We sincerely appreciate your questions and suggestions. Based on the issues you identified, we have implemented significant revisions to the manuscript. We commend your scientific rigor throughout this process.

Referee Comment 1

Line 50: The model uses a finite-volume scheme. Please explain why it is characterized as a local-stencil-based model.

Response to Referee:

Thank you for your suggestion. The term "local stencil-based model" refers to the computational characteristic of our finite-volume scheme, where updating a cell's state during explicit time marching requires only information from a local stencil surrounding that cell. This design eliminates the need for global communication, making each time step highly parallelizable. Accordingly, we have revised the sentence to:

"A finite-volume scheme requiring only information from a local stencil surrounding each cell to perform state updates, enabling massively parallel scalability."

Referee Comment 2

Line 52: It should be explicitly stated that the attractive property discussed here arises from the WENO scheme.

Response to Referee:

According to your advice, the new description is written as:

"A WENO (Weighted Essentially Non-Oscillatory) based, adaptive polynomial order reduction for essential non-oscillation."

Referee Comment 3

Figure 4: Panel (b) shows the spatial stencil for a quadratic polynomial, and panel (c) for a quartic polynomial. Please correct the caption.

Response to Referee:

The new version of Figure 4 (Figure 1 in this reply file) is shown as

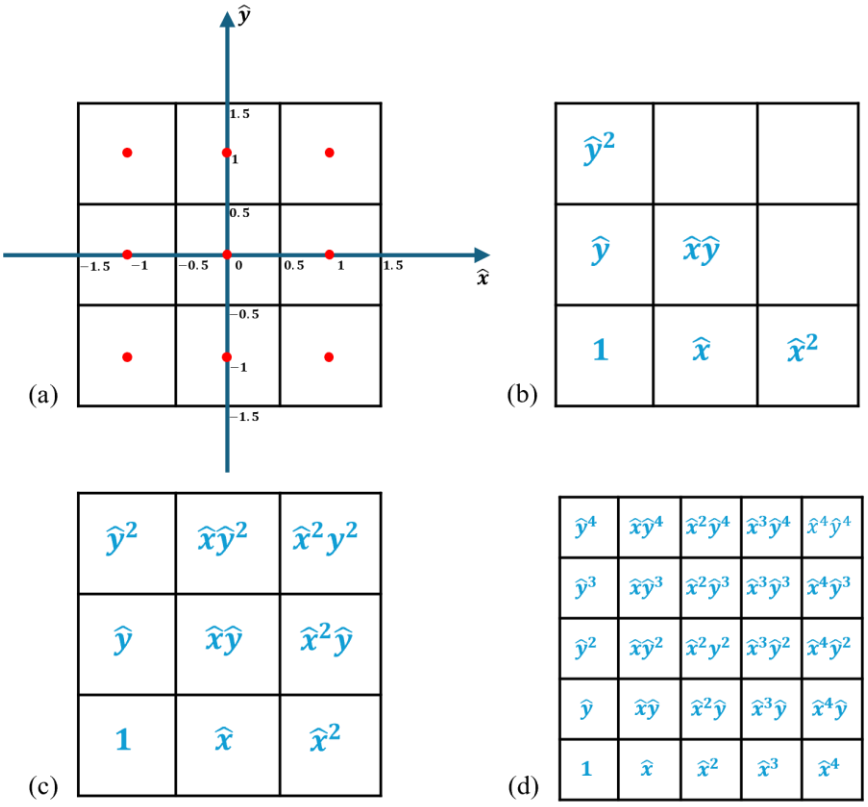

Figure 1 Reconstruction coordinate and polynomial terms on stencils. (a): Local reconstruction coordinate (the red points denote cell centers) (b): 2nd degree polynomial stencil; (c): TPP3 stencil; (d) TPP5 stencil

