The Flux-Differencing Discontinuous Galerkin Method Applied to an Idealized Fully Compressible Nonhydrostatic Dry Atmosphere

A. N. Souza[1], J. He[2], T. Bischoff[2], M. Waruszewski[3], L. Novak[2], V. Barra[2], T. Gibson[4], A. Sridhar[2], S. Kandala[2], S. Byrne[2], L. C. Wilcox[5], J. Kozdon[5], F. X. Giraldo[5], O. Knoth[6], J. Marshall[1], R. Ferrari[1], T. Schneider[2]

[1]Massachusetts Institute of Technology, Cambridge, Massachusetts, United States
[2]California Institute of Technology, Pasadena, California, United States
[3]Sandia National Laboratories, Albuquerque, New Mexico, United States
[4]University of Illinois Urbana-Champaign, Urbana and Champaign, Illinois, United States
[5]Naval Postgraduate School, Monterey, California, United States
[6]Leibniz Institute for Tropospheric Research, Leipzig, Saxony, Germany

Key Points:

- The Flux-Differencing Discontinuous Galerkin (FDDG) method offers a robust way to construct numerical discretizations in geophysically relevant configurations.
- FDDG allows for a computationally stable total energy formulation of the compressible Euler equations with gravity and rotation.
- FDDG simulates a dry convective boundary layer and the atmospheric general circulation without additional dissipation such as those given by diffusion or hyperdiffusion.

Corresponding author: A. N. Souza, `andrenogueirasouza@gmail.com`

Abstract

Dynamical cores used to study the circulation of the atmosphere employ various numerical methods ranging from finite-volume, spectral element, global spectral, and hybrid methods. In this work, we explore the use of Flux-Differencing Discontinuous Galerkin (FDDG) methods to simulate a fully compressible dry atmosphere at various resolutions. We show that the method offers a judicious compromise between high-order accuracy and stability for large-eddy simulations and simulations of the atmospheric general circulation. In particular, filters, divergence damping, diffusion, hyperdiffusion, or sponge-layers are not required to ensure stability; only the numerical dissipation naturally afforded by FDDG is necessary. We apply the method to the simulation of dry convection in an atmospheric boundary layer and in a global atmospheric dynamical core in the standard benchmark of Held and Suarez (1994).

Plain Language Summary

Numerical models cannot explicitly represent all degrees of freedom that characterize atmospheric flows due to limitations in computing power. One must allocate the available computational degrees of freedom to reduce the degradation of the solution. In this work, we explore the use of the discontinuous Galerkin numerical method, a hybrid approach that combines the accuracy of spectral methods with the flexibility of finite volume methods. We apply it to idealized dry atmospheric simulations and show that the method is robust and incorporates physical principles to best account for unresolved processes.

1 Introduction

Designing dynamical cores that meet the challenges imposed by simulating the continuous equations that govern geophysical flows has a long history (Williamson, 2007). Various numerical methods are employed to achieve accuracy, efficiency, and stability. However, careful compromises are required because these goals are often in conflict: significant dissipation helps with stability at the cost of accuracy, and high-order schemes deliver accuracy at the expense of computing cost. This work explores the discontinuous Galerkin (DG) method for simulating atmospheric motions. The goal is not to provide an in-depth introduction to the method. For this, there are excellent references (Hesthaven & Warburton, 2007; Winters et al., 2021; G. J. Gassner & Winters, 2021), which illustrate the method in the context of weak formulations of partial differential equations, finite-element and spectral methods, and discrete algebraic properties of numerical operators. Instead, we focus on compromises that achieve stable and accurate atmospheric solutions.

The DG method is similar to finite volume methods since both use a discontinuous function space to approximate a partial differential equation. However, functions are not approximated as piecewise constants within a control volume, as in finite-volume methods, but as piecewise polynomials whose shape is chosen to achieve high-order accuracy.[1]. The method is a generalization of finite volume methods with some flavor of spectral methods as it uses a function basis within each control volume (see e.g. Karniadakis & Sherwin, 2005; Hesthaven & Warburton, 2007; D. Kopriva, 2009; F. X. Giraldo, 2020). Figure 1 shows how increasing the polynomial order improves the DG approximation of a sinusoidal function. Notice the discontinuities at the control volume edges, which are a signature of the DG representation.

---

[1] Interestingly, the convergence of the mean value in each control volume has been shown by Cao et al. (2015) to exhibit superconvergence properties in special cases. For example, if the numerical scheme uses 6th-order polynomials, then the control volume-mean converges at a 13th-order rate.

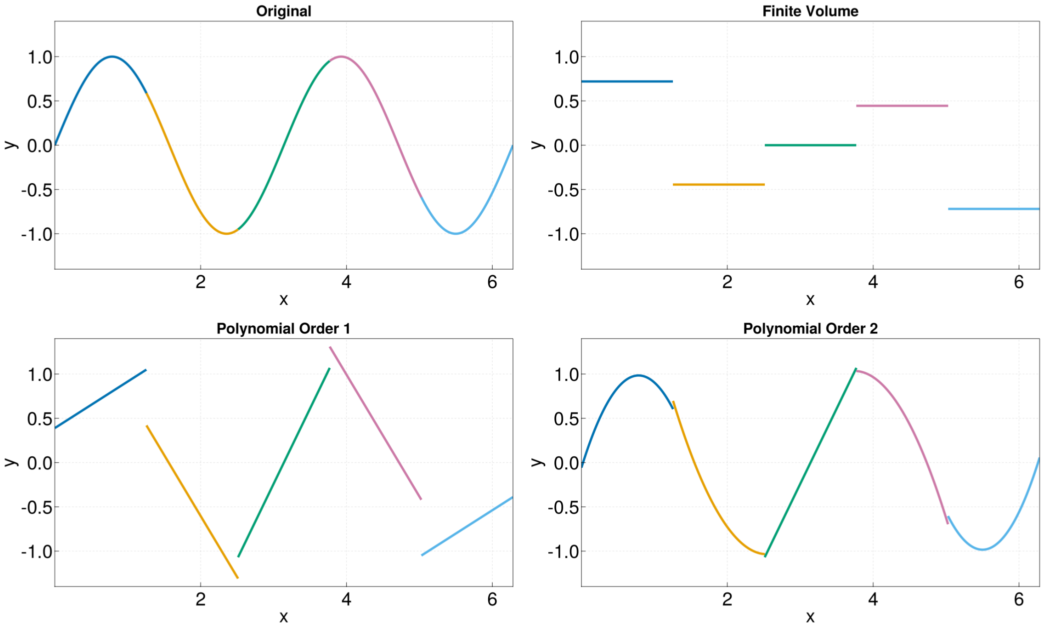

Figure 1. Projection of a function onto spaces of polynomials. The different colors represent different control volumes. The original function $y = \sin(x)$ is in the top left panel, and various projections are shown in the other panels. Note the discontinuities at the edges of the control volumes in all projections.

Finite volume flux-reconstruction methods can be applied to DG within the control volume and at the discontinuities between elements (Fisher & Carpenter, 2013). The flexibility of choosing a “volume” numerical flux and the usual “interface” numerical flux is leveraged to yield robust numerical simulations. Departing from standard practice to use central fluxes for the volume terms, we demonstrate choices among a new class of schemes, known as Flux-Differencing Discontinuous Galerkin (FDDG) methods (Winters et al., 2021), which provide the numerical stability and accuracy necessary for geophysical fluid dynamics applications, in which the flows in question are usually strongly underresolved. The resulting spatial discretization is different from other DG methods that have been applied to geophysical flows such as those of F. Giraldo et al. (2002) or Nair et al. (2005). What follows is along a new thread of methods, e.g. G. Gassner et al. (2015).

We rely on recent theoretical advancements in the formulation of FDDG methods. FDDG methods retain stability without needing additional diffusion, hyperdiffusion, or other numerical filters to guarantee stability. Instead, the numerical dissipation comes directly from the formulation of the numerical flux and the time-stepping method.

Of course, it is not always desirable to leave all dissipation to the numerical method itself; however, such implicit dissipation can be a desirable feature if the numerical dissipation mimics that owing to missing physical information and otherwise is minimal where information loss is minimal (Pressel et al., 2017). As a counterpoint, see, for example, Boyd (2001) for numerous reasons why it can be better to create a well-posed mathematical problem and use an optimally convergent numerical method. A robust numerical method saves human time since it is common in geophysical simulations to include the minimal necessary dissipation for stability; see Winters et al. (2021) for comments with respect to engineering and astrophysical examples. Tuning numerical filters to achieve a desired level of fidelity requires substantial effort, one that must often be repeated upon any change to model configuration. The automation of this effort through a well-chosen numerical method allows model developers to focus on the physics.

We apply the FDDG method to the compressible Euler equations with gravity in both rotating and non-rotating reference frames, which we take as a model for dry atmospheres. The prognostic variables are density, momentum, and, as the prognostic thermodynamic variable, total energy. The equations then are

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \tag{1}$$

$$\partial_t (\rho \boldsymbol{u}) + \nabla \cdot (\boldsymbol{u} \otimes \rho \boldsymbol{u} + p\mathbb{I}) = -\rho \nabla \Phi + \mathcal{S}_{\rho \boldsymbol{u}} (\rho, \rho \boldsymbol{u}, \rho e), \tag{2}$$

$$\partial_t (\rho e) + \nabla \cdot (\boldsymbol{u} (p + \rho e)) = \mathcal{S}_{\rho e} (\rho, \rho \boldsymbol{u}, \rho e), \tag{3}$$

where Φ is the geopotential, $\mathcal{S}_{\rho \boldsymbol{u}}$ are momentum sources (e.g., the Coriolis force), and $\mathcal{S}_{\rho e}$ constitutes sources of energy (e.g., radiation). Total energy is defined as the sum of kinetic, potential, and internal energy,

$$\rho e = \frac{1}{2} \rho \|\boldsymbol{u}\|^2 + \rho \Phi + c_v \rho T, \tag{4}$$

where c_v is the specific heat capacity of dry air at constant volume. We diagnose temperature from the prognostic variables and pressure using the ideal gas law, i.e.,

$$T = \frac{1}{c_v \rho} \left(\rho e - \frac{1}{2} \rho \|\boldsymbol{u}\|^2 - \rho \Phi \right) \text{ and } p = \rho R T. \tag{5}$$

This set of equations includes processes often filtered out in atmospheric general circulation models (AGCMs), such as sound waves. Retaining additional physics is key if the model is used for coarse resolution AGCM simulations, cloud-resolving high-resolution AGCM simulations, and high-resolution large-eddy S=simulations (LES) of boundary layers. The flexibility is especially crucial for simulating other planetary bodies or analogous "small-planet" versions of Earth.

