $\mu \equiv \sin(\phi)$, $\xi(\lambda, \mu)$, consists of a sum of spherical harmonic basis functions $P_s^r(\mu)e^{ir\lambda}$ weighted by the spectral coefficients ξ_s^r ,

$$\xi(\lambda, \mu) = \sum_{r=-R}^R \sum_{s=|r|}^{S(r)} \xi_s^r P_s^r(\mu) e^{ir\lambda}, \quad (11)$$

where the index r (respectively, s) refers to the latitudinal (respectively, longitudinal) mode. In (11), P_s^r is the normalized associated Legendre polynomial. Without loss of generality, we use a triangular truncation with $S(r) = R$. In Section 4.3.3, we will explain that a coarse representation of ξ can be obtained by simply truncating the number of modes – i.e., reducing R and S in (11) – to construct a hierarchy of spatial levels with different degrees of coarsening in MLSDC-SH. The transformation from physical to spectral space is achieved in two steps. The first step consists in taking the discrete Fourier transform of $\xi(\lambda, \mu)$ in longitude – i.e., over λ –, defined as

$$\xi^r(\mu) = \frac{1}{I} \sum_{\iota=1}^I \xi(\lambda_\iota, \mu) e^{-ir\lambda_\iota}, \quad (12)$$

where I denotes the number of grid points in the longitudinal direction, located at longitudes $\lambda_\iota = \frac{2\pi\iota}{I}$. Then, in the second step, the application of the discrete Legendre transformation in latitude yields

$$\xi_s^r = \sum_{j=1}^J \xi^r(\mu_j) P_s^r(\mu_j) w_j. \quad (13)$$

In (13), J is the number of Gaussian latitudes μ_j , chosen as the roots of the Legendre polynomial of degree J , P_J , and w_j denotes the Gaussian weight at latitude μ_j . This two-step global transform is applied to (7) to obtain a system of coupled ordinary differential equations involving the prognostic variables in spectral space, $\Theta_s^r = [\Phi_s^r, \zeta_s^r, \delta_s^r]$. Note that due to the symmetry of the spectral coefficients, it is sufficient to only include the indices $r \geq 0$. Hence, for $r \in \{0, \dots, R\}$ and $s \in \{r, \dots, R\}$, the equations are

$$\frac{\partial \Theta_s^r}{\partial t} = (\mathbf{L}_G)_s^r(\Theta) + (\mathbf{L}_F)_s^r(\Theta) + \mathbf{N}_s^r(\Theta), \quad (14)$$

where $(\mathbf{L}_G)_s^r$, $(\mathbf{L}_F)_s^r$, and \mathbf{N}_s^r are the discrete, spectral representations of the operators defined in (8), (9), and (10). The state variable in spectral space, Θ , is defined as a vector of size $K = 3R(R+1)/2$ as follows

$$\Theta \equiv [\Theta_0^0, \Theta_0^1, \dots, \Theta_{R-1}^R, \Theta_R^R]^T. \quad (15)$$

We refer to the work of Hack and Jakob (1992) for a thorough presentation of the scheme, including the full expression of the right-hand side of (14) in spectral space. More details about an efficient implementation of the global SH transform can be found in Temperton (1991); Rivier et al. (2002). The implementation of the spherical harmonics transformation used in this work is based on the SHTns library developed by Schaeffer (2013). Next, we proceed to the presentation of the discretization in time based on spectral deferred corrections.

4. Temporal discretization

4.1. Temporal splitting

The choice of a temporal splitting for the right-hand side of (14) is one of the key determinants of the performance of the scheme. Fully explicit schemes are based on inexpensive local updates but are limited by a severe stability restriction on the time step size. In the context of atmospheric modeling, this limitation is often caused by the presence of fast waves (e.g., sound or gravity waves) propagating in the system. Fully implicit schemes overcome the stability constraint on the time step size but rely on costly nonlinear global implicit solves to update all the degrees of freedom simultaneously (Evans et al., 2010; Jia et al., 2013; Lott et al., 2015).

Instead, implicit-explicit (IMEX) schemes only treat the stiff terms responsible for the propagation of the fast-moving waves implicitly, while the non-stiff terms that represent processes operating on a slower time scale are evaluated explicitly. This strategy reduces the cost of the implicit solves, relative to fully implicit solves, and allows for relatively large stable time steps. A common IMEX approach employed in non-hydrostatic atmospheric modeling is based on dimensional splitting and implicitly discretizes only the terms involved in the (fast) vertical dynamics (Ullrich and Jablonowski, 2012; Durran and Blossey, 2012; Weller et al., 2013; Giraldo et al., 2013; Lock et al., 2014; Gardner et al., 2018). Alternatively, the approach of Robert et al. (1972); Giraldo (2005) consists in linearizing the governing PDEs in the neighborhood of a reference state. The linearized piece is then discretized implicitly, and the term treated explicitly is obtained by subtracting the linearized piece from the nonlinear system. For the shallow-water equations, we directly discretize the fast linear terms on the right-hand side of (14) implicitly, while the other terms are evaluated explicitly. Specifically, we investigate an IMEX scheme based on the following splitting, for $r \in \{0, \dots, R\}$ and $s \in \{r, \dots, R\}$,

$$\frac{\partial \Theta_s^r}{\partial t} = (\mathbf{F}_I)_s^r(\Theta) + (\mathbf{F}_E)_s^r(\Theta), \quad (16)$$

in which the implicit right-hand side, $(\mathbf{F}_I)_s^r$, contains the terms representing linear wave motion induced by gravitational forces and the diffusion term. The explicit right-hand side, $(\mathbf{F}_E)_s^r$, contains the linear harmonic oscillator and the nonlinear terms. This temporal splitting leads to implicit and explicit right-hand sides defined as

$$(\mathbf{F}_I)_s^r \equiv (\mathbf{L}_G)_s^r, \quad (17)$$

$$(\mathbf{F}_E)_s^r \equiv (\mathbf{L}_F)_s^r + \mathbf{N}_s^r. \quad (18)$$

This implicit-explicit approach greatly simplifies the solution strategy for the implicit systems and circumvents the need for a global linear solver. The solution algorithm treats the geopotential separately from the divergence and vorticity variables. Since the Coriolis term and the nonlinear terms are treated explicitly, one can form a diagonal linear system in spectral space to update the geopotential, and then update locally the vorticity and divergence variables. This is explained in Section 4.4. We will investigate the stability and accuracy of the splitting with numerical examples in Section 5. Next, we describe the multi-level temporal integration scheme starting with the fundamentals of SDC.

4.2. IMEX Spectral Deferred Corrections

We start with a review of the fundamentals of the Spectral Deferred Corrections (SDC) scheme. SDC methods have been introduced in Dutt et al. (2000) and later extended to methods with different temporal splittings in Minion (2003); Bourlioux et al. (2003); Layton and Minion (2004). In Minion (2003), an implicit-explicit SDC method is described and referred to as *semi-implicit SDC* to contrast the method with subsequent *multi-implicit SDC* methods with multiple implicit terms introduced in Bourlioux et al. (2003). Here we employ the more used term *IMEX* to refer to SDC methods with an implicit-explicit splitting. The properties of IMEX SDC schemes for fast-wave slow-wave problems are analyzed in Ruprecht and Speck (2016). We consider a system of coupled ODEs in the generic form

$$\frac{\partial \Theta}{\partial t}(t) = \mathbf{F}_I(\Theta(t)) + \mathbf{F}_E(\Theta(t)), \quad t \in [t^n, t^n + \Delta t], \quad (19)$$

$$\Theta(t^n) = \Theta^n, \quad (20)$$

and its solution in integral form given by

$$\Theta(t) = \Theta^n + \int_{t^n}^t (\mathbf{F}_I + \mathbf{F}_E)(\Theta(a)) da = \Theta^n + \int_{t^n}^t \mathbf{F}(\Theta(a)) da, \quad (21)$$

where \mathbf{F}_I and \mathbf{F}_E are the implicit and explicit right-hand sides, respectively, with $\mathbf{F} = \mathbf{F}_I + \mathbf{F}_E$, and $\Theta(t)$ is the state variable at time t . In (21), the integral is applied componentwise. Denote by $\tilde{\Theta}(t)$ an approximation of $\Theta(t)$, and then define the correction $\Delta\Theta(t) = \Theta(t) - \tilde{\Theta}(t)$. The SDC scheme applied to the implicit-explicit temporal splitting described above iteratively improves the accuracy of the approximation based on a discretization of the update or correction equation

$$\begin{aligned} \tilde{\Theta}(t) + \Delta\Theta(t) &= \Theta^n + \int_{t^n}^t [\mathbf{F}_E(\tilde{\Theta}(a) + \Delta\Theta(a)) - \mathbf{F}_E(\tilde{\Theta}(a))] da \\ &+ \int_{t^n}^t [\mathbf{F}_I(\tilde{\Theta}(a) + \Delta\Theta(a)) - \mathbf{F}_I(\tilde{\Theta}(a))] da \\ &+ \int_{t^n}^t \mathbf{F}(\tilde{\Theta}(a)) da, \end{aligned} \quad (22)$$

where Θ^n is the (known) state variable at the beginning of the time step. In the update equation (22), the last integral is computed with a high-order Gaussian quadrature rule. However, the other integrals are approximated with simpler low-order quadrature rules. We mention here that all the quadrature rules used in (22) are based on a relatively small number of Gauss points – up to five in this work – compared to the quadrature rule used in the discrete Legendre transform to obtain (13). Each pass of the discrete version of the update equation (22), referred to as sweep, increases the formal order of accuracy by one until the order of accuracy of the quadrature applied to the third integral is reached (Hagstrom and Zhou, 2007; Xia et al., 2007; Christlieb et al., 2009).

To discretize the update equation (22), the correction algorithm uses a decomposition of the time interval $[t^n, t^{n+1}]$ into M subintervals using $M + 1$ temporal nodes, such that

$$t^n \equiv t^{n,0} < t^{n,1} < \dots < t^{n,M} = t^n + \Delta t \equiv t^{n+1}. \quad (23)$$

The points $t^{n,m}$ are chosen to correspond to Gaussian quadrature nodes. Throughout this paper, we use Gauss-Lobatto nodes. We use the shorthand notations $t^m = t^{n,m}$ and $\Delta t^m = t^{m+1} - t^m$. We denote by $\Theta^{m+1,(k+1)}$ the approximate solution at node $m + 1$ and at sweep $(k + 1)$. The terms in the first integral of (22) are treated explicitly, and therefore this integral is discretized with a forward Euler method. Conversely, the second integral in (22) is discretized implicitly. The general form of the discrete version of equation (22) is then

$$\begin{aligned} \Theta^{m+1,(k+1)} &= \Theta^n + \Delta t \sum_{j=1}^m \tilde{q}_{m+1,j}^E [\mathbf{F}_E(\Theta^{j,(k+1)}) - \mathbf{F}_E(\Theta^{j,(k)})] \\ &+ \Delta t \sum_{j=1}^{m+1} \tilde{q}_{m+1,j}^I [\mathbf{F}_I(\Theta^{j,(k+1)}) - \mathbf{F}_I(\Theta^{j,(k)})] \\ &+ \Delta t \sum_{j=0}^M q_{m+1,j} \mathbf{F}(\Theta^{j,(k)}). \end{aligned} \quad (24)$$

In (24), the coefficients $\tilde{q}_{m+1,j}^E$ correspond to forward-Euler time stepping. The coefficients $q_{m+1,j}$ correspond to the Lobatto IIIA optimal order collocation quadrature

$$q_{m+1,j} \equiv \frac{1}{\Delta t} \int_{t^{n,0}}^{t^{n,m+1}} L^j(a) da, \quad (25)$$

where L^j denotes the j^{th} Lagrange polynomial constructed using the SDC nodes (23). We note that the formulation of the correction given in (24) differs from that of Jia et al. (2013) in two ways. First,

our scheme is based on an implicit-explicit splitting, whereas that of Jia et al. (2013) is fully implicit. Second, for the choice of the quadrature weights used in the discretization of the implicit correction integral, we adopt the approach of Weiser (2015). Specifically, the weights $\tilde{q}_{m+1,j}^f$ in (24) are chosen to be the coefficients of the upper triangular matrix in the LU decomposition of $\mathbf{Q} = \{q_{ij}\} \in \mathbb{R}^{(M+1) \times (M+1)}$, while a diagonal matrix is used in Jia et al. (2013). This formulation leads to a faster convergence of the iterative process to the fixed-point solution and remains convergent even when the underlying problem is stiff. We refer to Weiser (2015) for a proof, and to Hamon et al. (ming) for numerical examples illustrating the improved convergence.