HOPE employs an equiangular cubed-sphere grid, where each panel undergoes uniform angular discretization into $n_c \times n_c$ cells. In the computational space (equiangular coordinates), each cell spans an angular interval of $\frac{\pi}{2nc}$, therefore

$$\Delta x = \Delta y = \frac{\pi}{2nc}$$

This uniformity ensures that all cells are geometrically identical in the computational space, thereby avoiding the need for cell-specific treatment during reconstruction studies. In the following part of this section, we set a new computational space for reconstruction process. The coordinate system $(\hat{x}, \hat{y})$ is established such that within each reconstruction stencil, the origin $(0,0)$ is located at the stencil center, the central cell spans$[-0.5,0.5]$ in both $\hat{x}$ and $\hat{y}$ directions, as shown in Figure 1 (a). All of the cells have the same size in $\hat{x}, \hat{y}$ directions:

$$\Delta\hat{x} = \Delta\hat{y} = 1$$

On the cubed-sphere grid, a fixed reconstruction scheme yields consistent stencils across all cells. This structural homogeneity renders the reconstruction operation computationally equivalent

to two-dimensional convolution, thereby enabling efficient GPU acceleration through PyTorch's built-in conv2d function.

To construct genuinely 2D reconstructions, the functional form of the reconstruction basis must be selected. A bivariate polynomial of degree d contains $\frac{(d+1)(d+2)}{2}$ terms. As illustrated in Figure 1 (b), the 6 terms of a bivariate quadratic polynomial ($d = 2$) are insufficient to cover a square stencil. To address this, we adopt Tensor Product Polynomials (TPP) as basis functions. We denote a TPP function containing $n \times n$ terms as TPPn. Determining the coefficients of TPPn requires information from a $n \times n$ block of cells. When using a TPP reconstruction stencil of size $n \times n$, HOPE achieves fifth-order accuracy when simulating smooth flow fields. We therefore designate a TPP reconstruction stencil of size $n \times n$ as an n-th order TPP stencil, the 3rd and 5th order TPP stencils are shown in Figure 1 (c)(d).

Referee Comment 4

Figure 7: The ghost cells are interpolated using a two-dimensional procedure, which involves solving a system of equations iteratively. I recommend the authors consider employing a one-dimensional interpolation scheme instead, as the quadrature points in ghost cells are arranged along lines connecting the corresponding points in neighboring cells. One-dimensional interpolation can simplify the interpolation and improve efficiency.

Response to Referee:

Thank you for your suggestion. While a one-dimensional interpolation scheme would indeed be efficient for models employing a dimension-by-dimension reconstruction approach, our testing indicates that a 1D ghost cell interpolation scheme cannot achieve accuracy beyond second order—consistent with findings from Ullrich et al. (2010). This limitation arises because HOPE integrates a two-dimensional reconstruction scheme with cell-boundary flux calculations. Using 1D reconstruction for ghost cell interpolation would cause a loss of two-dimensional information.

To illustrate, consider reconstruction along the x-direction: a pure 1D scheme computes ghost point values not as true pointwise quantities, but as integral averages along the y-direction within the cell. Recovering the actual point values would require an additional deconvolution operation

(Ullrich et al., 2010). Crucially, this process necessitates a wider interpolation stencil than the original scheme.

As you noted, our proposed ghost cell interpolation method appears to require solving linear systems for a closed-form expression. We acknowledge in the Appendix that direct inversion of such large-scale systems is impractical. HOPE's key innovation (Section 3.3.1) circumvents this by leveraging PyTorch's auto-differentiation:

1. The ghost cell interpolation matrix $\mathcal{G}$ is computed *row-wise* via automatic differentiation.
2. $\mathcal{G}$ is stored efficiently in Compressed Sparse Row (CSR) format.
3. Ghost cell interpolation then reduces to a single matrix-vector multiplication.

This approach dramatically reduces computational costs while preserving high-order accuracy.

Referee Comment 5

Subsection 3.5: As the model is based on a WENO scheme, I recommend using a TVD Runge-Kutta time integration.