In what follows, we highlight the FDDG choices that result in accurate and stable simulations using the same technique in three numerical experiments of the compressible Euler equations. First, we examine an LES of a dry convective boundary layer in a box with rigid walls at the top and bottom and doubly periodic horizontal boundary conditions. Second, we explore the use of FDDG for an idealized dry GCM configuration (Held & Suarez, 1994). Third, we perform a simulation of an atmosphere in a "small-planet" configuration where the scale separation between convective scales and large scales is reduced (Wedi & Smolarkiewicz, 2009).

2 Numerical Experiments

To solve the compressible Euler equations in three-dimensional domains, we use the FDDG formulation of Chan (2018) and Waruszewski et al. (2022) and construct metric terms as outlined by D. A. Kopriva (2006). See the review by G. J. Gassner and Winters (2021) for a general overview of the FDDG method.

The choice of numerical flux is critical in guaranteeing the stability of the simulations. As mentioned, FDDG allows for a selection of numerical fluxes for the interior of the control volume and the cell interfaces. In addition, there is flexibility in the choice of numerical flux for any interface between elements, as well as the direction of the flux, i.e., a flux along the gravity-aligned direction need not be the same as a flux orthogonal to the direction of gravity.

In general, we choose kinetic energy preserving (KEP) volume fluxes to guarantee the flow's nonlinear stability; see G. J. Gassner et al. (2016) for an explanation of this property. This is especially important for simulating highly underresolved turbulent flows, as is typical in geophysical fluid dynamics. We find the KEP property to be the key feature that greatly increases the robustness of simulations. Stated succinctly, a numerical flux satisfies the KEP property if the discrete kinetic energy equation mimics the continuous kinetic energy equation. The importance of preserving the discrete algebraic properties of the kinetic energy equation has been commented on before (Zang, 1991).

Numerical fluxes that do not satisfy the KEP property can have terms in the discrete kinetic energy equation that correspond to energy injection due to transport, a manifestation of aliasing errors. It is serendipitous that there are a large class of numerical fluxes that satisfy this property, but it is especially worth noting that traditional DG methods do not have the KEP property when applied to geophysically relevant simulations, leading to stability problems in underresolved flows. In order to control this error, past methods had to use numerical filters, explicit dissipation, or overintegration strategies. None of these corrections are necessary if one just simply uses an FDDG formulation that automatically satisfies the KEP property.

The surface fluxes add a penalty term either as a Roe flux or a Rusanov flux (Roe, 1981; Hesthaven & Warburton, 2007). Furthermore, we enact a special treatment of the gravity source term. The details of our choices are given in the Appendix; however, we state the combination of methods used for a particular simulation in the relevant subsection.

Simulations are performed in a Julia-based open-source codebase that can exploit heterogeneous and distributed CPU/GPU architectures, (Bezanson et al., 2017; Besard et al., 2019; Besard et al., 2019). Although the DG method is well suited for parallel-computing architectures (Abdi et al., 2019; Sridhar et al., 2021), the scale of our problem allowed us to perform all simulations on a single Nvidia Titan V GPU. All plots in this text were generated using the Julia package Makie.jl (Danisch & Krumbiegel, 2021).

In the following subsections, we illustrate the skill of FDDG methods to simulate

1. Convection in a dry boundary layer.
2. The dry atmospheric circulation in an Earth-like domain.
3. The dry atmospheric circulation on a small Earth.

The domain for the first simulation is a horizontally periodic Cartesian box, for the second simulation an Earth-like thin spherical shell, and for the third simulation a spherical shell with a 20 times decreased planetary radius and increased rotation rate.

Nonetheless, the same computational kernels are used for all simulations. The connectivity between the elements and metric terms is the only change to transform from one domain to another. In addition, for computational efficiency, we implicitly timestep vertical acoustic and gravity wave modes in the Earth-like domain.

2.1 Dry Convection in the Atmospheric Boundary Layer

We start by simulating a dry atmospheric boundary layer. The following simulation is similar in spirit to Margolin et al. (1999), but with additional simplifications. All parameters for the simulation and their physical meaning are summarized in Table 1. We use a cubic domain of volume L^3 with periodic boundary conditions in the horizontal direction and no-flux, no-penetration boundary conditions in the vertical direction. The geopotential is $\Phi = gz$ where z is the vertical coordinate and x, y are the horizontal coordinates.

We start with a linear potential temperature profile

$$\theta(z) = T_s + \Delta\theta \frac{z}{L}, \tag{6}$$

parameter	value	unit	description
L	3	km	domain length
g	9.81	m^2 s^{-1}	gravitational constant
R	287	m^2 s^{-2} K^{-1}	gas constant for dry air
p_0	10^5	kg m^{-1} s^{-2}	reference sea-level pressure
T_s	300	K	surface temperature
c_v	717.5	J kg^{-1} K^{-1}	specific heat capacity of dry air at constant volume
c_p	1004.5	J kg^{-1} K^{-1}	specific heat capacity of dry air at constant pressure
ℓ	100	m	radiative length scale
$\mathcal{Q}$	100	m^{3}s^{-3}	radiative forcing magnitude
$\Delta\theta$	10	K	potential temperature difference from top to bottom

Table 1. Parameter values for the convective boundary layer test case.

which, when combined with the ideal gas law and hydrostatic balance,

$$\rho R T = p, \quad \theta = T\left(\frac{p_0}{p}\right)^{R/c_p}, \quad \text{and} \quad \partial_z p = -\rho g, \tag{7}$$

implies that pressure is

$$p(x,y,z,t=0) = p_0\left(-\frac{gL}{\Delta\theta c_p}\log(\theta(z)/T_s) + 1\right)^{c_p/R}. \tag{8}$$

We also apply an initial random perturbation to the velocity field to induce a rapid transition to turbulence,

$$\boldsymbol{u}(x,y,z,t=0) = 0.01\mathcal{N}(0,1), \tag{9}$$

where $\mathcal{N}$ is a random normal variable at each grid point. Thus the initial condition for total energy is

$$\rho e(x,y,z,t=0) = \frac{1}{2}\rho\|\boldsymbol{u}\|^2 + \rho\Phi + c_v\rho T \tag{10}$$

where c_v is the specific heat capacity of dry air at constant volume.

We apply a radiative forcing to drive convective instability. The resulting equations are

$$\partial_t\rho + \nabla\cdot(\rho\boldsymbol{u}) = 0 \tag{11}$$

$$\partial_t\rho\boldsymbol{u} + \nabla\cdot(\boldsymbol{u}\otimes\rho\boldsymbol{u} + p\mathbb{I}) = -\rho g\hat{z} \tag{12}$$

$$\partial_t\rho e + \nabla\cdot(\boldsymbol{u}\,[\rho e + p]) = \rho\frac{\mathcal{Q}}{\ell}\exp(-z/\ell). \tag{13}$$

We use a Kennedy-Gruber flux for the volume terms and a Kennedy-Gruber flux with a Roe flux penalty term for the interface numerical fluxes (Kennedy & Gruber, 2008); see Appendix A for details. For time-stepping, the fourth-order low storage 14-stage Runge-Kutta method of Niegemann et al. (2012) is employed. The sound waves are resolved in the simulation. We emphasize that we have not included any viscosity or diffusivity and solely rely on the numerical dissipation of the FDDG method for stability.

The domain is partitioned into 24^3 elements, each of which has three-dimensional fourth-order polynomials, leading to a total of 120^3 degrees of freedom. The smallest grid spacing is 21 meters, leading to a timestep size of $\Delta t = 0.11$ seconds to ensure compliance with the acoustic CFL limit.

The radiative heating is strongest near the surface, leading to air parcels to become buoyant and rise. As the plumes rise, they laterally entrain air from the surrounding environment; we expect the fluid to develop a well-mixed region of potential temperature near the surface. As the plumes move through the well-mixed layer, they eventually reach a stably-stratified region and overshoot their level of neutral buoyancy. The plumes drum on the stratified layer above, developing a layer of downward potential temperature fluxes and high potential temperature variance. This process erodes the stratification, leading to diffusive growth of the well-mixed region over time.

We estimate the growth of the well-mixed region from classic energetic arguments as done by, for example, Stull (1988). First, we observe that the flux of potential temperature is approximately $\mathcal{Q}_\theta = \mathcal{Q}/c_p \approx 0.1$ [K m s^{-1}]. We define the boundary layer height to be the height of maximum stratification. The boundary layer height at a given moment in time, t, is given by the empirical scaling law

$$h \propto \sqrt{t\frac{\mathcal{Q}_\theta L}{\Delta\theta}} \tag{14}$$

where the entrainment layer modifies the constant of proportionality. Without accounting for the entrainment layer, one derives $h(t) = \sqrt{2t\frac{\mathcal{Q}_\theta L}{\Delta\theta}}$ as in Stull (1988). Accounting for the entrainment layer seems to only modify the constant “2”, e.g. (Van Roekel et al., 2018), as opposed to modifying the scaling law.

Specifically, we compare the boundary layer height given by $h(t) = \sqrt{Ct\frac{\mathcal{Q}_\theta L}{\Delta\theta}}$, with $C = 3$ as in Souza et al. (2020), to that of the simulation in Figure 2. We see that the simulation agrees well with the empirical scaling law. This agreement suggests that the implicit dissipation mechanisms of the FDDG method enable subgrid-scale modeling, similar to other methods such as a Smagorinsky closure or a non-oscillatory scheme (Margolin et al., 1999; Van Roekel et al., 2018).