Using these definitions, the integration scheme (24) is effectively an iterative solution method for the collocation problem defined by

$$\mathbf{A}(\vec{\Theta}) = \mathbf{1}_{M+1} \otimes \Theta^{n,0}. \quad (26)$$

The operator \mathbf{A} is

$$\mathbf{A}(\vec{\Theta}) \equiv \vec{\Theta} - \Delta t(\mathbf{Q} \otimes \mathbf{I}_K) \vec{\mathbf{F}}, \quad (27)$$

where \otimes denotes the Kronecker product and $\mathbf{I}_K \in \mathbb{R}^{K \times K}$ is the identity matrix. $\mathbf{1}_{M+1} \in \mathbb{R}^{M+1}$ is a vector of ones. Following the notation used in Bolten et al. (2017), the space-time vectors $\vec{\Theta} \in \mathbb{C}^{(M+1)K}$ and $\vec{\mathbf{F}} \in \mathbb{C}^{(M+1)K}$ are such that

$$\vec{\Theta} \equiv [\Theta^{n,0}, \dots, \Theta^{n,M}]^T, \quad (28)$$

$$\vec{\mathbf{F}} \equiv \vec{\mathbf{F}}(\vec{\Theta}) = [\mathbf{F}(\Theta^{n,0}), \dots, \mathbf{F}(\Theta^{n,M})]^T. \quad (29)$$

Next, we introduce the multi-level algorithm based on SDC that we will apply to the collocation problem (26).

4.3. Multi-Level Spectral Deferred Corrections (MLSDC)

Multi-Level Spectral Deferred Corrections (MLSDC) schemes are based on the idea of replacing some of the SDC iterations required to converge to the collocation problem (26) with SDC sweeps performed on a coarsened (and hence computationally cheaper) version of the problem. The solutions on different levels are coupled by the introduction of a Full Approximation Scheme (FAS) correction term explained below as in nonlinear multigrid methods. The combination of performing SDC sweeps on multiple space-time levels with a FAS correction term first appears as part of the PFASST method in Emmett and Minion (2012). The idea is generalized and analyzed in Speck et al. (2015) showing how MLSDC can improve the efficiency for certain problems compared to single-level SDC methods. In this work, only two-level MLSDC schemes are considered, and the study of MLSDC with three or more space-time levels to integrate the shallow-water equations is left for future work.

4.3.1. Full Approximation Scheme (FAS)

We define two space-time levels to solve the collocation problem (26), and we denote by $\ell = f$ (respectively, $\ell = c$) the fine level (respectively, the coarse level). We denote by $\vec{\Theta}_\ell \in \mathbb{C}^{(M_\ell+1)K_\ell}$ and $\vec{\mathbf{F}}_\ell \in \mathbb{C}^{(M_\ell+1)K_\ell}$ the space-time vector and right-hand side at level ℓ , respectively. The matrix $\mathbf{R}_f^c \in \mathbb{R}^{(M_c+1)K_c \times (M_f+1)K_f}$ is the linear restriction operator from the fine level to the coarse level. Here, K_ℓ represents the total number of spectral coefficients in (15) on level ℓ . As in nonlinear multigrid methods (Brandt, 1977), the coarse problem is modified by the introduction of a correction term, denoted by $\vec{\tau}_c$, that couples the solutions at the two space-time levels. Specifically, the coarse problem reads

$$\mathbf{A}_c(\vec{\Theta}_c) - \vec{\tau}_c = \mathbf{1}_{M_c+1} \otimes \Theta_c^{n,0}, \quad (30)$$

where the FAS correction term at the coarse level is defined as

$$\vec{\tau}_c \equiv \mathbf{A}_c(\mathbf{R}_f^c \vec{\Theta}_f) - \mathbf{R}_f^c \mathbf{A}_f(\vec{\Theta}_f) + \mathbf{R}_f^c \vec{\tau}_f, \quad (31)$$

with, for the two-level case, $\vec{\tau}_f = \mathbf{0}$ on the fine level. In (30)-(31), the operator \mathbf{A}_c denotes an approximation of \mathbf{A} at the coarse level. We note that

$$\begin{aligned} \mathbf{A}_c(\mathbf{R}_f^c \vec{\Theta}_f) - \vec{\tau}_c &= \mathbf{A}_c(\mathbf{R}_f^c \vec{\Theta}_f) - \mathbf{A}_c(\mathbf{R}_f^c \vec{\Theta}_f) + \mathbf{R}_f^c \mathbf{A}_f(\vec{\Theta}_f) - \mathbf{R}_f^c \vec{\tau}_f \\ &= \mathbf{R}_f^c (\mathbf{A}_f(\vec{\Theta}_f) - \vec{\tau}_f), \end{aligned} \quad (32)$$

which implies that the restriction of the fine solution, $\mathbf{R}_f^c \vec{\Theta}_f$, is a solution of the coarse problem. On the coarse problem (30), the modified SDC update for temporal node $m+1$ at sweep $(k+1)$ is

$$\begin{aligned}
 \Theta_c^{m+1,(k+1)} &= \Theta_c^{n,0} \\
 &+ \Delta t \sum_{j=1}^m (\tilde{q}_{m+1,j}^E)_c [\mathbf{F}_{E,c}(\Theta_c^{j,(k+1)}) - \mathbf{F}_{E,c}(\Theta_c^{j,(k)})] \\
 &+ \Delta t \sum_{j=1}^{m+1} (\tilde{q}_{m+1,j}^I)_c [\mathbf{F}_{I,c}(\Theta_c^{j,(k+1)}) - \mathbf{F}_{I,c}(\Theta_c^{j,(k)})] \\
 &+ \Delta t \sum_{j=0}^M (q_{m+1,j})_c \mathbf{F}_c(\Theta_c^{j,(k)}) + \tau_c^{m+1,(k)}.
 \end{aligned} \tag{33}$$

4.3.2. IMEX MLSDC algorithm

We are now ready to review the steps of the MLSDC algorithm of Emmett and Minion (2012); Speck et al. (2015) for the case of two space-time levels. In this section, $\Theta_\ell^{m,(k)}$ denotes the approximate solution at temporal node m , space-time level ℓ , and sweep (k) . $\vec{\Theta}_\ell^{(k)}$ is the space-time vector that contains the approximate solution at all temporal nodes on level ℓ . The vectors $\mathbf{F}_\ell^{m,(k)}$ and $\vec{\mathbf{F}}_\ell^{(k)}$ are defined analogously. The MLSDC iteration starts with an SDC sweep on the fine level. The iteration continues as in a V-cycle from the fine level to the coarse level, and then back to the fine level. The specifics of the MLSDC iteration with two space-time levels are detailed in Algorithm 1.

Algorithm 1: IMEX MLSDC iteration on two space-time levels denoted by “coarse” and “fine”.

Data: Initial data $\Theta_f^{0,(k)}$ and function evaluations $\vec{\mathbf{F}}_{I,f}^{(k)}$, $\vec{\mathbf{F}}_{E,f}^{(k)}$ from the previous MLSDC iteration (k) on the fine level.

Result: Approximate solution $\vec{\Theta}_\ell^{(k+1)}$ and function evaluations $\vec{\mathbf{F}}_{I,\ell}^{(k+1)}$, $\vec{\mathbf{F}}_{E,\ell}^{(k+1)}$ on all levels.

A) Perform a fine sweep

$$\vec{\Theta}_f^{(k+1)}, \vec{\mathbf{F}}_{I,f}^{(k+1)}, \vec{\mathbf{F}}_{E,f}^{(k+1)} \leftarrow \text{SweepFine}(\vec{\Theta}_f^{(k)}, \vec{\mathbf{F}}_{I,f}^{(k)}, \vec{\mathbf{F}}_{E,f}^{(k)})$$

B) Restrict, re-evaluate, and save restriction

for $m = 1, \dots, M_c$ do

$$\left| \begin{aligned}
 \Theta_c^{m,(k)} &\leftarrow \text{Restrict}(\Theta_f^{m,(k+1)}) \\
 \mathbf{F}_{I,c}^{m,(k)}, \mathbf{F}_{E,c}^{m,(k)} &\leftarrow \text{Evaluate_F}(\Theta_c^{m,(k)}) \\
 \vec{\Theta}_c^{m,(k)} &\leftarrow \Theta_c^{m,(k)} \\
 \vec{\mathbf{F}}_{I,c}^{m,(k)}, \vec{\mathbf{F}}_{E,c}^{m,(k)} &\leftarrow \mathbf{F}_{I,c}^{m,(k)}, \mathbf{F}_{E,c}^{m,(k)}
 \end{aligned} \right.$$

end

C) Compute FAS correction and sweep

$$\tau_c \leftarrow \text{FAS}(\vec{\mathbf{F}}_{I,f}^{(k)}, \vec{\mathbf{F}}_{E,f}^{(k)}, \vec{\mathbf{F}}_{I,c}^{(k)}, \vec{\mathbf{F}}_{E,c}^{(k)}, \tau_f)$$

$$\vec{\Theta}_c^{(k+1)}, \vec{\mathbf{F}}_{I,c}^{(k+1)}, \vec{\mathbf{F}}_{E,c}^{(k+1)} \leftarrow \text{SweepCoarse}(\vec{\Theta}_c^{(k)}, \vec{\mathbf{F}}_{I,c}^{(k)}, \vec{\mathbf{F}}_{E,c}^{(k)}, \tau_c)$$

D) Return to finest level before next iteration

for $m = 1, \dots, M_f$ do

$$\left| \begin{aligned}
 \Theta_f^{m,(k+1)} &\leftarrow \Theta_f^{m,(k+1)} + \text{Interpolate}(\Theta_c^{m,(k+1)} - \vec{\Theta}_c^{m,(k)}) \\
 \mathbf{F}_{I,f}^{m,(k+1)} &\leftarrow \mathbf{F}_{I,f}^{m,(k+1)} + \text{Interpolate}(\mathbf{F}_{I,c}^{m,(k+1)} - \vec{\mathbf{F}}_{I,c}^{m,(k)}) \\
 \mathbf{F}_{E,f}^{m,(k+1)} &\leftarrow \mathbf{F}_{E,f}^{m,(k+1)} + \text{Interpolate}(\mathbf{F}_{E,c}^{m,(k+1)} - \vec{\mathbf{F}}_{E,c}^{m,(k)})
 \end{aligned} \right.$$

end

The single-level SDC iteration only consists of the fine sweep of Step **A**. Both schemes share the

same initialization procedure. Specifically, before the first iteration, for $k = 0$, we initialize the algorithm described above by simply copying the initial data for the time step, denoted by $\Theta_f^{n,0}$, to all the other SDC nodes, that is, for $m \in \{0, \dots, M_f\}$:

$$\Theta_f^{m,(k=0)} := \Theta_f^{n,0}. \quad (34)$$

In Algorithm 1, the procedure *SweepFine* consists in applying (24) once on the fine level. The procedure *SweepCoarse* involves applying the correction described by (33) on the coarse level. We found that for the numerical examples considered in this work, doing multiple sweeps on the coarse level instead of one every time *SweepCoarse* is called does not improve the accuracy of MLSDC, but increases the computational cost. This is why the procedure *SweepCoarse* only involves one sweep per call. The procedure *Evaluate_F* involves computing the implicit and explicit right-hand sides. We highlight that the last step of Algorithm 1 does not involve any function evaluation. Instead, when we return to the fine level, we interpolate the coarse solution update as well as the coarse right-hand side corrections to the fine level. By avoiding M_f function evaluations, this reduces the computational cost of the algorithm without undermining the order of accuracy of the scheme, as shown with numerical examples in Section 5. We now discuss two key determinants of the performance of MLSDC-SH, namely, the coarsening strategy and the solver for the implicit systems.