Response to Referee:

Thank you for the valuable suggestion. The pairing of WENO with a TVD Runge-Kutta time integration scheme is indeed appropriate. HOPE is compatible with various explicit Runge-Kutta schemes, and our experiments have comprehensively evaluated multiple such methods. For the test cases presented in this paper, results demonstrate no discernible differences between using TVD Runge-Kutta and the WRF Runge-Kutta formulation.

Referee Comment 6

I suggest including results for solid rotation of a cosine bell along different directions. Please also provide time histories of normalized errors.

Response to Referee:

The Solid Body Rotation Cosine Bell (Case 1 from Williamson (1992)) is widely used to assess noise generated by panel boundaries, as noted by Chen and Xiao (2008), Ullrich et al. (2010).

Figure 2 presents the norm errors for a 12-day simulation at $\alpha = 0$; results for $\alpha = \pi/2$ are identical. The temporal evolution of L_1 and L_2 norm errors does not exhibit a pronounced signature attributable to panel boundaries. In contrast, the L_∞norm error evolution shows significant sensitivity to panel boundaries, varying considerably with grid resolution and reconstruction order. When using low resolution and low reconstruction order (TPP3 with C30 grid), oscillations induced by panel boundaries are relatively weak. However, as the model resolution or reconstruction order increases, the influence of panel boundaries on the L_∞ norm error manifests as a distinct four-peak pattern, corresponding to the four longitudinally aligned panel boundaries of the cubed-sphere grid.

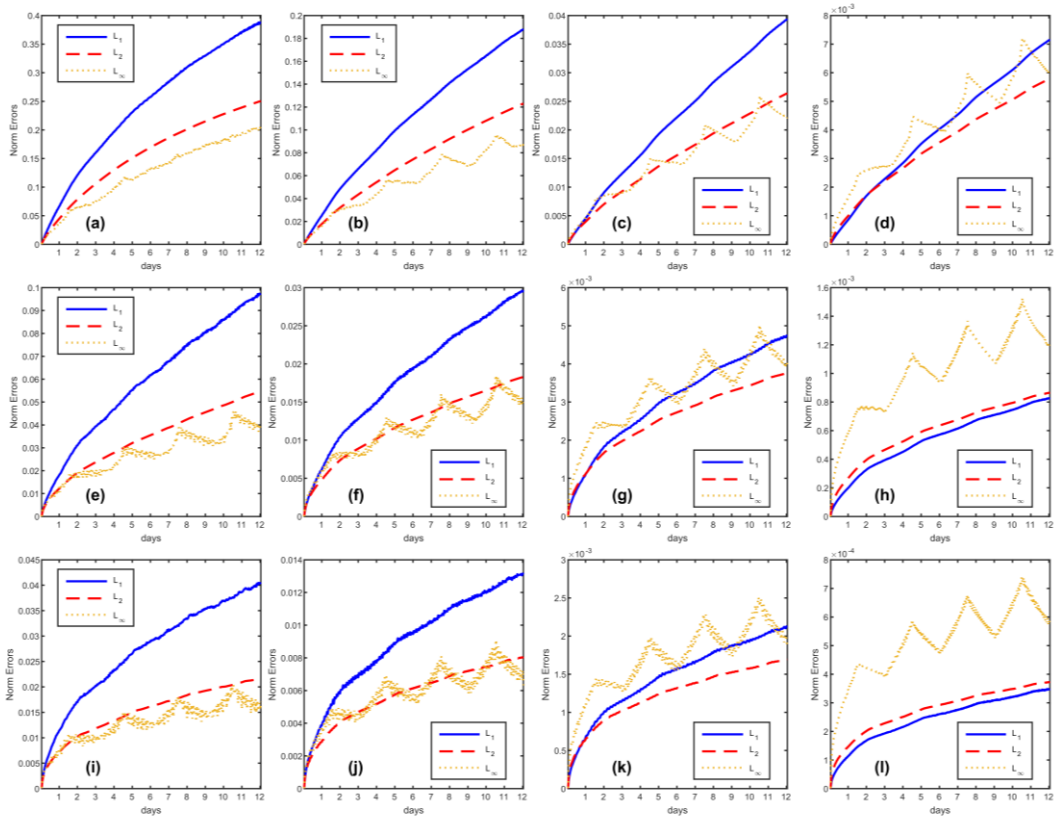

Figure 2 The variation of norm errors during simulation days for the cosine bell advection test case, with direction parameter $\alpha = 0$. The rows represent reconstruction schemes TPP3, TPP5 and TPP7, the columns stand for grid C30, C45, C90 and C180.