An instantaneous snapshot of the simulation after 5 hours is typified by Figure 3. The three-dimensional figure shows the mixed layer potential temperature as transparent, thereby emphasizing potential temperature anomalies. The visualization reveals the three-dimensional convective structure and small scorching plumes emanating from the surface. The top of the domain is is chosen to be the height at which the horizontally averaged potential temperature flux is most negative.

To the right of the three-dimensional figure are horizontal averages of potential temperature (top), vertical potential temperature flux (middle), and potential temperature variance (bottom). The horizontal average of potential temperature displays a well-mixed layer in the bottom kilometer of the domain, capped by an entrainment layer of enhanced stratification before easing into the background stratification. The vertical advective flux exhibits the expected linear structure in the mixed layer and is negative in the entrainment region. The negative flux arises from an anti-correlation between the vertical velocity and potential temperature, associated with plumes overshooting their region of neutral buoyancy. On average, this entrainment produces a negative flux whose maximum is approximately 17% of the input heat flux Q_θ. The negative flux minima is consistent with those commonly found in the literature, for example Margolin et al. (1999); Siebesma et al. (2007); Van Roekel et al. (2018), where the most negative flux is between 10%-20% of the heat input. The oscillations above the entrainment layer are due to gravity waves reflecting from the top of the domain. Furthermore, the plot shows that the temperature variance is largest in the entrainment layer.

2.2 Atmospheric Dynamical Core: The Held-Suarez Test

We next consider the GCM benchmark test proposed by Held and Suarez (1994), HS94 hereafter. The formulation of the problem allows for flexibility in hydrostatic vs

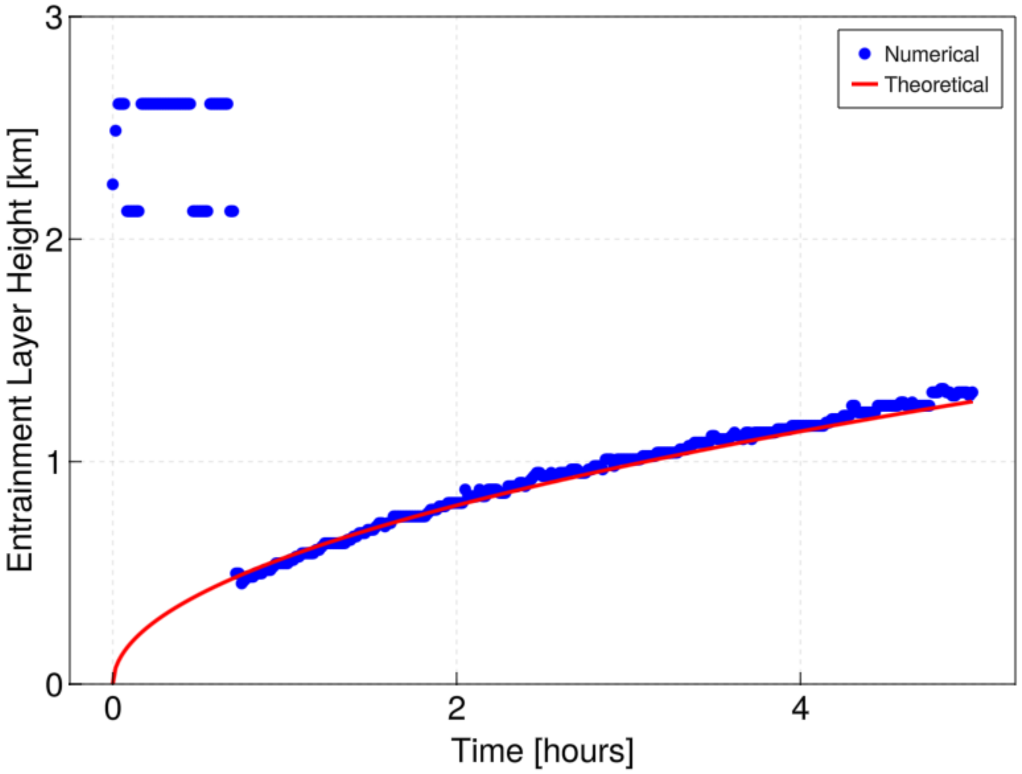

Figure 2. Convective Boundary Layer. The boundary layer height growth over time. Here we compare the empirical scaling law in red, given by $h(t) = \sqrt{3t\frac{Q_\theta L}{\Delta\theta}}$, to one calculated from the maximum potential temperature gradient in blue. A spin-up characterizes the first hour of simulation into the turbulent state. After the initial spin-up the simulation latches on to the empirical scaling law.

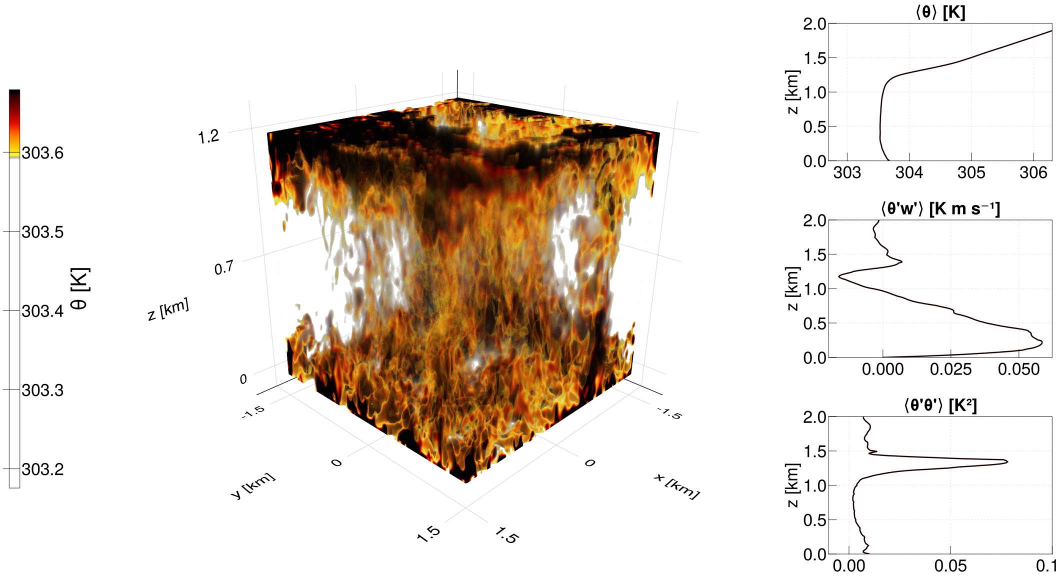

Figure 3. Convective Boundary Layer. A snapshot of potential temperature and its horizontally averaged statistics. The three-dimensional plot is a volume-rendering of the potential temperature where the white values on the color bar to the left are transparent in the volume-rendering visualization in the middle. The statistics on the right of the plot are horizontal averages of potential temperature (top), vertical potential temperature eddy flux (middle), potential temperature variance (bottom) at the same moment in time.

non-hydrostatic dynamics, dissipation mechanisms, prognostic variables, and boundary conditions. We choose to use an equation set that retains fully compressible dynamics and is formulated in terms of density, total energy, and Cartesian momentum as the prognostic variables, yielding the equations

$$\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0 \tag{15}$$

$$\partial_t (\rho \boldsymbol{u}) + \nabla \cdot (\boldsymbol{u} \otimes \rho \boldsymbol{u} + p \mathbb{I}) = -\rho \nabla \Phi - 2 \boldsymbol{\Omega} \times \rho \boldsymbol{u} + \mathbf{s}_{\rho \boldsymbol{u}} (\rho, \rho \boldsymbol{u}, \rho e) \tag{16}$$

$$\partial_t (\rho e) + \nabla \cdot (\boldsymbol{u} (p + \rho e)) = s_{\rho e} (\rho, \rho \boldsymbol{u}, \rho e) \tag{17}$$

where $\Phi = 2 G M_P r_{\text{planet}}^{-1} - G M_P r^{-1}$ is the geopotential, $\boldsymbol{\Omega} = \Omega \hat{\mathbf{z}}$ is the planetary angular velocity, and $\hat{\mathbf{z}}$ is the direction of the planetary axis of rotation. We do not make the traditional approximation, which assumes a thin atmospheric shell in which the distance from any point in the atmosphere to the center of the planet is taken to be equal to the planetary radius, leading to the Coriolis force having only horizontal components.

The HS94 forcing is applied to momentum and energy as follows

$$\mathbf{s}_{\rho \boldsymbol{u}} = -k_v \left(\mathbb{I} - \hat{r} \otimes \hat{r} \right) \rho \boldsymbol{u} \tag{18}$$

$$s_{\rho e} = -k_T \rho c_v \left(T - T_{\text{equilibrium}} \right), \tag{19}$$

where $T_{\text{equilibrium}}$ is the radiative equilibrium temperature depending on latitude (φ) and pressure $\sigma = p / p_0$,

$$T_{\text{equilibrium}} (\varphi, \sigma) = \max \left(T_{\min}, \left[T_{\text{equator}} - \Delta T_y \sin^2(\varphi) - \Delta \theta_z \ln(\sigma) \cos(\varphi) \right] \sigma^{R_d / c_p} \right), \tag{20}$$

and the parameters k_v, k_T are the inverse timescales for momentum damping and temperature relaxation, respectively, with

$$k_v = k_f \Delta \sigma \quad \text{and} \quad k_T = k_a + (k_s - k_a) \Delta \sigma \cos^4(\varphi), \tag{21}$$

parameter	value	unit	description
$\mathcal{X}$	1 or 20	-	scaling parameter
z_{top}	3×10^4	m	atmosphere height
r_{planet}	$6.371 \times 10^6/\mathcal{X}$	m	planetary radius
R	287	m^2 s^{-2} K^{-1}	gas constant for dry air
Ω	$2\pi/86400 \times \mathcal{X}$	s^{-1}	Coriolis magnitude
p_0	1×10^5	kg m^{-1} s^{-2}	reference sea-level pressure
T_{min}	200	K	minimum equilibrium temperature
$T_{equator}$	315	K	equatorial equilibrium temperature
σ_b	0.7	-	dimensionless damping height
c_v	717.5	J kg^{-1} K^{-1}	specific heat capacity of dry air at constant volume
c_p	1004.5	J kg^{-1} K^{-1}	specific heat capacity of dry air at constant pressure
k_f	$\mathcal{X}/86400$	s^{-1}	damping scale for momentum
k_a	$\mathcal{X}/(40 \times 86400)$	s^{-1}	polar relaxation scale
k_s	$\mathcal{X}/(4 \times 86400)$	s^{-1}	equatorial relaxation scale
ΔT_y	60	K	latitudinal temperature difference
$\Delta\theta_z$	10	K	vertical temperature difference
G	6.67408×10^{-11}	kg^{-1} m^3 s^{-2}	gravitational constant
M_P	$5.9722/\mathcal{X}^2 \times 10^{24}$	kg	planetary mass