4.3.3. Coarsening strategy and transfer functions

In this section, we describe the linear restriction and interpolation operators used in the MLSDC-SH algorithm to transfer the approximate solution from fine to coarse levels, and vice-versa. In this work, the spatial restriction and interpolation procedures are performed in spectral space and heavily rely on the decomposition (15) resulting from the SH basis. As explained below, this approach is based on the truncation of high-frequency modes, and therefore avoids the generation of spurious modes in the approximate solution that would propagate in the spectrum due to nonlinear wave interweaving over one coarse sweep.

We reiterate that K_ℓ denotes the number of spectral coefficients used in (15) at level ℓ and $M_\ell + 1$ denotes the number of SDC nodes at level ℓ . Therefore, the space-time vector storing the state of the system at level ℓ , denoted by $\vec{\Theta}_\ell$, is in $\mathbb{C}^{(M_\ell+1)K_\ell}$. The two-step restriction process from fine level $\ell = f$ to coarse level $\ell = c$ consists in applying a restriction operator in time, denoted by $(\mathbf{R}^t)_f^c$, followed by a restriction operator in space, denoted by $(\mathbf{R}^s)_f^c$, that is,

$$\vec{\Theta}_c = \mathbf{R}_f^c \vec{\Theta}_f = (\mathbf{R}^s)_f^c (\mathbf{R}^t)_f^c \vec{\Theta}_f. \quad (35)$$

In (35), the restriction operator in time is defined using the Kronecker product as

$$(\mathbf{R}^t)_f^c \equiv \mathbf{\Pi}_f^c \otimes \mathbf{I}_{K_f} \in \mathbb{R}^{(M_c+1)K_f \times (M_f+1)K_f}, \quad (36)$$

where $\mathbf{I}_{K_f} \in \mathbb{R}^{K_f \times K_f}$ is the identity matrix, and $\mathbf{\Pi}_f^c \in \mathbb{R}^{(M_c+1) \times (M_f+1)}$ is the rectangle matrix employed to interpolate a scalar function from the fine temporal discretization to the coarse temporal discretization. Using the Lagrange polynomials L_f^j on the fine temporal discretization, this matrix reads

$$(\mathbf{\Pi}_f^c)_{ij} = L_f^{j-1}(t_c^{i-1}), \quad (37)$$

using the SDC node $i - 1$ at the coarse level, denoted by t_c^{i-1} . We note that, in the special case of two, three, and five Gauss-Lobatto nodes, applying this restriction operator in time amounts to performing pointwise injection. The restriction operator in space consists in truncating the spectral representation of the primary variables (15) based on the SH transform to remove the high-frequency features from the approximate solution. This is achieved by applying the matrix

$$(\mathbf{R}^s)_f^c \equiv \mathbf{I}_{M_c+1} \otimes \mathbf{D}_f^c \in \mathbb{R}^{(M_c+1)K_c \times (M_c+1)K_f}. \quad (38)$$

In (38), $\mathbf{D}_f^c \in \mathbb{R}^{K_c \times K_f}$ is a rectangle truncation matrix defined as

$$(\mathbf{D}_f^c)_{ij} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise.} \end{cases} \quad (39)$$

In the interpolation procedure employed to transfer the approximate solution from the coarse level to the fine level, we start with the application of the interpolation operator in space, $(\mathbf{P}^s)_c^f$, followed by the application of the interpolation operator in time, $(\mathbf{P}^t)_c^f$,

$$\vec{\Theta}_f \equiv \mathbf{P}_c^f \vec{\Theta}_c = (\mathbf{P}^t)_c^f (\mathbf{P}^s)_c^f \vec{\Theta}_c. \quad (40)$$

The interpolation operator in space consists in padding the spectral representation of the primary variables at the coarse level with $K_f - K_c$ zeros and can be defined as the transpose of the restriction operator in space (see (38)), that is,

$$(\mathbf{P}^s)_c^f \equiv ((\mathbf{R}^s)_f^c)^T \in \mathbb{R}^{(M_c+1)K_f \times (M_c+1)K_c}. \quad (41)$$

Finally, the interpolation operator in time is analogous to (36) and reads

$$(\mathbf{P}^t)_c^f \equiv \mathbf{\Pi}_c^f \otimes \mathbf{I}_{K_f} \in \mathbb{R}^{(M_f+1)K_f \times (M_c+1)K_f}, \quad (42)$$

where the rectangle interpolation matrix $\mathbf{\Pi}_c^f$ is constructed with the Lagrange polynomials L_c^j on the coarse temporal discretization. For two, three, and five Gauss-Lobatto nodes, this amounts to performing pointwise injection at the fine nodes that correspond to the coarse nodes, and then polynomial interpolation to compute the solution at the remaining fine nodes. This completes the presentation of the MLSDC-SH algorithm for the time integration of the shallow-water equations on the rotating sphere. Next, we discuss the implicit solver used in this work.

4.4. Solver for the implicit systems in SDC and MLSDC

The time integration schemes of Sections 4.2 and 4.3 involve solving implicit linear systems in the form

$$\Theta^{m+1,(k+1)} - \Delta t \tilde{q}_{m+1,m+1}^I \mathbf{F}_I(\Theta^{m+1,(k+1)}) = \mathbf{b}, \quad (43)$$

where \mathbf{b} is obtained from (24) or (33), and where we have dropped the subscripts denoting the space-time levels for simplicity. The structure of the implicit linear systems results from the spatial discretization based on the SH transform, but also from the temporal splitting between implicit and explicit terms described in Section 4.1. The solution strategy for (43) is performed in spectral space and follows two steps briefly outlined below.

First, we algebraically form a reduced linear system containing only the geopotential unknowns – that is, $K/3$ degrees of freedom, where K denotes the total number of spectral coefficients needed to represent the three primary variables in (15). Given that the longitudinal and latitudinal coupling terms present in the Coriolis term and in the nonlinear operators are discretized explicitly, the $K/3$ geopotential degrees of freedom are fully decoupled from one another. The geopotential linear system is therefore diagonal and trivial to solve. Second, we have to solve for the remaining $2K/3$ vorticity and divergence degrees of freedom. This is again a trivial operation that does not require a linear solver, since we have to solve two diagonal linear systems to update the vorticity and divergence variables, respectively.

Therefore, solving the implicit system (43) is purely based on local operations during which the degrees of freedom are updated one at a time in spectral space. We refer to Schreiber and Loft (ming) for the detailed formulation of the geopotential, vorticity, and divergence diagonal linear systems.

4.5. Computational cost of SDC and MLSDC

In this section, we compare the computational cost of the MLSDC-SH scheme described in Section 4.3 to that of the single-level SDC scheme. We refer to the single-level SDC scheme with $M_f + 1$ temporal nodes and N_S fine sweeps as $\text{SDC}(M_f + 1, N_S)$. We denote by $\text{MLSDC}(M_f + 1, M_c + 1, N_{ML}, \alpha)$ the MLSDC-SH scheme with $M_f + 1$ nodes on the fine level, $M_c + 1$ nodes of the coarse level, N_{ML} iterations, and a spatial coarsening ratio, α , defined using (11) as $\alpha = R_c/R_f$. The parameters of the SDC and MLSDC-SH schemes are summarized in Tables 1 and 2, respectively.

To evaluate the theoretical computational cost of the MLSDC-SH scheme, we count the number of function evaluations and the number of solves involved in a time step. We denote by C_ℓ^s the cost of a solve at level ℓ , and by C_ℓ^{fi} (respectively, C_ℓ^{fe}) the cost of an implicit (respectively, explicit) function evaluation at level ℓ . We neglect the cost of computing the FAS correction. The quantities C_c^s , C_c^{fi} , C_c^{fe}

SDC($M_f + 1, N_S$)	
Parameter	Description
$M_f + 1$	SDC nodes on fine level
N_S	Number of SDC iterations

Table 1: Parameters for the SDC scheme. The SDC iteration only involves one sweep on the fine level.

MLSDC($M_f + 1, M_c + 1, N_{ML}, \alpha$)	
Parameter	Description
$M_f + 1$	SDC nodes on fine level
$M_c + 1$	SDC nodes on coarse level
N_{ML}	Number of MLSDC iterations
α	Spatial coarsening ratio

Table 2: Parameters for the MLSDC-SH scheme. The MLSDC-SH iteration is described in Alg. 1, and involves one sweep on the fine level and one sweep on the coarse level.

depend on the spatial coarsening ratio, $\alpha = R_c/R_f$. Here, R_ℓ represents the highest Fourier wavenumber in the east-west representation of (11) on level ℓ .

The cost of a time step with the two-level MLSDC($M_f + 1, M_c + 1, N_{ML}, \alpha$) is

$$\begin{aligned}
 C^{MLSDC(M_f+1, M_c+1, N_{ML}, \alpha)} &= N_{ML}M_f(C_f^s + C_f^{fi} + C_f^{fe}) \\
 &\quad + N_{ML}M_c(C_c^s + C_c^{fi} + C_c^{fe}) \\
 &\quad + N_{ML}M_c(C_c^{fi} + C_c^{fe}),
 \end{aligned} \tag{44}$$

where the term in the right-hand side of the first line represents the cost of the fine sweeps, the second term represents the cost of the coarse sweeps, and the third term accounts for the cost of the evaluation of the right-hand sides at the coarse nodes after the restriction. This can be compared with the cost of a time step in SDC($M_f + 1, N_S$), given by

$$C^{SDC(M_f+1, N_S)} = N_S M_f (C_f^s + C_f^{fi} + C_f^{fe}). \tag{45}$$

Furthermore, we assume that the cost of a linear solve is the same as the cost of evaluating the right-hand side, that is,

$$C_f^s = C_f^{fi} = C_f^{fe}, \tag{46}$$

with this assumption being motivated by Section 4.4. In addition, we will also assume that the computational cost of the operators is proportional to the number of spectral coefficients in (15), denoted by K_ℓ . There are three primary variables, and each of them is represented in the triangular truncation framework with $R_\ell(R_\ell + 1)/2$ spectral coefficients, where R_ℓ denotes the highest Fourier wavenumber in the east-west representation of (11). This yields $K_\ell = 3R_\ell(R_\ell + 1)/2$. Using this notation and the definition $\alpha = R_c/R_f$, we can obtain an expression of C_c^s as a function α , C_f^s , and R_f , by writing

$$C_c^s = \frac{K_c}{K_f} C_f^s = \frac{R_c(R_c + 1)}{R_f(R_f + 1)} C_f^s = \alpha^2 \frac{R_f + 1/\alpha}{R_f + 1} C_f^s. \tag{47}$$

Similarly, using the same assumptions, we obtain

$$C_c^{fi} = \alpha^2 \frac{R_f + 1/\alpha}{R_f + 1} C_f^{fi}, \quad C_c^{fe} = \alpha^2 \frac{R_f + 1/\alpha}{R_f + 1} C_f^{fe}. \tag{48}$$

Using these notations and assuming that MLSDC-SH and SDC achieve the same accuracy, the theoretical speedup obtained with MLSDC-SH, denoted by \mathcal{S}^{theo} , reads

$$\mathcal{S}^{theo} = \frac{C^{SDC(M_f+1, N_S)}}{C^{MLSDC(M_f+1, M_c+1, N_{ML}, \alpha)}} = \frac{N_S}{N_{ML}} \times \frac{1}{1 + \alpha^2 \frac{5(R_f + 1/\alpha)M_c}{3(R_f + 1)M_f}}. \tag{49}$$

That is, $\text{MLSDC}(3,2,2,1/2)$, based on two fine sweeps and two coarse sweeps, yields a theoretical speedup $\mathcal{S}^{theo} \approx 1.66$ compared to $\text{SDC}(3,4)$, which uses four fine sweeps. This corresponds to a reduction of 40 % in the wall-clock time. $\text{MLSDC}(5,3,4,1/2)$, based on four fine sweeps and four coarse sweeps, also results in a theoretical speedup $\mathcal{S}^{theo} \approx 1.66$ compared to $\text{SDC}(5,8)$, which relies on eight fine sweeps. This reasoning assumes that one MLSDC-SH iteration can replace two single-level SDC iterations and still achieve the same accuracy. This point is investigated in the next section using numerical examples.