Figure 3 shows the 12-day simulation norm errors for $\alpha = \pi/4$. In this test configuration, the cosine bell initially moves alone the interface between Panel 1 and Panel 5, and subsequently moves along the interface between Panel 3 and Panel 6. The temporal evolution of L_1 and L_2 norm errors display two gentle peaks, corresponding to the errors generated as the cosine bell crosses

these panel interfaces. Similar to Figure 2, the L_∞ norm error progressively exceeds the L_1 and L_2 norm errors as grid resolution and reconstruction order increase.

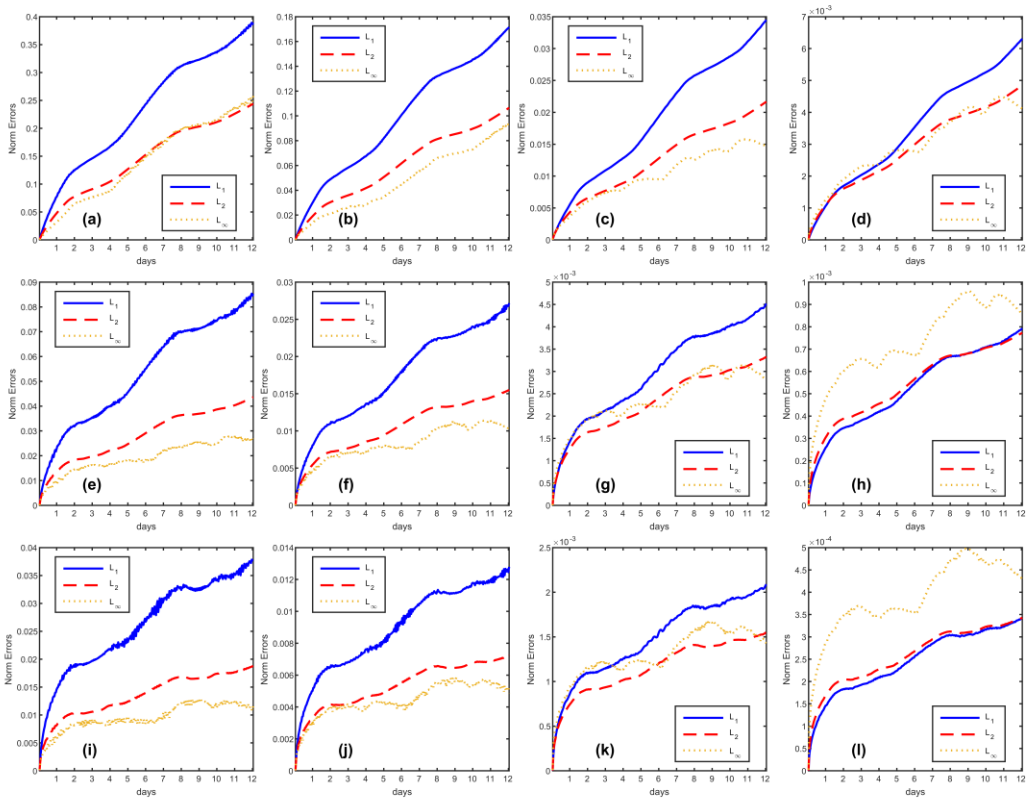

Figure 3 The variation of norm errors during simulation days for the cosine bell advection test case, with direction parameter $\alpha = \pi/4$. The rows represent reconstruction schemes TPP3, TPP5 and TPP7, the columns stand for grid C30, C45, C90 and C180.