Table 2. Parameter values for the Held-Suarez test case. The value $\mathcal{X} = 1$ corresponds to the standard test case, and $\mathcal{X} = 20$ is a small planet version of the Held-Suarez test case.

with $\Delta\sigma = \max\{0, (\sigma - \sigma_b)/(1 - \sigma_b)\}$. The temperature and pressure are diagnosed from total energy and the ideal gas law,

$$T = \frac{1}{c_v \rho}\left(\rho e - \rho\|\boldsymbol{u}\|^2 - \rho\Phi\right) \quad \text{and} \quad p = \rho R T. \tag{22}$$

The forcing terms differ only in quantitatively irrelevant aspects from the original formulation in HS94. In particular, we choose a constant pressure p_0 in the definition of σ instead of the instantaneous surface pressure. The parameter values are summarized in Table 2.

The domain is a piecewise polynomial approximation to a thin spherical shell of radius r_{planet} and height z_{top}. The thin spherical domain is partitioned into curved elements and uses an isoparametric representation of the domain and the cubed sphere mapping by Ronchi et al. (1996). In essence, this choice represents the domain as a piecewise polynomial function where the order of the polynomial corresponds to the order of the discretization (Winters et al., 2021). The metric terms are treated as in D. A. Kopriva (2006) and satisfy the discrete property that the divergence of a constant vector field is zero, i.e., the metric terms are free-stream preserving. The use of an isoparametric representation of the sphere with free-stream preserving metrics has a few subtleties. Since the vertical and horizontal directions are no longer discretely orthogonal, one must distinguish covariant and contravariant vertical directions.

We use no-flux boundary conditions for density and total energy. We use free-slip boundary conditions for the horizontal momenta and no-penetration boundary conditions for the vertical momentum. Our initial condition is a fluid that starts from rest in an isothermal atmosphere. We take the global temperature to be $T_I = 285K$, leading to

$$p(r) = p_0 \exp\left(-\frac{\Phi(r) - \Phi(r_{\text{planet}})}{RT_I}\right) \text{ and } \rho(r) = \frac{1}{RT_I} p(r). \tag{23}$$

We use implicit time-stepping in order to numerically filter vertically propagating sound waves and gravity waves. Specifically, we use the second-order Runge-Kutta IMEX scheme of F. X. Giraldo et al. (2013), but modify Equation 3.9 of their work by choosing $a_{32} = 1/2$ for an enhanced stability region. We use the Jacobian of both the surface and volume flux in the vertical for the implicit time-stepping component; see Appendix A2 for details. We linearize about the previous timestep, update the Jacobian for every column every 20 minutes of simulated time, and factorize it using a banded LU decomposition (Golub & Loan, 2013). Horizontal acoustic modes then limit the timestep. The largest Mach number, the ratio of the advective speed and the soundspeed, for the flow is roughly 0.25 in this simulation.

Aside from the inherent numerical dissipation resulting from the interface flux terms and implicit time-stepping, we use no additional forms of damping such as those in Jablonowski and Williamson (2011). In particular, we do not use any form of viscosity/hyperviscosity for small-scale damping. Furthermore, we do not include any divergence damping or filters. The method remains conservative up to rounding errors from finite-precision arithmetic. For the Held-Suarez benchmark, only density is conserved since it has no sources.

We run the HS94 test case with 6×10^2 elements in the horizontal on an equiangular cubed sphere, 10 evenly spaced elements in the vertical, polynomial order 4 within each element, totaling at $6{\times}50^2$ degrees of freedom in the horizontal and 40 degrees of freedom in the vertical. The minimum grid spacing is 120 kilometers in the horizontal and 650 meters in the vertical. We choose a timestep of 55 s to keep within the horizontal acoustic CFL limit. We discard the first 200 days of the simulation as spinup and average over the last 1000 days, as in HS94. We gather statistics by interpolating the cubed sphere grid to spherical coordinates and converting the Cartesian momentum to spherical velocities. As usual, we denote the zonal velocity component by u, the meridional velocity by v, and the vertical velocity by w. We gather statistics in height coordinates and for plotting we use the zonal and temporal average of pressure at the equator as the height. We comment that our choice of averaging does not use the Favre average, which is the more natural choice for comparison with pressure coordinate models.

In Figure 4 we show the long-time average of the zonal-mean zonal wind $\langle u \rangle$, temperature $\langle T \rangle$, temperature variance $\langle T'T' \rangle$, eddy momentum flux $\langle u'v' \rangle$, horizontal eddy kinetic energy $0.5\langle u'u' + v'v' \rangle$. The choice of fields is to directly compare with Figure 1 of Wan et al. (2008). The results here are in agreement with those reported in the literature (Held & Suarez, 1994; Chen et al., 1997; Ringler et al., 2000; Ullrich & Jablonowski, 2012). For example, the peak in westerly winds, temperature variance, and eddy kinetic energy are all within 10% of published results.

Perhaps the largest difference is in the meridional heat transport. In our simulations, the $\langle v'T' \rangle = -9$ K m s^{-1} contour remains disconnected above and below the “stretched height” = 400 hPa line. This difference could be due to the use of height coordinates for averaging rather than pressure coordinates, since a zonal average over a surface of constant height is different than that of constant pressure.

2.3 Small-Planet Held-Suarez

In addition to the typical HS94 configuration, we simulate a small planet with a large-scale climatology similar to that of HS94 by rescaling the equations in a manner similar to a DARE/hypohydrostatic rescaling of the equations as done by Kuang et al. (2005) and Pauluis et al. (2006), respectively. This rescaling is an exact similarity transformation of the hydrostatic primitive equations using the traditional approximation and thus only affects the balance between the non-hydrostatic and hydrostatic components of the flow. The test is the similar to the one as proposed by Wedi and Smolarkiewicz (2009) with minor modifications.

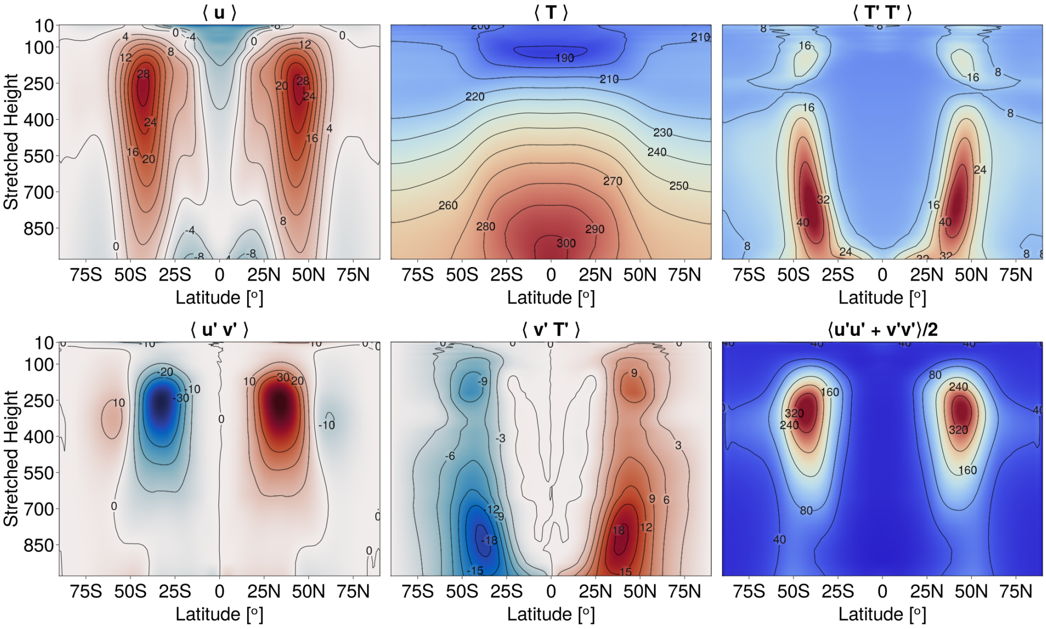

Figure 4. Temporal and zonal average Held-Suarez statistics. The "stretched height" is a global rescaling of height with the long time and zonal average of pressure at the equator, mimicking the effect of using pressure coordinates for ease of comparison with figures in the literature. The long-time average uses the last 1000 days of the simulation. We use 8 evenly spaced elements in the vertical and 6×10^2 elements in the horizontal with a polynomial order four basis in each direction.

We decrease the planetary radius by a factor of $\mathcal{X} = 20$ compared to Earth, increase the rotation rate by a factor of $\mathcal{X}$, and decrease the mass of the planet by a factor of $\mathcal{X}^2$. Furthermore, we increase all relaxation timescales in the problem by a factor of $\mathcal{X}$. The atmospheric height and temperature equilibrium remain the same. The parameter values are tabulated in Table 1. We will justify these choices shortly.

Changing the planetary radius, increasing the rotation rate, and keeping the same temperature equilibrium results in a planetary model with a similar thermal wind. This a natural consequence of the rescaling being an exact similarity transformation for the hydrostatic primitive equations. Indeed the thermal wind, u_{thermal} scales like

$$u_{\text{thermal}} \sim \frac{\Delta T}{\Omega \Delta H} \tag{24}$$

where $\Delta T / \Delta H$ is the latitudinal gradient of temperature. Observe that $\Delta H \propto r_{\text{planet}}$ and recall that the equilibrium temperature distribution is unchanged from the original configuration. Thus both ΔT and $\Omega \Delta H$ remain constant, and the resulting thermal wind is approximately the same across the two simulations. Consequently, the Rossby number $Ro \equiv u_{\text{thermal}}/(2\Omega r_{\text{planet}})$ remains the same.