We conclude this section by comparing the computational cost of MLSDC-SH to that of the fully implicit single-level SDC scheme based on the Spectral Element Method (SEM) presented in Jia et al. (2013) and referred to as SDC-SEM in the remainder of this paper. The fully implicit SDC-SEM iteration entails solving a large nonlinear system on the fine problem to update all the degrees of freedom simultaneously. This is achieved using the Jacobian-free Newton-Krylov (JFNK) method. As stated by the authors, the computational cost of this nonlinear solve can be large and heavily depends on the availability of a scalable preconditioner for the linear systems. Instead, the MLSDC-SH iteration involves trivial diagonal linear solves that can easily be parallelized. In addition, our multi-level time integration framework relies on a hierarchy of space-time levels to shift a significant fraction of the computational work to the coarser representation of the problem. This reduces the number of fine sweeps in the algorithm and therefore further reduces the cost of a time step.

As a result, we expect IMEX MLSDC-SH to be significantly less expensive than the fully implicit SDC-SEM on a per-timestep basis for moderate resolutions. Assuming that the linear systems can be efficiently preconditioned – as in Lott et al. (2015) – the key to the performance of SDC-SEM lies in its ability to take much larger stable time steps than MLSDC-SH to compensate for its relatively high cost on a per-timestep basis. Exploration of this trade-off requires a careful analysis that will be presented in future work.

5. Numerical examples

We assess the performance of MLSDC-SH with state-of-the-art test cases for the development of dynamical cores. All the test cases are nonlinear. They are selected to focus on particular challenges that arise with MLSDC-SH . The first test case in Section 5.1 targets geostrophically balanced modes. It evaluates the effects of multi-level mode truncation and the relation to the diffusion used in the simulations. The second test case in Section 5.2 studies the observed order of convergence of MLSDC-SH upon refinement in time for waves propagating on the rotating sphere. While the first two test cases are mainly dominated by the linear parts, the following benchmarks assess the performance of MLSDC-SH in the presence of stronger nonlinear interactions. The Rossby-Haurwitz benchmark in Section 5.3 studies the advection of a wave that propagates around the sphere without changing shape. This is followed by the unstable barotropic wave benchmark in Section 5.4 with an initially linear balanced flow perturbed by the introduction of a Gaussian bump in the geopotential field. All these benchmarks provide a key insight into the numerical properties of MLSDC-SH in the context of atmospheric simulations.

5.1. Steady zonal jet

We first study the behavior of the multi-level SDC scheme on a steady test case derived from Galewsky et al. (2004). This test case consists in the simulation of a steady, analytically specified mid-latitude jet with an unperturbed, balanced height field. This test assesses the ability of the numerical schemes to maintain this balanced state for 144 hours. The vorticity field obtained with the single-level $\text{SDC}(5,8)$ with a modal resolution of $R_f = S_f = 256$ is in Fig. 1, along with the corresponding vorticity spectrum in Fig. 2.

This steady numerical test is used to illustrate the order of convergence of MLSDC-SH upon refinement in time. We will consider the computational cost of the multi-level scheme in subsequent examples. We highlight that we do not use the steady geostrophic balance test case of Williamson et al. (1992) here because it is based on an initial vorticity field that can be represented with a few modes only. Instead, the vorticity field of Fig. 2 has a spectrum that spans a larger number of modes. This is key for our analysis because it allows us to better study the impact of the coarsening strategy based on spectral coefficient truncation (see Section 4.3.3) on the convergence rate upon temporal refinement. To our best

knowledge, this is the first time that a study of the effect of the coarsening strategy on the observed order of convergence of MLSDC is conducted.

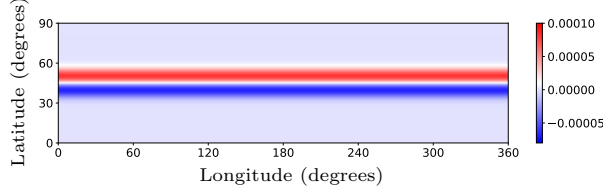


Figure 1: Steady zonal jet: vorticity field with a resolution of $R_f = S_f = 256$ after 144 hours. This solution is obtained with the single-level SDC(5,8). The diffusion coefficient is $\nu_B = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$.

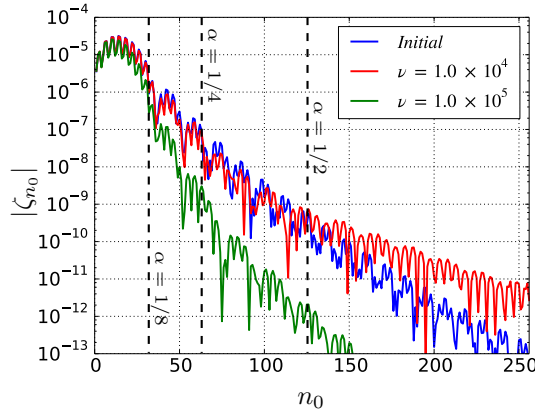


Figure 2: Steady zonal jet: max-spectrum of the vorticity field at the beginning of the simulation and after 144 hours for different values of the diffusion coefficient. The quantity on the y-axis is defined as $|\zeta_{n_0}| = \max_r |\zeta_{n_0}^r|$.

We perform a refinement study in time to assess the impact of the spatial coarsening ratio, α , on the observed order of convergence of the MLSDC-SH scheme. The study is done with a fixed fine resolution of $R_f = S_f = 256$. We consider two configurations, denoted by \mathcal{A} and \mathcal{B} , in which the diffusion coefficient is set to $\nu_{\mathcal{A}} = 1.0 \times 10^4 \text{ m}^2 \cdot \text{s}^{-1}$ and $\nu_{\mathcal{B}} = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$, respectively. In each configuration, the reference solution is obtained with SDC(5,8) using a time step size of $\Delta t_{ref} = 90 \text{ s}$. Fig. 3 shows the norm of the error in the vorticity field with respect to the reference solution as a function of the time step size. We use the L_∞ -norm to illustrate the connection between the convergence rate of MLSDC-SH upon refinement in time and the magnitude of the spectral coefficients of the vorticity that are truncated during the spatial restriction from the fine level to the coarse level. For this test case, we focus on MLSDC(3,2,2, α), which relies on three fine temporal nodes, two coarse temporal nodes, and uses two iterations (with one fine sweep and one coarse sweep per iteration). In both configurations, the observed order of convergence of MLSDC(3,2,2, α) varies significantly as a function of the spatial coarsening ratio, α .

In configuration \mathcal{A} , MLSDC(3,2,2,1/2) achieves fourth-order convergence upon refinement in time for stable time steps larger than 120 s, but exhibits only second-order convergence for shorter time steps. The reduction in the observed order of convergence can be explained by considering the vorticity spectrum of Fig. 2. Specifically, we note that for the time step size range defined by $\Delta t \leq 90 \text{ s}$, the MLSDC(3,2,2,1/2) scheme reaches an L_∞ -norm of the error smaller than 10^{-9} . We see in Fig. 2 that this threshold corresponds to the order of magnitude of the truncated terms during the restriction to the coarse level when $R_c = S_c = 128$. In MLSDC(3,2,2,1/4) and MLSDC(3,2,2,1/8), the truncated coefficients in the vorticity spectrum are relatively large which causes the observed convergence rate upon refinement in time to be reduced to second order in the entire range of stable time step sizes. Conversely, MLSDC(3,2,2,4/5) achieves the same convergence rate as SDC(3,4) over the full time step range (not shown here for brevity).

In configuration \mathcal{B} , the use of a larger diffusion coefficient significantly reduces the magnitude of the spectral coefficients associated with the high-frequency modes. Therefore, the MLSDC(3,2,2,1/2) scheme achieves fourth-order convergence in the entire time step range considered here. MLSDC(3,2,2,1/4)

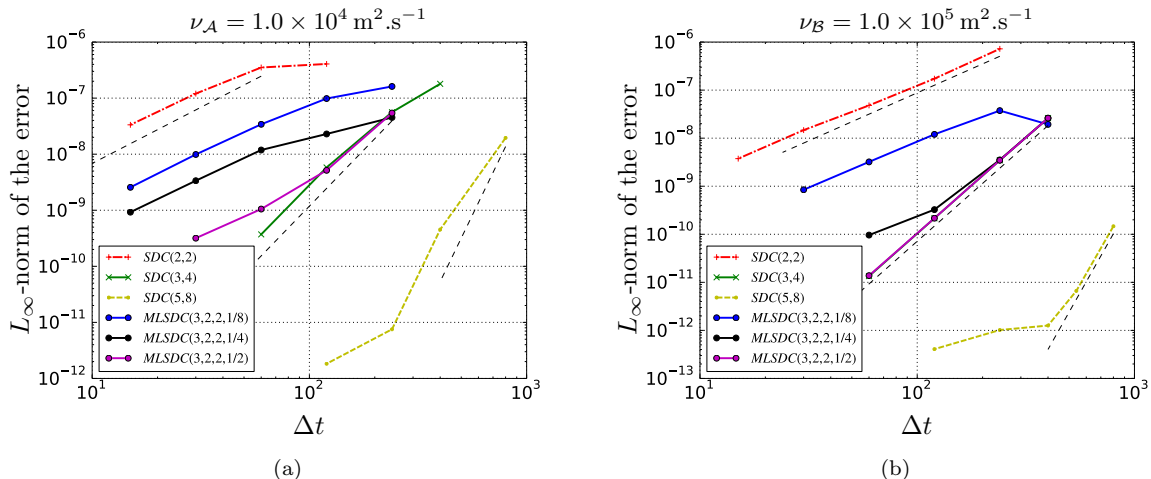


Figure 3: Steady zonal jet: L_∞ -norm of the error in the vorticity field with respect to the reference solution as a function of time step size. When the norm of the error is smaller than the magnitude of the spectral terms truncated during spatial coarsening (given in Fig. 2), MLSDC(3,2,2, α) exhibits only second-order convergence upon refinement in time. Above this threshold, we observe fourth-order convergence.

achieves fourth-order convergence in a larger fraction of the range of stable time step sizes, but still exhibits a reduction of its observed order of convergence when the norm of the error reaches the magnitude of the terms that are truncated during the restriction procedure. MLSDC(3,2,2,1/8) is still limited to second-order convergence. For this test case, the order of convergence of the MLSDC-SH scheme for the geopotential and divergence variables, not shown here, is similar to that observed for the vorticity.

The key insight of this section is that the accuracy of MLSDC-SH as the time step is reduced depends on the interplay between two key factors, namely the spectrum of the fine solution and the magnitude of the spatial coarsening ratio. They determine the range of scales of the fine solution that can be captured by the coarse correction, and as a result have a strong impact on the observed order of convergence of MLSDC-SH upon refinement in time. In particular, the presence of large high-frequency modes in the fine solution imposes a lower limit on the spatial coarsening ratio to preserve the high-order convergence of the multi-level scheme. Next, we study the computational cost of the MLSDC-SH scheme using unsteady test cases, starting with Gaussian dome propagation.