Because the Cosine Bell field lacks infinite continuity, the convergence rate of the norm errors cannot exceed second order in our tests, regardless of the reconstruction order employed. This observation aligns with the key point emphasized in our paper: high-order numerical methods achieve their design accuracy only when the flow field is sufficiently smooth. Discontinuities in the flow field violate the fundamental premise of polynomial reconstruction (as discontinuities impair the continuity of higher derivatives, leading to non-convergence of the Taylor series). This inherent sensitivity to smoothness is precisely the factor causing norm errors to be influenced by cubed-sphere panel boundaries. When using low-order reconstruction schemes at low resolutions, the Tensor Product Polynomial (TPP) reconstruction employs lower-degree polynomials and is consequently less sensitive to the smoothness of the flow field. Conversely, high-order TPP reconstruction requires the flow field to possess higher-order continuity to maintain accuracy; it is thus more sensitive to discontinuities. Insufficiently smooth flow fields can introduce numerical

oscillations with high-order schemes. Therefore, while TPP5 and TPP7 yield lower L_∞ norm error magnitudes than TPP3, they exhibit more pronounced oscillations caused by the cubed-sphere panel boundaries.

Referee Comment 7

Williamson test case 2: It would also be helpful to present results obtained using the corresponding linear scheme (i.e., by applying optimal weights in WENO schemes directly). Displaying the absolute error distributions will be helpful to evaluate the grid imprinting.

Response to Referee:

Thank you for your valuable suggestion. We have visualized the simulation errors and obtained meaningful insights.

As shown in Figure 4, errors near the panel boundaries of the cubed-sphere grid are significantly higher than those in the central regions, confirming the presence of grid imprinting. Furthermore, we implemented the AUSM$^+$-up Riemann solver (consistent with the scheme described in Ullrich et al. (2010)) as an alternative to LMARS. While computationally more complex, AUSM$^+$-up substantially reduces simulation errors. Comparative analysis of Figure 4 (a) and (b) demonstrates that the maximum absolute error decreases from 8.792×10^{-5} (LMARS) to 2.4129×10^{-5} (AUSM$^+$-up), while convergence rates remain unchanged.

Performance benchmarks using HOPE's Fortran implementation on a C90 grid show that simulating 12 days with a 200-second integration time step requires 49.4 seconds for LMARS versus 57.34 seconds for AUSM$^+$-up. This demonstrates that Riemann solver selection critically impacts simulation outcomes, consistent with the discussions in Ullrich et al. (2010).

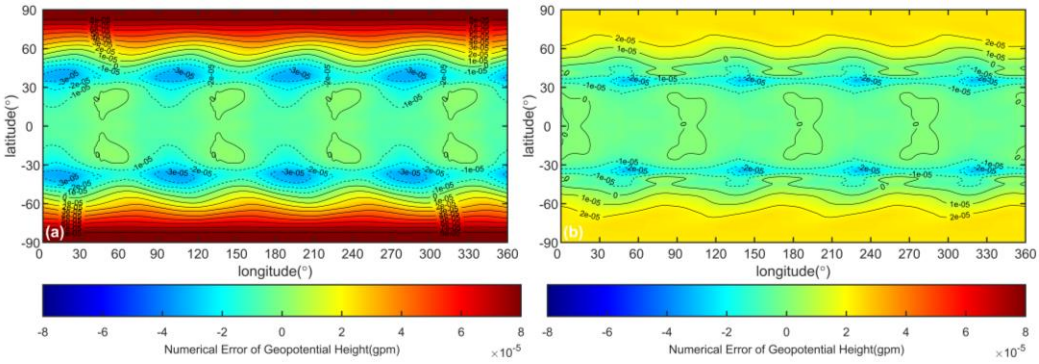

Figure 4 Numerical errors (simulation result minus exact solution) of geopotential height for steady state flow (Williamson test case 2) with Riemann solvers (a) LMARS and (b) AUSM$^+$-up. The

reconstruction scheme is TPP5.