Changing the planetary mass is necessary to retain an Earth-like hydrostatically balanced state. The gradient of the geopotential scales like $\nabla \Phi \sim r_{\text{planet}}^{-2}$ and thus the planetary mass must scale by a factor of $\mathcal{X}^{-2}$ to maintain the same force. We could have achieved a similar result by simply taking the geopotential to be $\Phi = gr$, but we saw no need to use this linearization.

We keep the same number of grid points, $6 \times 50^2 \times 40$ degrees of freedom, leading to a minimum grid spacing of 6 kilometers in the horizontal and 650 meters in the vertical. For the small planet, we use explicit time-stepping—the same low storage Runge-Kutta method of Niegemann et al. (2012)—which affords timesteps of size $dt = 6.5$ s, which corresponds to an acoustic Courant number of 3.6 in the vertical and 0.38 in the horizontal. Small timesteps are less of a limitation because planetary-scale dynamics are $\mathcal{X} = 20$ times faster than Earth's. Thus we only need to simulate 60 Earth days, which corresponds to 1200 small-planet days. The initial condition uses the same formula as before, Equation 23. We discard the first 20% of the simulation and average over the rest.

Figure 5 shows that statistics are relatively unchanged with respect to those in Figure 4, except for the zonal velocity, which has a vigorous easterly flow along the equator. We attribute the change in the zonal mean climatology of the zonal velocity to the increased vertical velocity, which in turn affects the non-traditional terms in the Coriolis force; these terms are not negligible in the small planet. See Marshall et al. (1997) for an explanation of the underlying physics in the ocean context. An enhanced easterly flow in the small planet configuration has been observed before. For example, see Figure 18 of Wedi and Smolarkiewicz (2009).

We confirm this statement by neglecting the non-traditional components of the planetary angular velocity,

$$\boldsymbol{\Omega}_{\text{traditional}} = (\hat{r} \cdot \boldsymbol{\Omega})\hat{r} \tag{25}$$

and comparing the zonal mean velocity statistics of the three simulations in Figure 6. We do not modify the metric terms thus the approximation is inconsistent, nonetheless it serves to illustrate the point. We see that the zonal mean velocity statistic of the original HS94 setup corresponds to that of the small planet with the "traditional" planetary angular velocity but not that of the small planet with the full-planetary angular velocity. This effect is a consequence of the decreased aspect ratio of the vertical vs horizontal domain in the small planet, which in turn increases the magnitude of the vertical velocity by a factor $\mathcal{X}$. Stated differently, even though the full Coriolis force is present in the Earth-like domain, the vertical velocity component is too weak to make a substantial difference, as expected for this test-case.

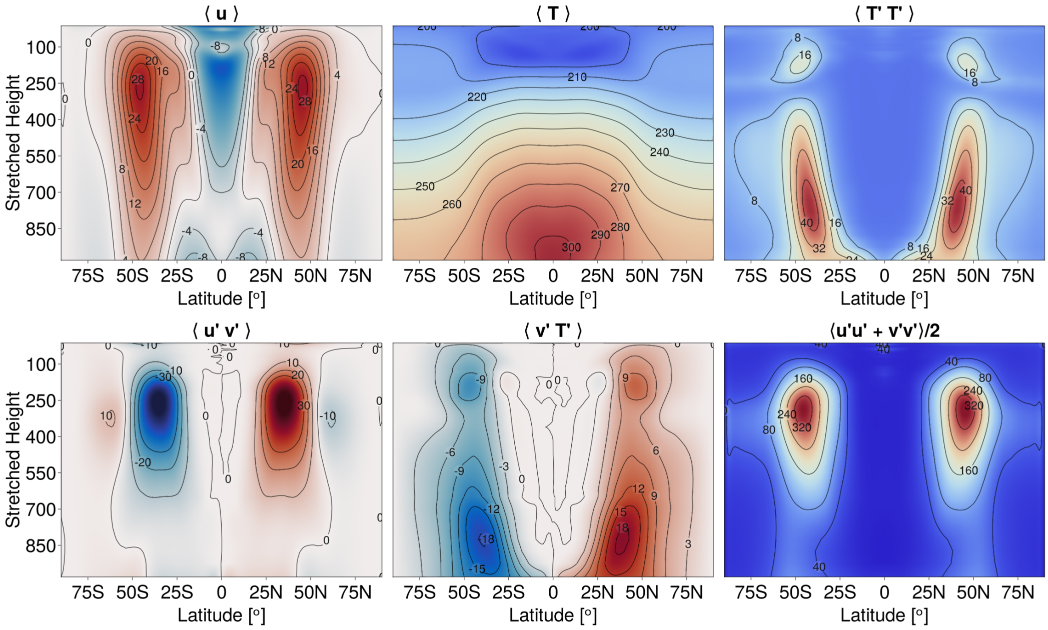

Figure 5. Small Planet Held-Suarez. The long time and zonal average Held-Suarez statistics in a “small planet” configuration. The “stretched height” is a rescaling of height with the long time and zonal average of pressure at the equator, mimicking the effect of using pressure coordinates for ease of comparison with figures in the literature. Time averages are taken over the last 1000 days of the simulation. We use 8 evenly spaced elements in the vertical and 6×10^2 elements in the horizontal with a polynomial order four basis in each direction.

We reemphasize no further code tuning is required to retain stability. Upon modification of the domain and appropriate parameters, the only necessary change was a reduction of timestep to stay within the acoustic CFL of the small planet. The ability to easily change planetary parameters allows for a systematic investigation of scaling laws of planetary systems with respect to rotation rates, planetary radii, and atmospheric heights.

3 Conclusions

We have presented the application of a discontinuous Galerkin method to an idealized dry atmosphere for local large-eddy simulations and global circulation modeling. We have shown that the statistics generated from using a fully-compressible code, with density, total energy, and Cartesian momentum as prognostic variables, are similar to other models in local and global settings. Furthermore, we did not require stabilization mechanisms outside those naturally afforded by the discontinuous Galerkin numerical method and time-stepping.

The main limitations of the numerical method are not associated with the spatial discretization per se but rather the need to develop efficient time-stepping strategies for modern computer architectures that can overcome limitations induced by acoustic waves, especially in the presence of topography. Different architectures may necessitate different algorithms to achieve an optimal time-to-solution. There are many approaches for obtaining a better time-to-solution that are worth exploring, e.g., fully implicit time stepping (Nguyen et al., n.d.) and multi-rate methods (Knoth & Wensch, 2014). Furthermore, switching between different flux-differencing methods in the vertical vs. horizon-

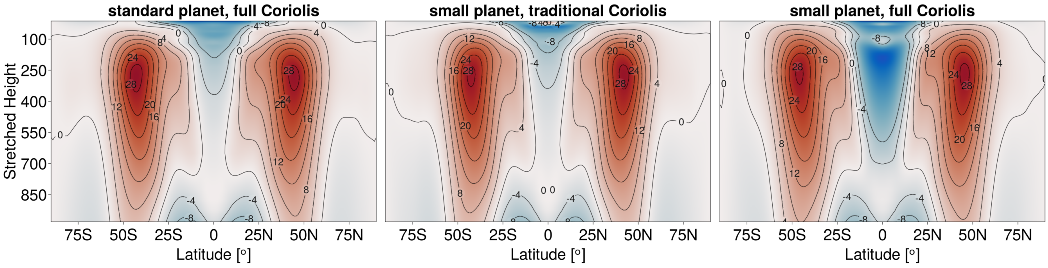

Figure 6. Small Planet Held-Suarez. A comparison between the long time and zonal average of the zonal velocity between three different configurations. The left-most plot is the typical HS94 setup utilizing the full Coriolis force, the middle plot is the small planet HS94 setup with the traditional approximation to the Coriolis force, and is the small planet with the full Coriolis force. We see that keeping the traditional approximation replicates the zonal velocity statistics of the Earth-like planet at the expense of being unphysical with respect to the small planet.

tal may yield larger timesteps due to better linearization properties, (G. Gassner et al., 2020; Ranocha & Gassner, 2021). An alternative option is to use lower order methods, such as staggered grid finite volume or lower polynomial orders, for the implicit vertical discretization, which may yield a faster time-to-solution.

The present study is limited to an idealized dry atmosphere, and moisture, topography, and radiation are necessary for realistic simulations. Positivity-preserving methods such as those outlined in Light and Durran (2016) need to be used, and topographic effects can also be handled (Baldauf, 2021). Previous studies of discontinuous Galerkin methods have involved designing numerical fluxes that preserve desired discrete properties. It would also be interesting to compare candidate methods for geophysical flows.

It is possible to bridge the gap between existing parameterizations and novel numerics by leveraging the sub-cell finite-volume interpretation of the Flux-Differencing Discontinuous Galerkin method. This interpretation is similar to using a "physics grid" as in Herrington et al. (2019) but simpler in its implementation. Another option is to develop new parameterizations that leverage the subgrid-scale shape functions of the spectral element method, akin to using a higher-order moment closure.

Flux-Differencing Discontinuous Galerkin methods are an interesting alternative discretization for Earth system modeling. They enable large-eddy simulation modeling with its natural subgrid-scale dissipation mechanisms, allow for flexible representation of the domain, and can be implemented efficiently on parallel hardware architectures. Developing efficient implicit timesteping methods in order to overcome the limitations due to gravity and sound waves are a remaining challenge, but we hope that the extra robustness and higher-order accuracy provided by FDDG methods will eventually allow for an overall faster time to solution.

Acknowledgments

We thank Keaton Burns and Greg L. Wagner for their encouragement and advice throughout the writing the of this manuscript. Our work is supported by the generosity of Eric and Wendy Schmidt by recommendation of the Schmidt Futures program, and by the National Science Foundation under AGS grants 1835860, 1835576, and 1835881.