5.2. Propagation of a Gaussian dome

We now consider an initial condition derived from the third numerical experiment of Swartrauber (2004). The velocities are initially equal to zero ($u = v = 0$). We place a Gaussian dome in the initial geopotential field, such that

$$h(\lambda, \phi) = \bar{h} + Ae^{-\alpha(d/a)^2}, \quad (50)$$

where a denotes the Earth radius. The distance d is defined as

$$d = \sqrt{x^2 + y^2 + z^2}, \quad (51)$$

with

$$x = a(\cos(\lambda) \cos(\phi) - \cos(\lambda_c) \cos(\phi_c)), \quad (52)$$

$$y = a(\sin(\lambda) \cos(\phi) - \sin(\lambda_c) \cos(\phi_c)), \quad (53)$$

$$z = a(\sin(\phi) - \sin(\phi_c)). \quad (54)$$

This corresponds to a Gaussian dome centered at $\lambda_c = \pi$ and $\phi_c = \pi/4$. We use realistic values for the Earth radius, the gravitational acceleration, and the angular rate of rotation Ω involved in the Coriolis force. We set $\bar{h} = 29400$ m and $A = 6000$ m, which is about ten times larger than in the original test case. The simulation of the collapsing dome is run for one day to study the behavior of MLSDC-SH on

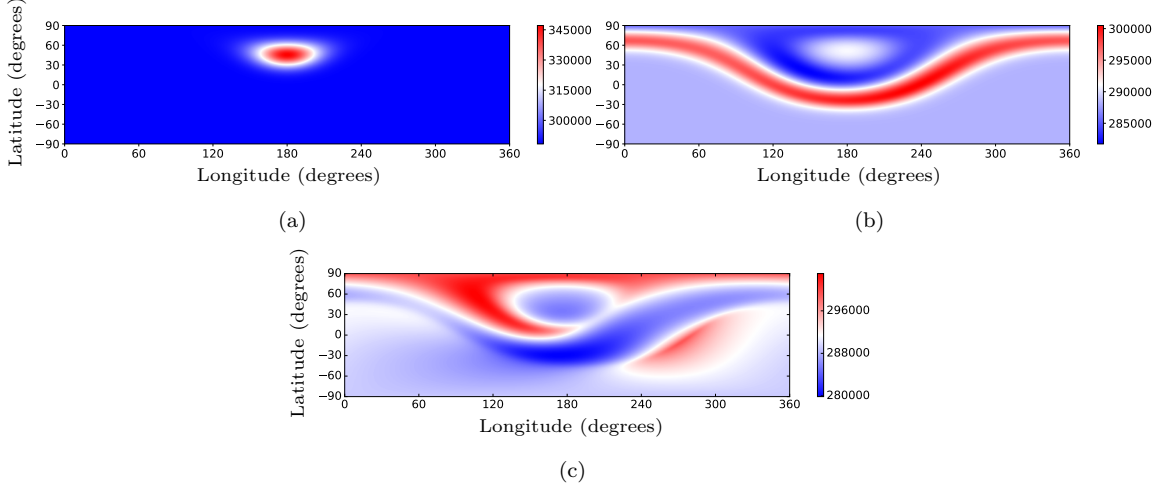


Figure 4: Gaussian dome: geopotential field with a resolution of $R_f = S_f = 256$ at the start of the simulation in 4(a), after 9000 s in 4(b), and after one day in 4(c). This solution is obtained with the single-level SDC(5,8). The diffusion coefficient is $\nu = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$.

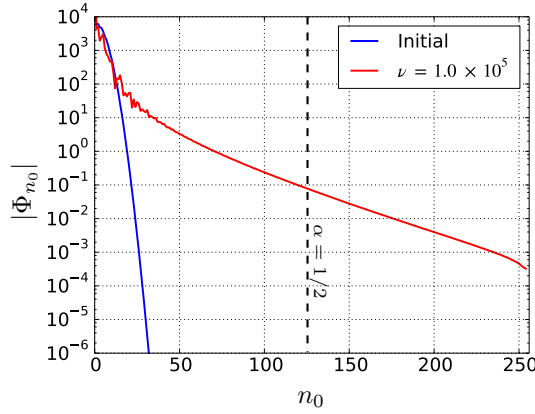


Figure 5: Gaussian dome: max-spectrum of the geopotential field at the beginning of the simulation and after one day. We use a diffusion coefficient $\nu = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$. The quantity on the y-axis is defined as $|\Phi_{n_0}| = \max_r |\Phi_{n_0}^r|$.

an advection-dominated test case. The geopotential field at different times is in Fig. 4, and the spectrum is in Fig. 5.

To assess the accuracy of the MLSDC-SH scheme, we perform a refinement study in time for a fixed spatial resolution ($R_f = S_f = 256$) over one day. The reference solution is obtained with the single-level SDC(5,8) and a time step size of $\Delta t_{ref} = 60 \text{ s}$. The diffusion coefficient is set to $\nu = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$. For this test case, we focus again on the geopotential and vorticity fields, because the results for the divergence field are qualitatively similar. For completeness, the study includes the results obtained with a second-order implicit-explicit Runge-Kutta scheme based on the temporal splitting (16) to (18). The L_∞ -norm of the error with respect to the reference solution, as a function of time step size, is shown in Fig. 6.

We see that MLSDC(3,2,2,1/2) achieves fourth-order convergence and an error of the same magnitude as that obtained with the single-level SDC(3,4) for time step sizes such that $\Delta t \geq 60 \text{ s}$. For a smaller time step size of 30s, the error induced by the truncation of high-frequency modes during coarsening reduces the observed convergence of MLSDC(3,2,2,1/2) to second order. As in the previous numerical example, the reduction in the observed convergence rate occurs when the L_∞ -norm of the error reaches the magnitude of the spectral coefficients truncated during spatial coarsening – about 10^{-1} for the

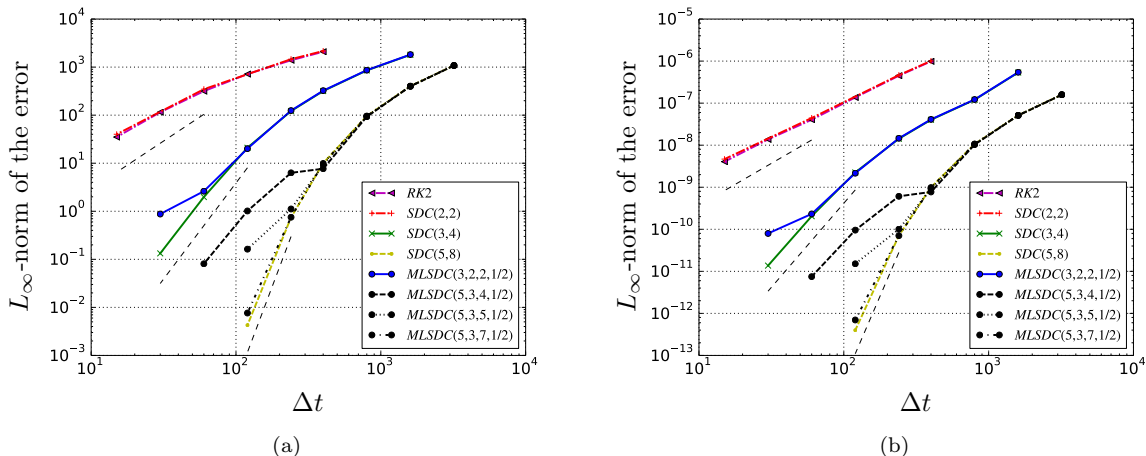


Figure 6: *Gaussian dome: L_∞ -norm of the error in the geopotential field in 6(a) and the vorticity field in 6(b) with respect to the reference solution as a function of time step size. $MLSDC(3,2,2,1/2)$ (respectively, $MLSDC(5,3,4,1/2)$) achieves the same order of convergence as the $SDC(3,4)$ (respectively, $SDC(5,8)$) for larger time steps. For smaller time steps, more $MLSDC$ -SH iterations are necessary to achieve the same order of convergence as the corresponding single-level schemes.*

geopotential according to Fig. 5. Numerical results not included for brevity indicate that this reduction in accuracy caused by spatial coarsening persists even in the absence of temporal coarsening ($M_f = M_c$). Still, in Fig. 6, the magnitude of the error obtained with $MLSDC(3,2,2,1/2)$ in this range of small time steps remains significantly smaller than that obtained with the single-level second-order $SDC(2,2)$. Fig. 6 also shows that $MLSDC(5,3,4,1/2)$ achieves the same convergence rate as the single-level $SDC(5,8)$ whenever $\Delta t \geq 400$ s. For smaller time step sizes, the observed convergence of $MLSDC(5,3,4,1/2)$ is reduced to fourth order. But, Fig. 6 demonstrates that doing more iterations with $MLSDC(5,3,7,1/2)$ is sufficient to recover eighth-order convergence in the asymptotic range.

To interpret these results, we distinguish two regimes in the temporal refinement study of Fig. 6. For very large time steps, large errors are caused by the fact that the fine and coarse corrections do not resolve the large temporal scales accurately and overstep the small scales present in the problem. Reducing the time step size in this regime reduces the errors associated with the large temporal scales. These scales can be resolved on both the coarse level and the fine level which explains why $MLSDC$ -SH and SDC converge at the same rate. The second regime starts for smaller time steps once these large scales have been resolved accurately. The error is then dominated by small-scale features that can be resolved by the fine correction but cannot be captured by the coarse correction due to its lower spatial resolution of the coarse problem. This undermines the observed order of convergence of $MLSDC$ -SH as the time step size becomes very small.

In Fig. 7, we investigate the computational cost the $MLSDC$ -SH scheme by measuring the wall-clock time of the simulations. $MLSDC(3,2,2,1/2)$ is more efficient than the $SDC(2,2)$ and $SDC(3,4)$ schemes for the time step sizes considered here despite the reduced observed order of convergence when $\Delta t \leq 60$ s. $MLSDC(5,3,4,1/2)$ is more efficient than $SDC(5,8)$ whenever $\Delta t \geq 120$ s. Below this time step size, its efficiency deteriorates slightly. Since the $MLSDC$ -SH scheme does not necessarily match the convergence rate of SDC in the simulations, we compute an observed speedup, \mathcal{S}^{obs} , as the ratio of the computational cost of SDC over that of $MLSDC$ -SH for a given error norm in Fig. 7. For an error norm of 10^0 in the geopotential field, $MLSDC(3,2,2,1/2)$ achieves an observed speedup $\mathcal{S}^{obs} \approx 1.58$ – i.e., a reduction of 37 % in wall-clock time – compared to $SDC(3,4)$. Considering that the cost of the FAS correction has been neglected in (49), this is close to the theoretical speedup $\mathcal{S}^{theo} \approx 1.66$ computed in Section 4. For the same magnitude of the error norm, $MLSDC(5,3,4,1/2)$ achieves an observed speedup $\mathcal{S}^{obs} \approx 1.50$ compared to $SDC(5,8)$, which represents a reduction of 33 % in wall-clock time. This is again relatively close to the theoretical speedup $\mathcal{S}^{theo} \approx 1.66$. Although this is not included in the figure for clarity, we point out that $MLSDC(5,3,5,1/2)$ and $MLSDC(5,3,7,1/2)$ are more accurate, but also more expensive

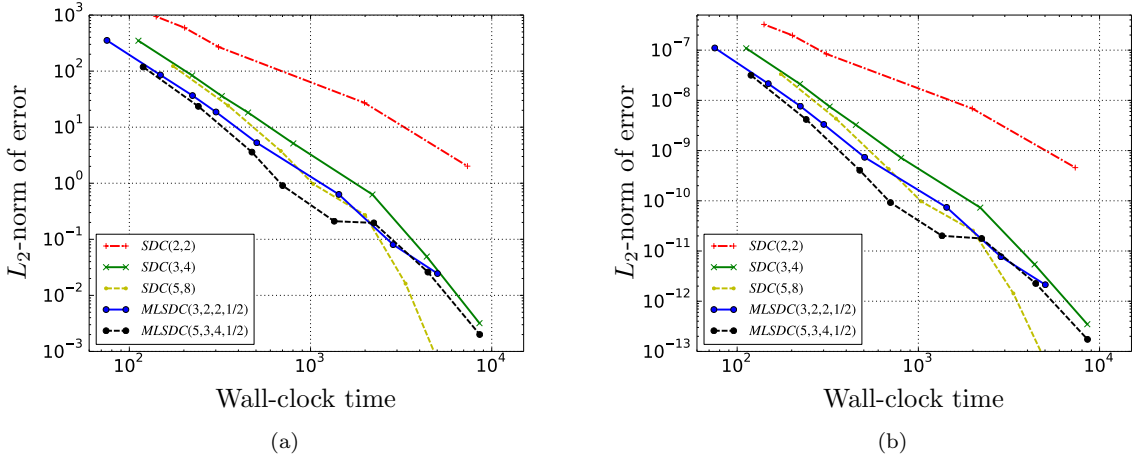


Figure 7: Gaussian dome: L_2 -norm of the error in the geopotential field in 7(a) and the vorticity field in 7(b) with respect to the reference solution as a function of the computational cost for the Gaussian dome test case. $MLSDC(3,2,2,1/2)$ is more efficient than $SDC(3,4)$ in the full time step range. $MLSDC(5,3,4,1/2)$ is more efficient than $SDC(5,8)$ when the norm of the error is above 10^{-1} for the geopotential, and above 10^{-11} for the vorticity. Below these thresholds, $MLSDC(5,3,4,1/2)$ only performs as efficiently as $MLSDC(3,2,2,1/2)$ because of the reduction in its observed order of convergence upon refinement in time.

than $MLSDC(5,3,4,1/2)$ in the range of time step sizes considered here.