Referee Comment 8

Williamson test cases 5 and 6: I recommend reporting the time histories of normalized errors of total energy and potential enstrophy.

Response to Referee:

We measure the conservation errors by defining the normalized error ϵ_r of the variable η as $\epsilon_r = \frac{I_g(\eta^n) - I_g(\eta^0)}{I_g(\eta^0)}$, where η^0 and η^n stand for η value at initial time and time slot n, respectively. The global integral is defined as:

$$I_g(\eta^n) = \sum_{p=1}^{n_p} \sum_{j=1}^{n_y} \sum_{i=1}^{n_x} \sqrt{G}_{i,j,p} \bar{\eta}_{i,j,p}$$

where $\bar{\eta}_{i,j,p}$ represents the average value of η in cell (i, j, p)

In the 15-day simulation of zonal flow over an isolated mountain the total energy exhibited a gradual increase over the integration time, while both the total potential enstrophy and the total zonal angular momentum showed gradual dissipation as the simulation progressed. The AUSM$^+$-up scheme demonstrated stronger energy dissipation compared to the LMARS scheme, as illustrated in Figure 5。

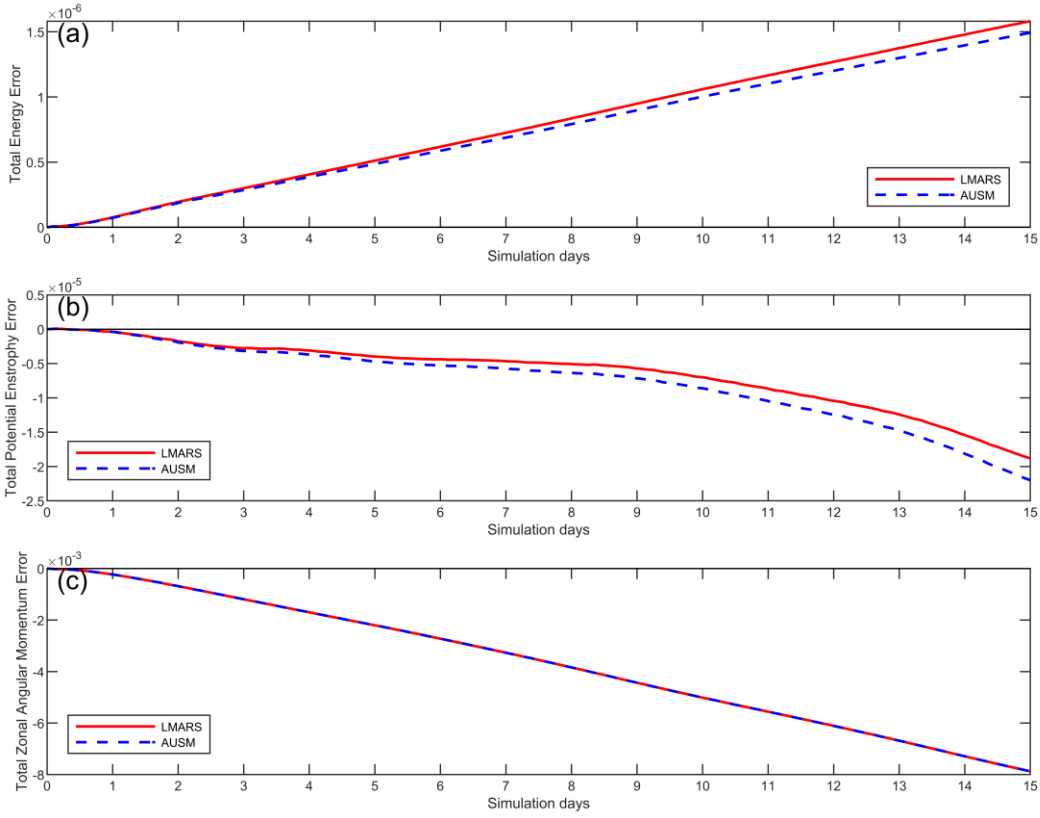

Figure 5 Time series of normalized conservation errors for the zonal flow over isolated mountain simulation on the C90 grid over days 0 to 100. (a) Normalized total energy error. (b) Normalized total potential enstrophy error. (c) Normalized total zonal angular momentum error.