Appendix A Discontinuous Galerkin Details

In this appendix, we collect choices of numerical fluxes and linear models. To highlight our choices, we use the compressible Euler equations with gravity,

$$\partial_t \rho + \nabla \cdot \rho \boldsymbol{u} = 0 \tag{A1}$$

$$\partial_t \rho \boldsymbol{u} + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u} + p\mathbb{I}) = -\rho \nabla \Phi \tag{A2}$$

$$\partial_t \rho e + \nabla \cdot (\boldsymbol{u}\, [\rho e + p]) = 0 \tag{A3}$$

$$(\gamma - 1)\left(\rho e - \frac{1}{2}\rho \|\boldsymbol{u}\|^2 - \rho \Phi\right) = p \tag{A4}$$

where $\gamma = 7/5$ and Φ is the geopotential. The source terms that do not involve gradients are collocated with grid-points and require no further description.

To describe the numerical fluxes we use the same notation as G. J. Gassner et al. (2016). Thus for a scalar field ψ with + as the "exterior" value and − as the "interior" value (Hesthaven & Warburton, 2007; G. J. Gassner et al., 2016), we take the averaging operator $\{\cdot\}$ and jump operator $[\![\cdot]\!]$ to mean

$$\{\psi\} \equiv \frac{\psi^+ + \psi^-}{2} \text{ and } [\![\psi]\!] \equiv \frac{\psi^+ - \psi^-}{2}. \tag{A5}$$

The averaging and jump operators are applied componentwise for vector and tensor Cartesian fields. We point out that our definition of jump, $[\![\cdot]\!]$, has a factor of two that is different from most other conventions.

The flux-differencing and metric term implementations are done in skew-symmetric form as outlined by Chan (2018) and Waruszewski et al. (2022). The metric terms are constructed to be free-stream preserving, (D. A. Kopriva, 2006; D. Kopriva, 2009).

A1 Numerical Fluxes

For the volume terms we use the Kennedy-Gruber flux (Kennedy & Gruber, 2008), with a modification to the gravity source term,

$$\mathcal{F}_\rho = \{\rho\}\{\boldsymbol{u}\} \tag{A6}$$

$$\mathcal{F}_{\rho \boldsymbol{u}} = \{p\}\,\mathbb{I} + \{\rho\}\{\boldsymbol{u}\} \otimes \{\boldsymbol{u}\} + \{\rho\}\,[\![\Phi]\!]\mathbb{I} \tag{A7}$$

$$\mathcal{F}_{\rho e} = \{\boldsymbol{u}\}\left(\{\rho\}\{e\} + \{p\}\right), \tag{A8}$$

where $\mathbb{I}$ is the identity matrix. The modification to the gravity source term was motivated by combining the entropy stable scheme of Waruszewski et al. (2022) with the Kennedy-Gruber flux.

We decompose the numerical flux normal to an interface between elements into two components by using the flux above as the "central" component[2] and a penalty term, which adds dissipation in a manner similar to upwinding. We choose different penalty terms for the vertical vs. horizontal directions when evolving the compressible Euler-Equations on the sphere. Distinguishing between vertical and horizontal fluxes is natural given the anisotropy of the Earth-like computational domain: a spherical shell with radius $\mathcal{O}(10^4)$ kilometers and height $\mathcal{O}(10)$ kilometers. This domain typically leads to pancake-like grid elements whose breadth is roughly 100 times its height.

In the direction associated with vertical grid points we use a Rusanov penalty whose wavespeed is based on a reference pressure and reference density. The reference density and pressure are updated every 20 minutes of simulated time with the instantaneous val-

---

[2] The geopotential is continuous on an interface; thus $[\![\Phi]\!] = 0$

ues. Specifically we add the following numerical fluxes,

$$c = \sqrt{\gamma p^{\mathrm{ref}}/\rho^{\mathrm{ref}}},\ \mathcal{F}^{\rho} = \{c\}_{\infty} [\![\rho]\!],\ \mathcal{F}^{\rho \boldsymbol{u}} = \{c\}_{\infty} [\![\rho \boldsymbol{u}]\!], \text{ and } \mathcal{F}^{\rho e} = \{c\}_{\infty} [\![\rho e]\!] \tag{A9}$$

where $\{c\}_{\infty} = \max\{c^{+}, c^{-}\}$. In the directions orthogonal to the vertical direction we use Roe fluxes,

$$\begin{aligned}
c &= \sqrt{\gamma p/\rho} && \text{(A10)}\\
w_1 &= |\{u_n\}_\rho - \{c\}_\rho| \left([\![p]\!] - \{\rho\}_\rho\{c\}_\rho[\![u_n]\!]\right) / \left(2\{c\}_\rho^2\right) && \text{(A11)}\\
w_2 &= |\{u_n\}_\rho + \{c\}_\rho| \left([\![p]\!] - \{\rho\}_\rho\{c\}_\rho[\![u_n]\!]\right) / \left(2\{c\}_\rho^2\right) && \text{(A12)}\\
w_3 &= |\{u_n\}_\rho| \left([\![\rho]\!] - [\![p]\!]/\{c\}_\rho^2\right) && \text{(A13)}\\
w_4 &= |\{u_n\}_\rho|\{\rho\}_\rho && \text{(A14)}\\
\mathcal{F}^{\rho} &= w_1 + w_2 + w_3 && \text{(A15)}\\
\mathcal{F}^{\rho \boldsymbol{u}} &= w_1 \left(\{\boldsymbol{u}\}_\rho - \{c\}_\rho \hat{n}\right) + w_2 \left(\{\boldsymbol{u}\}_\rho + \{c\}_\rho \hat{n}\right) + w_3\{\boldsymbol{u}\}_\rho + w_4 \left([\![\boldsymbol{u}]\!] - [\![u_n]\!]\hat{n}\right) && \text{(A16)}\\
\mathcal{F}^{\rho e} &= w_1 \left(\{(\rho e + p)/\rho\}_\rho - \{c\}_\rho\{u_n\}_\rho\right) + w_2 \left(\{(\rho e + p)/\rho\}_\rho + \{c\}_\rho\{u_n\}_\rho\right) && \text{(A17)}\\
&\quad + w_3 \left(\{\boldsymbol{u}\}_\rho \cdot \{\boldsymbol{u}\}_\rho/2 + \Phi\right) + w_4 \left(\{\boldsymbol{u}\}_\rho \cdot [\![\boldsymbol{u}]\!] - \{u_n\}_\rho[\![u_n]\!]\right) && \text{(A18)}
\end{aligned}$$

where the averaging, $\{\cdot\}_\rho$ is

$$\{\psi\}_\rho = \frac{\sqrt{\rho^{+}}\psi^{+} + \sqrt{\rho^{-}}\psi^{-}}{\sqrt{\rho^{+}} + \sqrt{\rho^{-}}} \tag{A19}$$

for all fields ψ except for ρ in which case

$$\{\rho\}_\rho = \sqrt{\rho^{+}\rho^{-}}. \tag{A20}$$

The variable $\hat{n} = \hat{n}(x, y, z)$ is the normal vector to a point on an element face (unit vectors of the contravariant basis) and $u_n = \boldsymbol{u} \cdot \hat{n}$ is the velocity component normal to a face.

On the boundaries of the sphere we set the density and energy fluxes to zero and for momentum we use the exterior + state and interior − state as

$$\rho \boldsymbol{u}^{+} = \left(\mathbb{I} - 2\hat{n} \otimes \hat{n}\right) \rho \boldsymbol{u}^{-} \tag{A21}$$

where $\hat{n}$ is the wall-normal unit vector. We then use central fluxes to compute the flux on the boundary. Equation A21 amounts to using the reflection principle on the wall-normal velocity, while also implementing no-flux boundary conditions for the tangential velocities. See Hesthaven and Warburton (2007) for further clarification on the reflection principle.

A2 Jacobian for Implicit Timestepping

To calculate the Jacobian of the compressible Euler equations with gravity it suffices to focus on the numerical flux,

$$\begin{aligned}
\mathcal{F}_{\rho} &= \{\rho\}\{\boldsymbol{u}\} && \text{(A22)}\\
\mathcal{F}_{\rho \boldsymbol{u}} &= \{p\}\mathbb{I} + \{\rho\}\{\boldsymbol{u}\} \otimes \{\boldsymbol{u}\} + \{\rho\}[\![\Phi]\!]\mathbb{I} && \text{(A23)}\\
\mathcal{F}_{\rho e} &= \{\boldsymbol{u}\}\left(\{\rho\}\{e\} + \{p\}\right). && \text{(A24)}
\end{aligned}$$

First we make the observation that variables such as $\boldsymbol{u}$, e, and p are nonlinear functions diagnosed from the prognostic variables ρ, $\rho \boldsymbol{u}$, and ρe,

$$\boldsymbol{u} = \frac{\rho \boldsymbol{u}}{\rho}\ ,\ e = \frac{\rho e}{\rho},\ \text{and } p = (\gamma - 1)\left(\rho e - \frac{\rho \boldsymbol{u} \cdot \rho \boldsymbol{u}}{2\rho} - \rho \Phi\right). \tag{A25}$$

Thus the linearization of Equations A22-A24 will involve linearizations of $\boldsymbol{u}$, e, and p. Furthermore, we can make use of the identities $\{a + b\} = \{a\} + \{b\}$ since we are using simple averages for the numerical flux. For example, the linearization of the mass conservation flux with respect to reference states ρ_r and $(\rho u)_r$ is calculated by including infinitesimal perturbations ρ and $\rho \boldsymbol{u}$, e.g.