5.3. Rossby-Haurwitz wave

In this section, we apply $MLSDC$ -SH to the Rossby-Haurwitz wave test case included in Williamson et al. (1992) and also considered in Jia et al. (2013). The initial analytically specified velocity field is non-divergent, and is computed with wavenumber 4. The initial geopotential field is obtained by solving the balance equation. The resulting Haurwitz pattern moves from east to west. We consider a fine resolution defined by $R_f = S_f = 256$ and we use a diffusion coefficient $\nu = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$. In Fig. 8 (respectively, Fig. 9), we show the solution (respectively, the vorticity spectrum) obtained with $SDC(5,8)$ after one day.

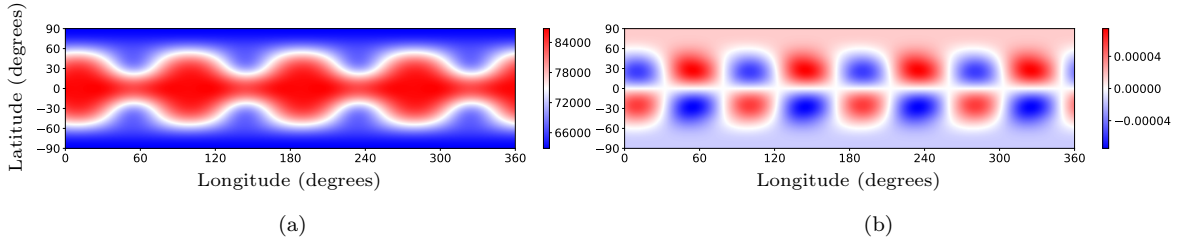


Figure 8: Rossby-Haurwitz wave: geopotential field in 8(a) and vorticity field in 8(b) with a resolution of $R_f = S_f = 256$ after one day. This solution is obtained with the single-level $SDC(5,8)$. The diffusion coefficient is $\nu = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$.

As in the previous sections, we carry out a refinement study in time using a reference solution obtained with $SDC(5,8)$ over one day using a time step size $\Delta t_{ref} = 120 \text{ s}$. The results, shown in Fig. 10, differ between the geopotential variable and the vorticity variable. Specifically, for the former, $MLSDC(3,2,2,1/2)$ achieves fourth-order convergence upon refinement in time and the same error magnitude as $SDC(3,4)$ for the full time step size range considered here. $MLSDC(5,3,4,1/2)$ is also more accurate than in the previous examples and reaches fifth-order convergence. But, for the vorticity variable, $MLSDC(3,2,2,1/2)$ exhibits a reduction in its convergence rate when $\Delta t \leq 120 \text{ s}$, which is slightly earlier than $SDC(3,4)$. Here again, this reduction is caused by the truncation of high-frequency modes during coarsening (see

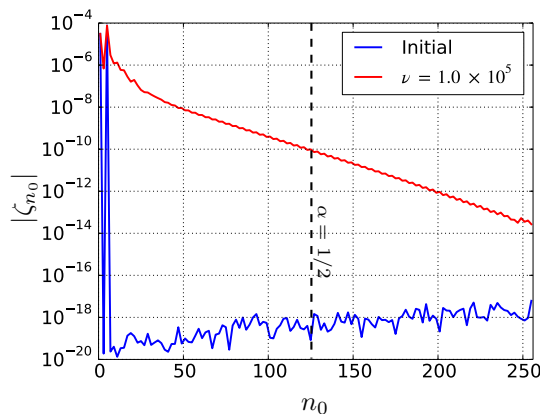


Figure 9: *Rossby-Haurwitz wave: max-spectrum of the vorticity field for the Rossby-Haurwitz wave test case at the beginning of the simulation and after one day. The quantity on the y-axis is defined as $|\zeta_{n_0}| = \max_r |\zeta_{n_0}^r|$. To simplify the figure, we only show $|\zeta_{n_0}|$ for the even n_0 . The odd n_0 correspond to negligible $|\zeta_{n_0}|$ and are therefore omitted.*

the vorticity spectrum in Fig. 9). MLSDC(5,3,4,1/2) is not in the asymptotic range and has already converged for the range of time step sizes considered here, which explains the flat line in Fig. 10(b).

In terms of wall-clock time, the most efficient scheme for both variables is MLSDC(5,3,4,1/2), as shown in Fig. 11. This multi-level scheme achieves a very low error norm for both variables while performing a large portion of the computations on the coarse level. MLSDC(3,2,2,1/2) is less efficient than MLSDC(5,3,4,1/2) but still more efficient than SDC(3,4) on the full range of time steps. In particular, for an error norm of 10^{-4} in the geopotential field, MLSDC(3,2,2,1/2) achieves an observed speedup $S^{obs} \approx 1.50$ compared to SDC(3,4), that is, a reduction in wall-clock time of 35 %. This is in good agreement with the theoretical speedup $S^{theo} \approx 1.66$ computed in Section 4.5. The MLSDC(3,2,2,1/2) performance deteriorates for the vorticity variable for smaller time steps and the speedup compared to SDC(3,4) decreases because of the reduction in its observed order of convergence. For an error norm of 10^{-12} in the vorticity field, the observed speedup is $S^{obs} \approx 1.50$, but it is reduced to $S^{obs} \approx 1.13$ for an error norm of 10^{-14} . Next, we conclude the analysis of MLSDC-SH with a challenging unsteady test case representative of atmospheric flows.

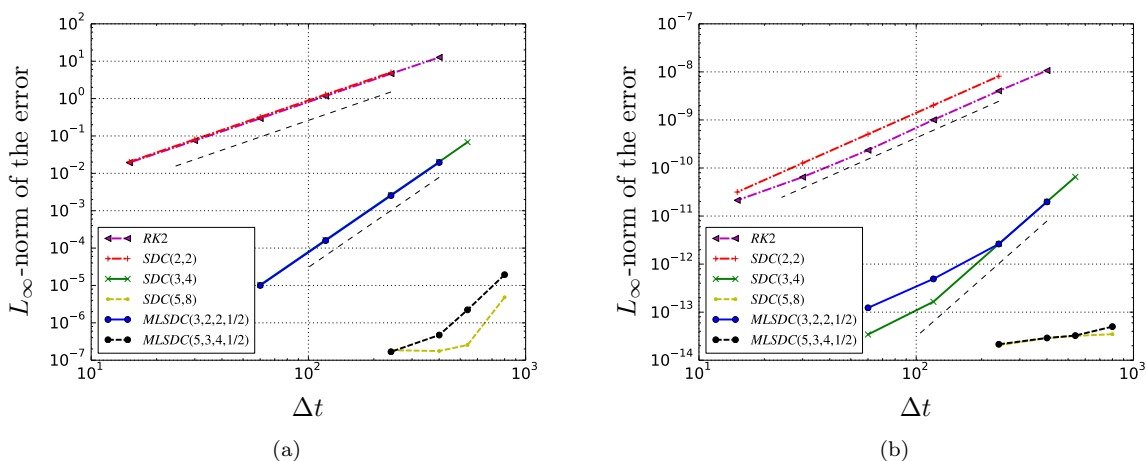


Figure 10: *Rossby-Haurwitz wave: L_∞ -norm of the error in the geopotential field in 10(a) and the vorticity field in 10(b) with respect to the reference solution as a function of time step size. For this example, MLSDC(3,2,2,1/2) (respectively, MLSDC(5,3,4,1/2)) achieves the same observed order of convergence as the single-level SDC(3,4) (respectively, SDC(5,8)).*

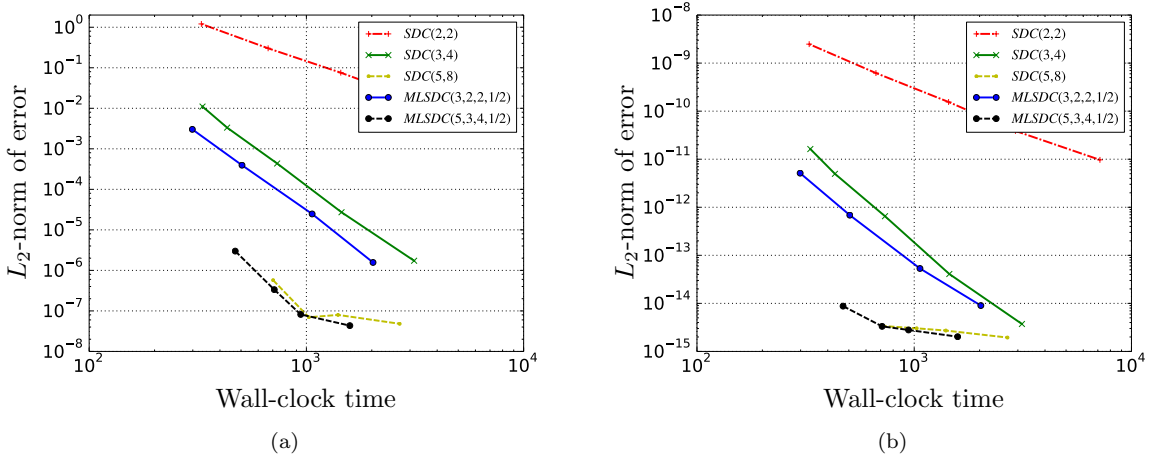


Figure 11: Rossby-Haurwitz wave: L_2 -norm of the error in the geopotential field in 11(a) and the vorticity field in 11(b) with respect to the reference solution as a function of the computational cost. The most efficient scheme is $MLSDC(5,3,4,1/2)$. We also note that $MLSDC(3,2,2,1/2)$ is more efficient than the single-level $SDC(3,4)$ for the range of time step sizes considered here.

5.4. Nonlinear evolution of an unstable barotropic wave

In this section, we consider the barotropic instability test case proposed in Galewsky et al. (2004). This is done by introducing a localized bump in the height field to perturb the balanced state described in Section 5.1. The perturbation first triggers the development of gravity waves and then leads to the formation of complex vortical dynamics. These processes operate on multiple time scales and are representative of the horizontal features of atmospheric flows. We run the simulations using two configurations, \mathcal{B} and \mathcal{C} , based on a diffusion coefficient $\nu_{\mathcal{B}} = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$ – as in Galewsky et al. (2004) – and $\nu_{\mathcal{C}} = 2.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$, respectively. The reference solutions for the refinement studies detailed below are obtained with $SDC(5,8)$ with a time step size $\Delta t_{ref} = 60 \text{ s}$. In Jia et al. (2013), the largest time step size used in the single-level SDC scheme combined with the Spectral Element Method (SEM) based on 24 elements along each cube edge and a polynomial basis of degree seven is 1200 s, which is the same order of magnitude as the largest time step size used here. The vorticity fields after 122 hours and 144 hours for configuration \mathcal{B} are shown in Fig. 12. Fig. 13 presents the corresponding spectrum of the vorticity field at the same times.

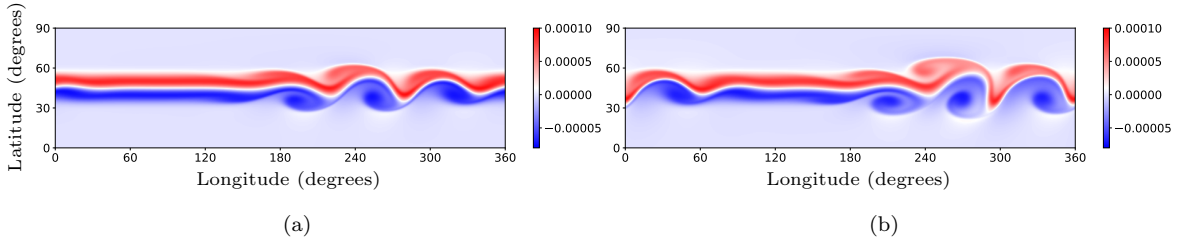


Figure 12: Unstable barotropic wave: vorticity field with a resolution of $R_f = S_f = 256$ after 120 hours in 12(a) and 144 hours in 12(b). This solution is obtained with the single-level $SDC(5,8)$. The diffusion coefficient is $\nu_{\mathcal{B}} = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$.