A 100-day simulation of the Rossby-Haurwitz wave was conducted using a C90 grid (1° resolution). The total energy simulated with the TPP3, TPP5, TPP7, and TPP9 schemes underwent dissipation to varying degrees. By day 100, the normalized total energy errors reached $-1.49 \times 10^{-3}, -1.33 \times 10^{-5}, -1.71 \times 10^{-6}, -4.20 \times 10^{-7}$, respectively, indicating significantly stronger dissipation for the TPP3 scheme compared to the other higher-order schemes Figure 6 (a)。Figure 6 (b) presents a scaled view of the energy evolution for TPP5, TPP7, and TPP9, clearly demonstrating that increasing the reconstruction order progressively reduces energy dissipation. Furthermore, following the RH wave collapse, a significant drop in total energy was observed for the TPP5 scheme (after approximately 90 days) and the TPP7 scheme (after approximately 95 days).

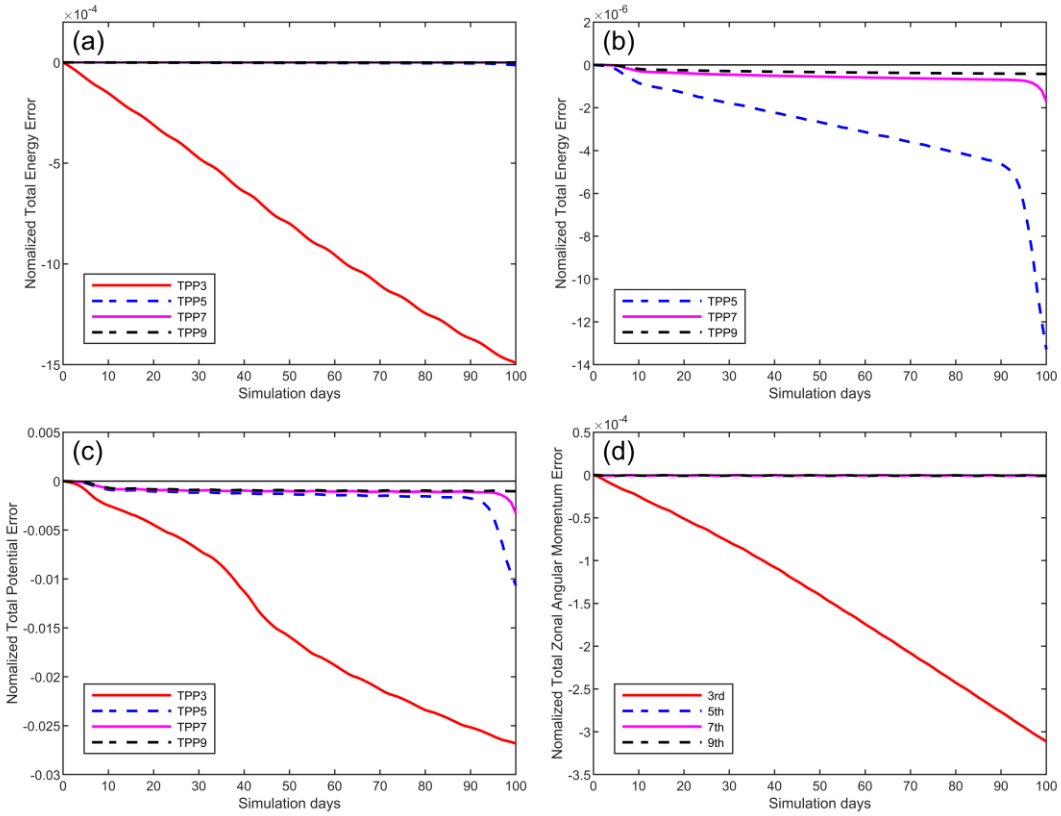

Figure 6 Time series of normalized conservation errors for the Rossby-Haurwitz wave simulation on the C90 grid over days 0 to 100, with LMARS scheme as Riemann solver. (a) Normalized total energy error for TPP3, TPP5, TPP7 and TPP9. (b) The total energy normalized error for TPP5, TPP7 and TPP9. (c) Normalized potential enstrophy error for TPP3, TPP5, TPP7 and TPP9. (d) Normalized total zonal angular momentum error for TPP3, TPP5, TPP7 and TPP9.