$$\begin{aligned}
\mathcal{F}^L_\rho &= \{\rho_r + \rho\}\left\{\frac{(\rho \boldsymbol{u})_r + \rho \boldsymbol{u}}{\rho_r + \rho}\right\} - \{\rho_r\}\left\{\frac{(\rho \boldsymbol{u})_r}{\rho_r}\right\} && \text{(A26)}\\
&= \{\rho_r\}\left\{\frac{(\rho \boldsymbol{u})_r + \rho \boldsymbol{u}}{\rho_r + \rho}\right\} - \{\rho_r\}\left\{\frac{(\rho \boldsymbol{u})_r}{\rho_r}\right\} + \{\rho\}\left\{\frac{(\rho \boldsymbol{u})_r + \rho \boldsymbol{u}}{\rho_r + \rho}\right\} && \text{(A27)}\\
&= \{\rho_r\}\left\{\frac{(\rho \boldsymbol{u})_r + \rho \boldsymbol{u}}{\rho_r + \rho} - \frac{(\rho \boldsymbol{u})_r}{\rho_r}\right\} + \{\rho\}\left\{\frac{(\rho \boldsymbol{u})_r}{\rho_r}\right\} && \text{(A28)}\\
&= \{\rho_r\}\left\{\frac{\rho \boldsymbol{u}}{\rho_r} - \rho\frac{\boldsymbol{u}_r}{\rho_r}\right\} + \{\rho\}\left\{\frac{(\rho \boldsymbol{u})_r}{\rho_r}\right\}, && \text{(A29)}
\end{aligned}$$

where in the last line we made use of

$$\frac{1}{\rho_r + \rho} = \frac{1}{\rho_r} - \frac{\rho}{\rho_r^2} \text{ and } \frac{(\rho \boldsymbol{u})_r + \rho \boldsymbol{u}}{\rho_r + \rho} = \frac{(\rho \boldsymbol{u})_r}{\rho_r} + \frac{\rho \boldsymbol{u}}{\rho_r} - \rho\frac{(\rho \boldsymbol{u})_r}{(\rho_r)^2}. \tag{A30}$$

We condense equation A29 by defining the reference velocity $\boldsymbol{u}_r$ and linearized velocity $\boldsymbol{u}_L$ as

$$\boldsymbol{u}_r \equiv \frac{(\rho \boldsymbol{u})_r}{\rho_r} \text{ and } \boldsymbol{u}_L \equiv \frac{\rho \boldsymbol{u}}{\rho_r} - \rho\frac{\boldsymbol{u}_r}{\rho_r}, \tag{A31}$$

so that

$$\mathcal{F}^L_\rho = \{\rho_r\}\{\boldsymbol{u}_L\} + \{\rho\}\{\boldsymbol{u}_r\}. \tag{A32}$$

Similarly we define linearized and reference values as

$$\begin{aligned}
e_r &\equiv \frac{(\rho e)_r}{\rho_r}, \; p_r \equiv (\gamma - 1)\left((\rho e)_r - \frac{(\rho \boldsymbol{u})_r \cdot (\rho \boldsymbol{u})_r}{2\rho_r} - \rho_r \Phi\right), && \text{(A33)}\\
e_L &\equiv \frac{\rho e}{\rho_r} - \rho\frac{e_r}{\rho_r}, \text{ and } p_L \equiv (\gamma - 1)\left(\rho e - \frac{1}{2}\left(\rho \boldsymbol{u}_r \otimes \boldsymbol{u}_L + \rho \boldsymbol{u} \otimes \boldsymbol{u}_r\right) - \rho\Phi\right). && \text{(A34)}
\end{aligned}$$

In total, the Jacobian of equations A22-A24 with respect to a reference state $\rho_r, (\rho \boldsymbol{u})_r, \rho e_r$, yields the linearized numerical fluxes

$$\begin{aligned}
\mathcal{F}^L_\rho &= \{\rho_r\}\{\boldsymbol{u}_L\} + \{\rho\}\{\boldsymbol{u}_r\} && \text{(A35)}\\
\mathcal{F}^L_{\rho \boldsymbol{u}} &= (\{p_L\} + \{\rho\}[\![\Phi]\!])\mathbb{I} + \{\rho\}\{\boldsymbol{u}_r\} \otimes \{\boldsymbol{u}_r\} + \{\rho_r\}\{\boldsymbol{u}_r\} \otimes \{\boldsymbol{u}_L\} + \{\rho_r\}\{\boldsymbol{u}_L\} \otimes \{\boldsymbol{u}_r\} && \text{(A36)}\\
\mathcal{F}^L_{\rho e} &= \{\boldsymbol{u}_L\}(\{\rho_r\}\{e_r\} + \{p_r\}) + \{\boldsymbol{u}_r\}(\{\rho\}\{e_r\} + \{\rho_r\}\{e_L\} + \{p_L\}). && \text{(A37)}
\end{aligned}$$

We see by inspection that the above system is indeed linear with respect to ρ, $\rho \boldsymbol{u}$, and ρe.

For the surface term component of the numerical flux, we use linearized versions of the surface flux used in the full equations plus a reference state based Rusanov flux for the penalty term. Each column has its own reference state and the resulting linear systems are factored and solved directly. The reference state itself is constructed from instantaneous values of density, horizontal-momentum, and total-energy. Projecting out the vertical momentum from the reference state makes the method slightly more robust.

References

Abdi, D. S., Wilcox, L. C., Warburton, T. C., & Giraldo, F. X. (2019). A gpu-accelerated continuous and discontinuous galerkin non-hydrostatic atmospheric model. *The International Journal of High Performance Computing*

Applications, *33*(1), 81-109. Retrieved from `https://doi.org/10.1177/1094342017694427` doi: 10.1177/1094342017694427

Baldauf, M. (2021). A horizontally explicit, vertically implicit (hevi) discontinuous galerkin scheme for the 2-dimensional euler and navier-stokes equations using terrain-following coordinates. *Journal of Computational Physics*, *446*, 110635. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999121005301` doi: https://doi.org/10.1016/j.jcp.2021.110635

Besard, T., Churavy, V., Edelman, A., & Sutter, B. D. (2019). Rapid software prototyping for heterogeneous and distributed platforms. *Advances in Engineering Software*, *132*, 29 - 46. Retrieved from `http://www.sciencedirect.com/science/article/pii/S0965997818310123` doi: https://doi.org/10.1016/j.advengsoft.2019.02.002

Besard, T., Foket, C., & De Sutter, B. (2019, April). Effective extensible programming: Unleashing julia on gpus. *IEEE Transactions on Parallel and Distributed Systems*, *30*(4), 827-841. doi: 10.1109/TPDS.2018.2872064

Bezanson, J., Edelman, A., Karpinski, S., & Shah, V. B. (2017). Julia: A Fresh Approach to Numerical Computing. *SIAM Review*, *59*(1), 65–98. doi: 10/f9wkpj

Boyd, J. P. (2001). *Chebyshev and Fourier spectral methods* (Second ed.). Mineola, NY: Dover Publications.

Cao, W., Shu, C.-W., Yang, Y., & Zhang, Z. (2015). Superconvergence of discontinuous galerkin methods for two-dimensional hyperbolic equations. *SIAM Journal on Numerical Analysis*, *53*(4), 1651-1671. Retrieved from `https://doi.org/10.1137/140996203` doi: 10.1137/140996203

Chan, J. (2018). On discretely entropy conservative and entropy stable discontinuous galerkin methods. *Journal of Computational Physics*, *362*, 346-374. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999118301153` doi: https://doi.org/10.1016/j.jcp.2018.02.033

Chen, M., Rood, R. B., & Takacs, L. L. (1997). Impact of a semi-lagrangian and an eulerian dynamical core on climate simulations. *Journal of Climate*, *10*(9), 2374 - 2389. Retrieved from `https://journals.ametsoc.org/view/journals/clim/10/9/1520-0442_1997_010_2374_ioasla_2.0.co_2.xml` doi: 10.1175/1520-0442(1997)010⟨2374:IOASLA⟩2.0.CO;2

Danisch, S., & Krumbiegel, J. (2021). Makie.jl: Flexible high-performance data visualization for julia. *Journal of Open Source Software*, *6*(65), 3349. Retrieved from `https://doi.org/10.21105/joss.03349` doi: 10.21105/joss.03349

Fisher, T. C., & Carpenter, M. H. (2013). High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains. *Journal of Computational Physics*, *252*, 518-557. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999113004385` doi: https://doi.org/10.1016/j.jcp.2013.06.014

Gassner, G., Svärd, M., & Hindenlang, F. J. (2020). Stability issues of entropy-stable and/or split-form high-order schemes. *ArXiv*, *abs/2007.09026*.

Gassner, G., Winters, A., & Kopriva, D. (2015, 08). A well balanced and entropy conservative discontinuous galerkin spectral element method for the shallow water equations. *Applied Mathematics and Computation*, *272*. doi: 10.1016/j.amc.2015.07.014

Gassner, G. J., & Winters, A. R. (2021). A novel robust strategy for discontinuous galerkin methods in computational fluid mechanics: Why? when? what? where? *Frontiers in Physics*, *8*, 612. Retrieved from `https://www.frontiersin.org/article/10.3389/fphy.2020.500690` doi: 10.3389/fphy.2020.500690

Gassner, G. J., Winters, A. R., & Kopriva, D. A. (2016). Split form nodal discontinuous galerkin schemes with summation-by-parts property for the compressible euler equations. *Journal of Computational Physics*, *327*, 39-66.

Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999116304259` doi: https://doi.org/10.1016/j.jcp.2016.09.013

Giraldo, F., Hesthaven, J., & Warburton, T. (2002). Nodal high-order discontinuous galerkin methods for the spherical shallow water equations. *Journal of Computational Physics*, *181*, 499-525. doi: https://doi.org/10.1006/jcph.2002.7139

Giraldo, F. X. (2020). *An introduction to element-based galerkin methods on tensor-product bases: Analysis, algorithms, and applications.* Springer, Cham.

Giraldo, F. X., Kelly, J. F., & Constantinescu, E. M. (2013, 01). Implicit-explicit formulations of a three-dimensional nonhydrostatic unified model of the atmosphere (numa). *SIAM Journal on Scientific Computing*, *35*. doi: 10.1137/120876034

Golub, G. H., & Loan, C. F. V. (2013). *Matrix computations* (4th ed.). Baltimore, MD, USA: Johns Hopkins University Press. Retrieved from `http://www.cs.cornell.edu/cv/GVL4/golubandvanloan.htm`

Held, I., & Suarez, M. (1994). A proposal for the intercomparison of the dynamical cores of atmospheric general circulation models. *Bulletin of the American Meteorological Society*, *75*(10), 1825–1830. (Copyright: Copyright 2018 Elsevier B.V., All rights reserved.) doi: 10.1175/1520-0477(1994)075⟨1825:APFTIO⟩2.0.CO;2

Herrington, A. R., Lauritzen, P. H., Taylor, M. A., Goldhaber, S., Eaton, B. E., Bacmeister, J. T., … Ullrich, P. A. (2019). Physics–dynamics coupling with element-based high-order galerkin methods: Quasi-equal-area physics grid. *Monthly Weather Review*, *147*(1), 69 - 84. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/147/1/mwr-d-18-0136.1.xml` doi: 10.1175/MWR-D-18-0136.1

Hesthaven, J., & Warburton, T. (2007, 01). Nodal discontinuous galerkin methods: Algorithms, analysis, and applications. In (Vol. 54).