As in Section 5.1, we first highlight the connection between the spectrum of the vorticity field and the observed order of convergence of the MLSDC-SH scheme upon refinement in time. This is done with $MLSDC(3,2,2,1/2)$ – that is, MLSDC-SH with three nodes on the fine level, two nodes on the coarse level, two iterations, and $R_c = S_c = 128$ – in the refinement study in time shown in Fig. 14. When $\nu_{\mathcal{B}} = 1.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$, the magnitude of the truncated terms in the vorticity spectrum is of the order of 10^{-8} (see Fig. 13). Since this is also the order of the L_∞ -norm of the error for the largest stable time step ($\Delta t = 400 \text{ s}$), $MLSDC(3,2,2,1/2)$ achieves only second-order convergence for the range of time step sizes

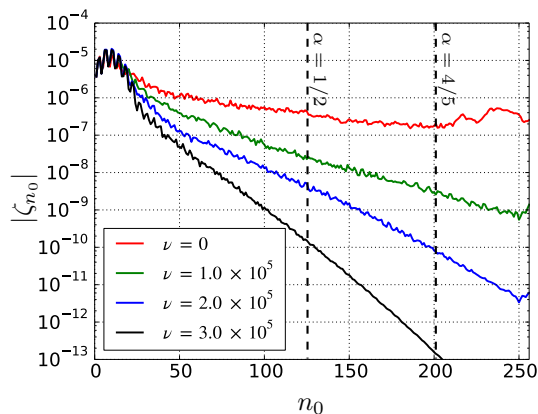


Figure 13: *Unstable barotropic wave: max-spectrum of the vorticity field after 144 hours for various values of the diffusion coefficient. The quantity on the y-axis is defined as $|\zeta_{n_0}| = \max_r |\zeta_{n_0}^r|$. We observe a strong damping of the high-frequency modes when the diffusion coefficient is large. This has a significant impact on the observed order of convergence of MLSDC-SH as shown in Fig. 14.*

considered here. With $\nu_c = 2.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$, the magnitude of the truncated terms is of the order of 10^{-9} , and MLSDC(3,2,2,1/2) reaches fourth-order convergence until this threshold is reached for $\Delta t = 320 \text{ s}$. For completeness, we have run the same test with $\nu = 3.0 \times 10^5 \text{ m}^2 \cdot \text{s}^{-1}$, in which case this threshold is lower, which allows MLSDC(3,2,2,1/2) to exhibit fourth-order convergence for $\Delta t \geq 120 \text{ s}$. In the following paragraphs, we show that this reduction in the observed order of convergence can be overcome by doing additional MLSDC-SH iterations. For instance, we demonstrate that MLSDC(3,2,3,1/2) recovers fourth-order convergence in configurations \mathcal{B} and \mathcal{C} .

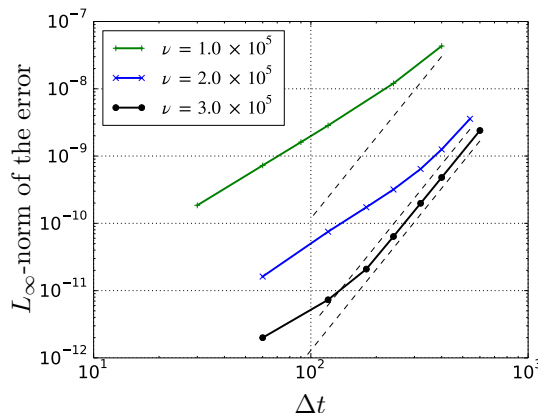


Figure 14: *Unstable barotropic wave: L_∞ -norm of the error in the vorticity field with respect to the reference solution as a function of the time step size. These results are obtained with MLSDC(3,2,2,1/2). We note again that a reduction in the observed order of convergence of MLSDC-SH occurs when the norm of the error is smaller than the magnitude of the spectral coefficients that are truncated during spatial coarsening (see Fig. 13).*

We now use this knowledge of the spectrum of the vorticity field to motivate our choice of the spatial coarsening ratio in each configuration. The goal is to make the coarse sweeps as inexpensive as possible without undermining the observed order of convergence of the MLSDC-SH scheme upon refinement in time. In configuration \mathcal{B} , we choose a relatively modest spatial coarsening ratio $\alpha_{\mathcal{B}} = 4/5$ to account for the presence of large spectral coefficients associated with the high-frequency modes. At the coarse level, this choice yields $R_c^{\mathcal{B}} = S_c^{\mathcal{B}} = 204$. In configuration \mathcal{C} , we can choose a more aggressive coarsening strategy with $\alpha_{\mathcal{C}} = 1/2$, leading to $R_c^{\mathcal{C}} = S_c^{\mathcal{C}} = 128$ as in the previous test cases. These choices are such that the magnitude of the truncated spectral vorticity coefficients have the same order of magnitude, that is $|\zeta_{204}^{\mathcal{B}}| \approx |\zeta_{128}^{\mathcal{C}}| \approx 3 \times 10^{-9}$ in Fig. 13. The results of the refinement study in L_∞ -norm are shown

in Fig. 15 for the geopotential. The asymptotic rates observed for the divergence and the vorticity are qualitatively similar to those of the geopotential and are therefore omitted for brevity.

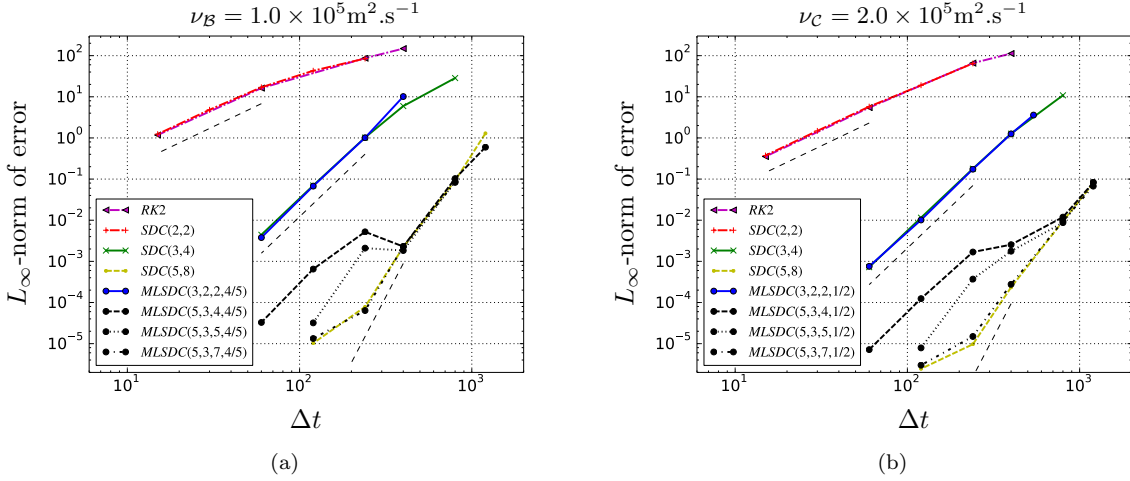


Figure 15: Unstable barotropic wave: L_∞ -norm of the error in the geopotential field with respect to the reference solution as a function of the time step size. $MLSDC(3,2,2,\alpha)$ achieves the same order of convergence as $SDC(3,4)$ in both configurations. $MLSDC(5,3,4,\alpha)$ is also more accurate than $SDC(3,4)$, but achieves the same accuracy as $SDC(5,8)$ for larger time steps only. Doing additional $MLSDC$ -SH iterations – in this case, seven iterations with $MLSDC(5,3,7,\alpha)$ – is needed to match the accuracy of $SDC(5,8)$ in the full range of time step sizes.

In Fig. 15, $MLSDC(3,2,2,\alpha)$ exhibits the same observed order of convergence and error magnitude as $SDC(3,4)$ for both diffusion configurations. $MLSDC(5,3,4,\alpha)$ also converges at a fourth-order rate in the asymptotic range, but achieves a significantly smaller error magnitude than $MLSDC(3,2,2,\alpha)$. $MLSDC(5,3,4,\alpha)$ is as accurate as $SDC(5,8)$ for larger time step sizes. But, to achieve the same observed order of convergence as $SDC(5,8)$ in the entire range of time step sizes, seven iterations – with $MLSDC(5,3,7,\alpha)$ – are needed. Finally, we note that numerical examples not shown here for brevity confirm that the observed order of convergence of $MLSDC$ -SH increases significantly for a larger coefficient α – i.e., less aggressive spatial coarsening –, but this also drastically increases the computational cost.

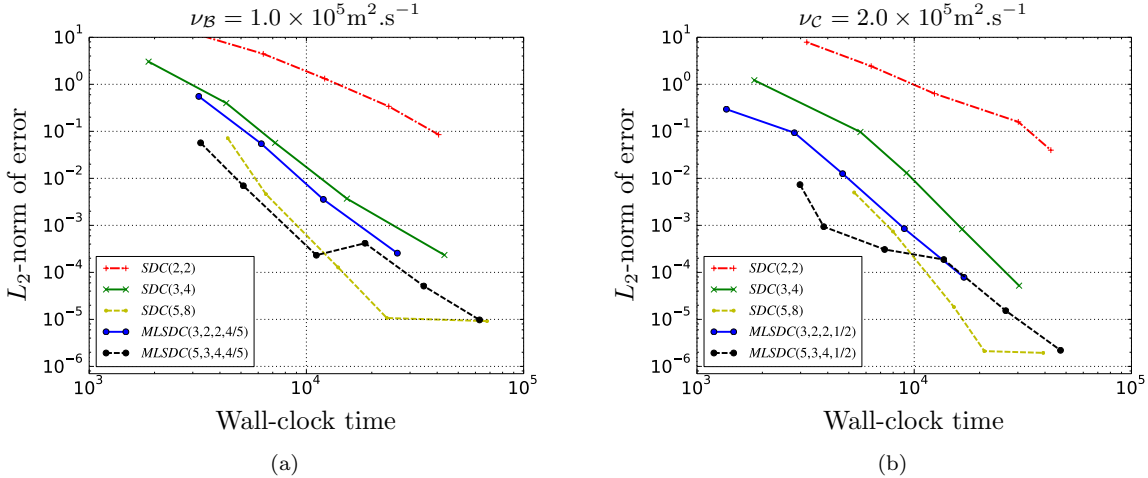


Figure 16: Unstable barotropic wave: L_2 -norm of the error in the geopotential field with respect to the reference solution as a function of the computational cost. $MLSDC(3,2,2,\alpha)$ is more efficient than $SDC(3,4)$. The cost reduction is larger when the spatial coarsening is more aggressive ($\alpha = 1/2$). $MLSDC(5,3,4,\alpha)$ is more efficient than $SDC(5,8)$ for larger error magnitudes. In the range of relatively larger errors, $MLSDC(5,3,4,\alpha)$ is the most efficient scheme among those considered here.