Analysis of the normalized total potential enstrophy error (Figure 6 (c)) and the normalized zonal angular momentum error (Figure 6 (d)) over time yields conclusions consistent with those for total energy. Specifically, the TPP3 scheme exhibited substantially higher dissipation than the higher-order schemes, confirming that employing higher-order reconstruction schemes effectively minimizes dissipation. Notably, significant dissipation surges occurred in these quantities following the RH wave collapse.

Referee Comment 9

Genuine 2D scheme: The manuscript emphasizes the benefits of using a genuine two-dimensional discretization. The benefits should be demonstrated through Williamson's standard test suite, rather than a dam-break problem, which is not representative of global atmospheric dynamics. Additionally, please quantify the computational cost differences between the dimension-by-dimension and genuinely 2D schemes.

Response to Referee:

The motivation behind designing the dam break test case was specifically to highlight the differences between the genuine two-dimensional scheme and dimension-by-dimension approaches. These differences proved to be understated in Williamson's Case 5 (zonal flow over an isolated mountain) and Case 6 (Rossby-Haurwitz wave).

We recognize that Multi-Moment Finite Volume Method (MCV) algorithms can achieve high accuracy even with dimension-by-dimension discretization. This is because MCV inherently computes the tendency of point values (PV) through its spatial discretization.

Conversely, in the HOPE model, the prognostic variable is the Volume Integral Average (VIA), not PV. When a dimension-by-dimension reconstruction scheme is applied in this context (specifically for Case 2, steady-state geostrophic flow), the accuracy cannot surpass second order. This limitation parallels the situation mentioned in Fig. 2 of Ullrich et al. (2010). The underlying reason, consistent with our response to Referee Comment 4, is that performing a reconstruction in the x-direction targeting the VIA does not directly yield PV. Instead, it produces line integral averages along the y-direction. Consequently, achieving high accuracy necessitates additional convolution or deconvolution operations.

Furthermore, Shi et al. (2002) report in their Table 3.1 that for the simulation of the 2D Vortex Evolution problem, the convergence rates of genuine finite-volume methods and dimension-by-dimension methods are similar. However, the L_∞ error of the genuine method is significantly lower.

In our own performance tests using WENO with a stencil width of 5 on a C90 grid, simulating 1 day with a 200s time step via RK3 TVD, the dimension-by-dimension scheme took 16.4 seconds, while the genuine 2D scheme required 80.9 seconds (Fortran code, Dual E5-2699V4 processors). It is crucial to note that the current research primarily demonstrates the feasibility of the scheme; the model implementation has not undergone high-performance optimization. Therefore, these timing results should be considered preliminary, and significant improvements in computational efficiency are anticipated during future development.

References

Chen, C. and Xiao, F.: Shallow Water Model on Cubed-Sphere by Multi-Moment Finite Volume Method, Journal of Computational Physics, 227, 5019-5044, 10.1016/j.jcp.2008.01.033, 2008.

Shi, J., Hu, C., and Shu, C.-W.: A Technique of Treating Negative Weights in WENO Schemes, Journal of Computational Physics, 175, 108-127, 10.1006/jcph.2001.6892, 2002.

Ullrich, P. A., Jablonowski, C., and van Leer, B.: High-Order Finite-Volume Methods for the Shallow-Water Equations on the Sphere, Journal of Computational Physics, 229, 6104-6134, 10.1016/j.jcp.2010.04.044, 2010.