Jablonowski, C., & Williamson, D. L. (2011). The pros and cons of diffusion, filters and fixers in atmospheric general circulation models. In P. Lauritzen, C. Jablonowski, M. Taylor, & R. Nair (Eds.), *Numerical techniques for global atmospheric models* (pp. 381–493). Berlin, Heidelberg: Springer Berlin Heidelberg. Retrieved from `https://doi.org/10.1007/978-3-642-11640-7_13` doi: 10.1007/978-3-642-11640-7_13

Karniadakis, G., & Sherwin, S. (2005). *Spectral/hp element methods for computational fluid dynamics.* Oxford University Press, New York.

Kennedy, C. A., & Gruber, A. (2008). Reduced aliasing formulations of the convective terms within the navier–stokes equations for a compressible fluid. *Journal of Computational Physics*, *227*(3), 1676-1700. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999107004251` doi: https://doi.org/10.1016/j.jcp.2007.09.020

Knoth, O., & Wensch, J. (2014). Generalized split-explicit runge–kutta methods for the compressible euler equations. *Monthly Weather Review*, *142*(5), 2067 - 2081. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/142/5/mwr-d-13-00068.1.xml` doi: 10.1175/MWR-D-13-00068.1

Kopriva, D. (2009). *Implementing spectral methods for partial differential equations.* Springer, New York.

Kopriva, D. A. (2006). Metric identities and the discontinuous spectral element method on curvilinear meshes. *Journal of Scientific Computing*, *26*(3), 301. Retrieved from `https://doi.org/10.1007/s10915-005-9070-8` doi: 10.1007/s10915-005-9070-8

Kuang, Z., Blossey, P. N., & Bretherton, C. S. (2005). A new approach for 3d cloud-resolving simulations of large-scale atmospheric circulation. *Geophysical Research Letters*, *32*(2). Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2004GL021024` doi: https://doi.org/10.1029/2004GL021024

Light, D., & Durran, D. (2016). Preserving nonnegativity in discontinuous galerkin approximations to scalar transport via truncation and mass aware rescaling (tmar). *Monthly Weather Review*, *144*(12), 4771 - 4786. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/144/12/mwr-d-16-0220.1.xml` doi: 10.1175/MWR-D-16-0220.1

Margolin, L. G., Smolarkiewicz, P. K., & Sorbjan, Z. (1999). Large-eddy simulations of convective boundary layers using nonoscillatory differencing. *Physica D: Nonlinear Phenomena*, *133*, 390-397.

Marshall, J., Hill, C., Perelman, L., & Adcroft, A. (1997). Hydrostatic, quasi-hydrostatic, and nonhydrostatic ocean modeling. *Journal of Geophysical Research: Oceans*, *102*(C3), 5733-5752. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/96JC02776` doi: https://doi.org/10.1029/96JC02776

Nair, R. D., Thomas, S. J., & Loft, R. D. (2005). A discontinuous galerkin global shallow water model. *Monthly Weather Review*, *133*, 876-888. doi: 10.1175/MWR2903.1

Nguyen, C., Terrana, S., & Peraire, J. (n.d.). Wall-resolved implicit large eddy simulation of transonic buffet over the oat15a airfoil using a discontinuous galerkin method. In *Aiaa scitech 2020 forum.* Retrieved from `https://arc.aiaa.org/doi/abs/10.2514/6.2020-2062` doi: 10.2514/6.2020-2062

Niegemann, J., Diehl, R., & Busch, K. (2012). Efficient low-storage runge-kutta schemes with optimized stability regions. *J. Comput. Phys.*, *231*, 364-372.

Pauluis, O., Frierson, D. M. W., Garner, S. T., Held, I. M., & Vallis, G. K. (2006). The hypohydrostatic rescaling and its impacts on modeling of atmospheric convection. *Theoretical and Computational Fluid Dynamics*, *20*(5), 485–499. Retrieved from `https://doi.org/10.1007/s00162-006-0026-x` doi: 10.1007/s00162-006-0026-x

Pressel, K. G., Mishra, S., Schneider, T., Kaul, C. M., & Tan, Z. (2017). Numerics and subgrid-scale modeling in large eddy simulations of stratocumulus clouds. *J. Adv. Model. Earth Sys.*, *9*, 1342–1365.

Ranocha, H., & Gassner, G. J. (2021). Preventing pressure oscillations does not fix local linear stability issues of entropy-based split-form high-order schemes. *Communications on Applied Mathematics and Computation*. Retrieved from `https://doi.org/10.1007/s42967-021-00148-z` doi: 10.1007/s42967-021-00148-z

Ringler, T. D., Heikes, R. P., & Randall, D. A. (2000). Modeling the atmospheric general circulation using a spherical geodesic grid: A new class of dynamical cores. *Monthly Weather Review*, *128*(7), 2471 - 2490. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/128/7/1520-0493_2000_128_2471_mtagcu_2.0.co_2.xml` doi: 10.1175/1520-0493(2000)128⟨2471:MTAGCU⟩2.0.CO;2

Roe, P. (1981). Approximate riemann solvers, parameter vectors, and difference schemes. *Journal of Computational Physics*, *43*(2), 357-372. Retrieved from `https://www.sciencedirect.com/science/article/pii/0021999181901285` doi: https://doi.org/10.1016/0021-9991(81)90128-5

Ronchi, C., Iacono, R., & Paolucci, P. (1996). The "cubed sphere": A new method for the solution of partial differential equations in spherical geometry. *Journal of Computational Physics*, *124*(1), 93-114. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999196900479` doi: https://doi.org/10.1006/jcph.1996.0047

Siebesma, A., Soares, P., & Teixeira, J. (2007, 04). A combined eddy-diffusivity mass-flux approach for the convective boundary layer. *Journal of The Atmospheric Sciences - J ATMOS SCI*, *64*. doi: 10.1175/JAS3888.1

Souza, A. N., Wagner, G. L., Ramadhan, A., Allen, B., Churavy, V., Schloss, J., . . . Ferrari, R. (2020). Uncertainty quantification of ocean parameterizations: Application to the k-profile-parameterization for penetrative convection. *Journal of Advances in Modeling Earth Systems*, *12*(12), e2020MS002108. Retrieved from `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2020MS002108` (e2020MS002108 10.1029/2020MS002108) doi: https://doi.org/10.1029/2020MS002108

Sridhar, A., Tissaoui, Y., Marras, S., Shen, Z., Kawczynski, C., Byrne, S., ... Schneider, T. (2021). Large-eddy simulations with climatemachine v0.2.0: a new open-source code for atmospheric simulations on gpus and cpus. *Geoscientific Model Development Discussions*, *2021*, 1–41. Retrieved from `https://gmd.copernicus.org/preprints/gmd-2021-335/` doi: 10.5194/gmd-2021-335

Stull, R. (1988). *An introduction to boundary layer meteorology.* Springer Netherlands. Retrieved from `https://books.google.com/books?id=eRRz9RNvNOkC`

Ullrich, P. A., & Jablonowski, C. (2012). Mcore: A non-hydrostatic atmospheric dynamical core utilizing high-order finite-volume methods. *Journal of Computational Physics*, *231*(15), 5078-5108. Retrieved from `https://www.sciencedirect.com/science/article/pii/S0021999112002057` doi: https://doi.org/10.1016/j.jcp.2012.04.024

Van Roekel, L., Adcroft, A. J., Danabasoglu, G., Griffies, S. M., Kauffman, B., Large, W., ... Schmidt, M. (2018). The kpp boundary layer scheme for the ocean: Revisiting its formulation and benchmarking one-dimensional simulations relative to les. *Journal of Advances in Modeling Earth Systems*, *10*(11), 2647-2685. doi: 10.1029/2018MS001336

Wan, H., Giorgetta, M. A., & Bonaventura, L. (2008). Ensemble held-suarez test with a spectral transform model: Variability, sensitivity, and convergence. *Monthly Weather Review*, *136*(3), 1075 - 1092. Retrieved from `https://journals.ametsoc.org/view/journals/mwre/136/3/2007mwr2044.1.xml` doi: 10.1175/2007MWR2044.1

Waruszewski, M., Kozdon, J., Wilcox, L., Gibson, T., & Giraldo, F. X. (2022). Entropy stable discontinuous galerkin methods for balance laws in non-conservative form: Applications to euler with gravity. *Journal of Computational Physics*, *(in press)*.

Wedi, N. P., & Smolarkiewicz, P. K. (2009). A framework for testing global non-hydrostatic models. *Quarterly Journal of the Royal Meteorological Society*, *135*(639), 469-484. Retrieved from `https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.377` doi: https://doi.org/10.1002/qj.377

Williamson, D. L. (2007). The evolution of dynamical cores for global atmospheric models. *Journal of the Meteorological Society of Japan. Ser. II*, *85B*, 241-269. doi: 10.2151/jmsj.85B.241

Winters, A., Kopriva, D., Gassner, G., & Hindenlang, F. (2021). Construction of modern robust nodal discontinuous galerkin spectral element methods for the compressible navier-stokes equations. *CISM International Centre for Mechanical Sciences, Courses and Lectures*, *602*, 117-196.

Zang, T. A. (1991). On the rotation and skew-symmetric forms for incompressible flow simulations. *Applied Numerical Mathematics*, *7*(1), 27-40. Retrieved from `https://www.sciencedirect.com/science/article/pii/0168927491901026` doi: https://doi.org/10.1016/0168-9274(91)90102-6