In Fig. 16, we show the L_2 -norm of the error as a function of the wall-clock time of the simulations

for the geopotential. We see that $\text{MLSDC}(3,2,2,\alpha)$ is significantly less expensive than $\text{SDC}(3,4)$ in the full range of time step sizes. This cost reduction is larger with $\text{MLSDC}(3,2,2,\alpha_{\mathcal{C}})$ since configuration \mathcal{C} allows for a more aggressive spatial coarsening strategy than configuration \mathcal{B} . In both configurations, the observed speedup of $\text{MLSDC}(3,2,2,\alpha)$ compared to $\text{SDC}(3,4)$ is close to the theoretical speedup. Specifically, in configuration \mathcal{B} , the observed speedup is $\mathcal{S}_{\mathcal{B}}^{obs} \approx 1.28$ for an error norm of 3×10^{-3} in the geopotential field whereas the theoretical speedup – obtained with (49) evaluated with $\alpha_{\mathcal{B}} = 4/5$ – is $\mathcal{S}_{\mathcal{B}}^{theo} \approx 1.30$. In configuration \mathcal{C} , $\text{MLSDC}(3,2,2,1/2)$ achieves $\mathcal{S}_{\mathcal{C}}^{obs} \approx 1.56$ for an error norm of 8×10^{-5} in the geopotential, for a theoretical speedup $\mathcal{S}_{\mathcal{C}}^{theo} \approx 1.66$. $\text{MLSDC}(5,3,4,\alpha)$ is the most efficient scheme for relatively large error magnitudes and also achieves observed speedups close to the theoretical speedup. But, the performance of $\text{MLSDC}(5,3,4,\alpha)$ deteriorates for lower error magnitudes. We found that doing additional iterations, for instance with $\text{MLSDC}(3,2,3,\alpha)$ or $\text{MLSDC}(5,3,5,\alpha)$, does not improve the efficiency of MLSDC-SH .

6. Conclusions and future work

We have studied a high-order implicit-explicit iterative multi-level time integration scheme for the nonlinear shallow-water equations on the rotating sphere. Our algorithm relies on the Multi-Level Spectral Deferred Corrections (MLSDC) scheme of Emmett and Minion (2012); Speck et al. (2015) combined with a spatial discretization performed with the global Spherical Harmonics (SH) transform. MLSDC-SH applies a sequence of updates distributed on a hierarchy of space-time levels obtained by coarsening the problem in space and in time. This approach makes it possible to shift a significant portion of the computational work to the coarse representation of the problem to reduce the time-to-solution while preserving accuracy.

We have discussed the requirements of consistent inter-level transfer operators which play a crucial role in MLSDC-SH . Our approach consists in exploiting the canonical basis of the multi-level scheme. This SH-based algorithm leads to restriction and interpolation procedures performed in spectral space to transfer the solution between different spatio-temporal levels. The proposed restriction and interpolation methods do not introduce spurious modes that would, driven by nonlinear interactions, propagate across the spectrum. Our results show that this is one the key features needed to obtain an efficient MLSDC-SH scheme. The development of restriction and interpolation operators for other non-global spatial discretization schemes is left for future work.

We have shown that MLSDC-SH is efficient for the nonlinear wave-propagation-dominated problems arising from the discretized Shallow-Water Equations (SWE) on the rotating sphere. Our numerical studies are based on challenging test cases that are representative of the horizontal effects present in the full atmospheric dynamics. With a steady zonal jet test case, we have first examined the impact of the coarsening strategy on the observed accuracy of MLSDC-SH upon refinement in time. Then, using unsteady numerical examples, we have shown that MLSDC-SH can achieve up to eighth-order convergence upon refinement in time, and that MLSDC-SH can take stable time steps that are as large as those of the single-level SDC schemes. We have also demonstrated that MLSDC-SH is more efficient than the single-level SDC schemes, and in particular requires fewer function evaluations. Our results show that MLSDC-SH can reduce the wall-clock time of the simulations by up to 37% compared to single-level SDC schemes.

As a final note, we mention here that MLSDC is one of the key building blocks of the Parallel Full Approximation Scheme in Space and in Time (PFASST). The present work therefore lays the foundations of a parallel-in-time integration of the full shallow-water equations on the sphere with PFASST.

7. Code availability

The code used to generate the simulations presented here is publicly available (Schreiber, 2018).

8. Acknowledgements

The work of François Hamon and Michael Minion was supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research, Applied Mathematics program under

contract number DE-AC02005CH11231. Part of the simulations were performed using resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231. Martin Schreiber gratefully acknowledges support from the Computational and Information Systems Laboratory (CISL) Visitor Program at the National Center for Atmospheric Research in Boulder, CO. We thank Andreas Kreienbuehl for discussions on his work on MLSDC and PFASST using the Spectral Element Method.

References

- Bolten, M., Moser, D., and Speck, R. (2017). A multigrid perspective on the parallel full approximation scheme in space and time. *Numerical Linear Algebra with Applications*, 24(6):e2110–n/a.
- Bourke, W. (1972). An efficient, one-level, primitive-equation spectral model. *Monthly Weather Review*, 100(9):683–689.
- Bourlioux, A., Layton, A. T., and Minion, M. L. (2003). High-order multi-implicit spectral deferred correction methods for problems of reactive flow. *Journal of Computational Physics*, 189(2):651–675.
- Brandt, A. (1977). Multi-level adaptive solutions to boundary-value problems. *Mathematics of computation*, 31(138):333–390.
- Christlieb, A., Ong, B., and Qiu, J.-M. (2009). Comments on high-order integrators embedded within integral deferred correction methods. *Communications in Applied Mathematics and Computational Science*, 4(1):27–56.
- Crockatt, M. M., Christlieb, A. J., Garrett, C. K., and Hauck, C. D. (2017). An arbitrary-order, fully implicit, hybrid kinetic solver for linear radiative transport using integral deferred correction. *Journal of Computational Physics*, 346:212–241.
- Durran, D. R. and Blossey, P. N. (2012). Implicit–explicit multistep methods for fast-wave–slow-wave problems. *Monthly Weather Review*, 140(4):1307–1325.
- Dutt, A., Greengard, L., and Rokhlin, V. (2000). Spectral deferred correction methods for ordinary differential equations. *BIT Numerical Mathematics*, 40(2):241–266.
- Emmett, M. and Minion, M. L. (2012). Toward an efficient parallel in time method for partial differential equations. *Communications in Applied Mathematics and Computational Science*, 7(1):105–132.
- Evans, K. J., Taylor, M. A., and Drake, J. B. (2010). Accuracy analysis of a spectral element atmospheric model using a fully implicit solution framework. *Monthly Weather Review*, 138(8):3333–3341.
- Falgout, R. D., Friedhoff, S., Kolev, T. V., MacLachlan, S. P., and Schroder, J. B. (2014). Parallel time integration with multigrid. *SIAM Journal on Scientific Computing*, 36(6):C635–C661.
- Galewsky, J., Scott, R. K., and Polvani, L. M. (2004). An initial-value problem for testing numerical models of the global shallow-water equations. *Tellus A*, 56(5):429–440.
- Gardner, D. J., Guerra, J. E., Hamon, F. P., Reynolds, D. R., Ullrich, P. A., and Woodward, C. S. (2018). Implicit–explicit (IMEX) Runge–Kutta methods for non-hydrostatic atmospheric models. *Geoscientific Model Development*, 11(4):1497.
- Giraldo, F. X. (2005). Semi-implicit time-integrators for a scalable spectral element atmospheric model. *Quarterly Journal of the Royal Meteorological Society*, 131(610):2431–2454.
- Giraldo, F. X., Kelly, J. F., and Constantinescu, E. M. (2013). Implicit-explicit formulations of a three-dimensional nonhydrostatic unified model of the atmosphere (NUMA). *SIAM Journal on Scientific Computing*, 35(5):B1162–B1194.
- Hack, J. J. and Jakob, R. (1992). *Description of a global shallow water model based on the spectral transform method*. National Center for Atmospheric Research.
- Hagstrom, T. and Zhou, R. (2007). On the spectral deferred correction of splitting methods for initial value problems. *Communications in Applied Mathematics and Computational Science*, 1(1):169–205.
- Hamon, F. P., Day, M. S., and Minion, M. L. (forthcoming). Concurrent implicit spectral deferred correction scheme for low-Mach number combustion with detailed chemistry. *Combustion Theory and Modelling*. <http://dx.doi.org/10.1080/13647830.2018.1524156>.
- Jia, J., Hill, J. C., Evans, K. J., Fann, G. I., and Taylor, M. A. (2013). A spectral deferred correction method applied to the shallow water equations on a sphere. *Monthly Weather Review*, 141(10):3435–3449.

- Kanamitsu, M., Tada, K., Kudo, T., Sato, N., and Isa, S. (1983). Description of the JMA operational spectral model. *Journal of the Meteorological Society of Japan. Ser. II*, 61(6):812–828.
- Layton, A. T. and Minion, M. L. (2004). Conservative multi-implicit spectral deferred correction methods for reacting gas dynamics. *Journal of Computational Physics*, 194(2):697–715.
- Lions, J.-L., Maday, Y., and Turinici, G. (2001). Résolution d’EDP par un schéma en temps pararéel. *Comptes Rendus de l’Académie des Sciences-Series I-Mathematics*, 332(7):661–668.
- Lock, S.-J., Wood, N., and Weller, H. (2014). Numerical analyses of Runge–Kutta implicit–explicit schemes for horizontally explicit, vertically implicit solutions of atmospheric models. *Quarterly Journal of the Royal Meteorological Society*, 140(682):1654–1669.
- Lott, P. A., Woodward, C. S., and Evans, K. J. (2015). Algorithmically scalable block preconditioner for fully implicit shallow-water equations in CAM-SE. *Computational Geosciences*, 19(1):49–61.
- Minion, M. L. (2003). Semi-implicit spectral deferred correction methods for ordinary differential equations. *Communications in Mathematical Sciences*, 1(3):471–500.
- Rivier, L., Loft, R., and Polvani, L. M. (2002). An efficient spectral dynamical core for distributed memory computers. *Monthly Weather Review*, 130(5):1384–1396.
- Robert, A., Henderson, J., and Turnbull, C. (1972). An implicit time integration scheme for baroclinic models of the atmosphere. *Monthly Weather Review*, 100(5):329–335.
- Ruprecht, D. and Speck, R. (2016). Spectral deferred corrections with fast-wave slow-wave splitting. *SIAM Journal on Scientific Computing*, 38(4):A2535–A2557.
- Schaeffer, N. (2013). Efficient spherical harmonic transforms aimed at pseudospectral numerical simulations. *Geochemistry, Geophysics, Geosystems*, 14(3):751–758.
- Schreiber, M. (2018). Shallow-Water Equations Environment for Tests (SWEET). <https://github.com/schreiberx/sweet>.
- Schreiber, M. and Loft, R. (forthcoming). A parallel time-integrator for solving the linearized shallow water equations on the rotating sphere. *Numerical Linear Algebra with Applications*.
- Speck, R., Ruprecht, D., Emmett, M., Minion, M. L., Bolten, M., and Krause, R. (2015). A multi-level spectral deferred correction method. *BIT Numerical Mathematics*, 55(3):843–867.
- Swarztrauber, P. N. (2004). Shallow water flow on the sphere. *Monthly Weather Review*, 132(12):3010–3018.
- Temperton, C. (1991). On scalar and vector transform methods for global spectral models. *Monthly Weather Review*, 119(5):1303–1307.
- Ullrich, P. and Jablonowski, C. (2012). Operator-split Runge–Kutta–Rosenbrock methods for nonhydrostatic atmospheric models. *Monthly Weather Review*, 140(4):1257–1284.
- Wedi, N. P., Hamrud, M., and Mozdzyński, G. (2013). A fast spherical harmonics transform for global NWP and climate models. *Monthly Weather Review*, 141(10):3450–3461.
- Weiser, M. (2015). Faster SDC convergence on non-equidistant grids by DIRK sweeps. *BIT Numerical Mathematics*, 55(4):1219–1241.
- Weller, H., Lock, S.-J., and Wood, N. (2013). Runge–Kutta IMEX schemes for the horizontally explicit/vertically implicit (HEVI) solution of wave equations. *Journal of Computational Physics*, 252:365–381.
- Williamson, D. L., Drake, J. B., Hack, J. J., Jakob, R., and Swarztrauber, P. N. (1992). A standard test set for numerical approximations to the shallow water equations in spherical geometry. *Journal of Computational Physics*, 102(1):211–224.
- Winkel, M., Speck, R., and Ruprecht, D. (2015). A high-order Boris integrator. *Journal of Computational Physics*, 295:456–474.
- Xia, Y., Xu, Y., and Shu, C. W. (2007). Efficient time discretization for local discontinuous Galerkin methods. *Discrete and Continuous Dynamical Systems Series B*, 8(3):677.