Finite difference schemes for several PDE's for the atmosphere and the ocean

Jean-Pierre Croisille

Université de Lorraine, CNRS, IECL, F-57000 Metz, France

May 23 2023

Outline

Joined work with

- J.-B. Bellet (Metz), M. Brachet (Poitiers), B. Portelenelle (Troyes)
- M. Ben-Artzi (Jerusalem), D. Fishelov (Tel Aviv)

Equiangular Cubed Sphere

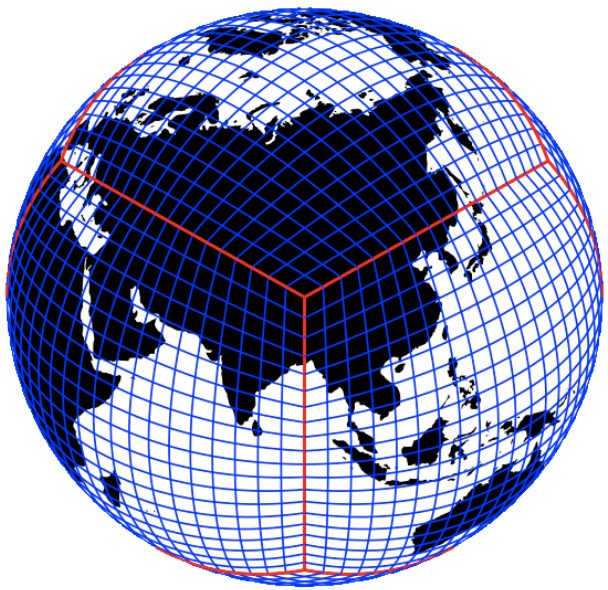
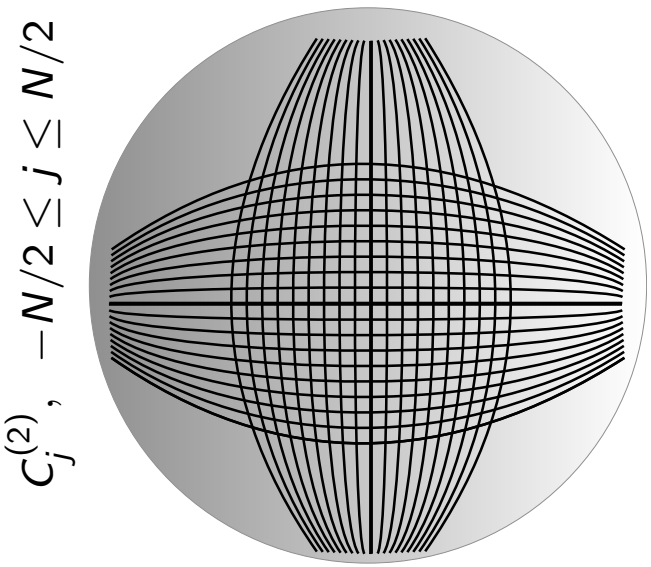


Figure: The equiangular Cubed Sphere.

Equiangular Cubed Sphere (cont.)



- Grid points located at the intersection of two series of great circles.
- The integer N is the measure of the spatial accuracy.
- Circles in *vertical* position are $C_i^{(1)}$, $-N/2 \leq i \leq N/2$.
- Circles in *horizontal* position are labeled $C_j^{(2)}$, $-N/2 \leq j \leq N/2$.

The 6 panels

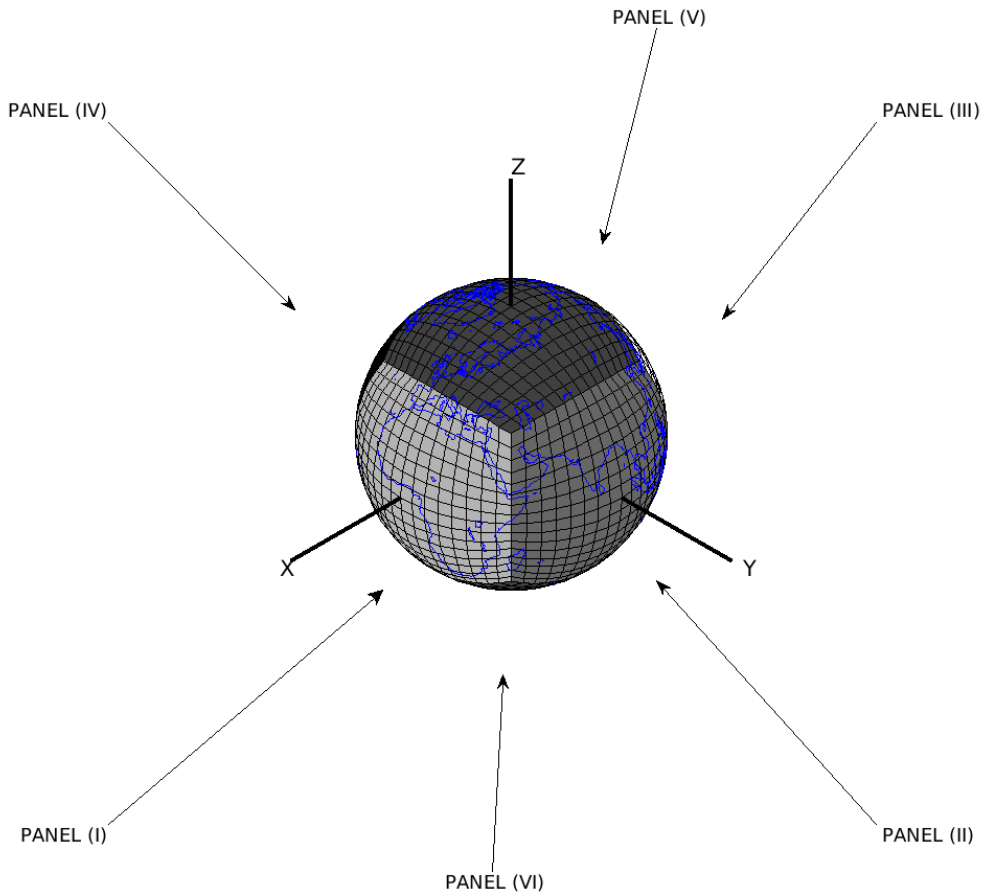


Figure: The 6 panels on the Cubed Sphere grid with resolution $N = 16$ (16^2 cells by panel).

Discrete derivative

Calculate Hermitian derivatives along coordinate great circles on the Cubed Sphere:

- "horizontal" coordinate ξ
- "vertical" coordinate η.

Exemple: fourth order discrete derivative formula

$u_{x,j} \simeq u'(x_j)$ is defined by

$$\frac{1}{6}u_{x,j-1} + \frac{2}{3}u_{x,j} + \frac{1}{6}u_{x,j+1} = \frac{u_{j+1} - u_{j-1}}{2h} \tag{1}$$

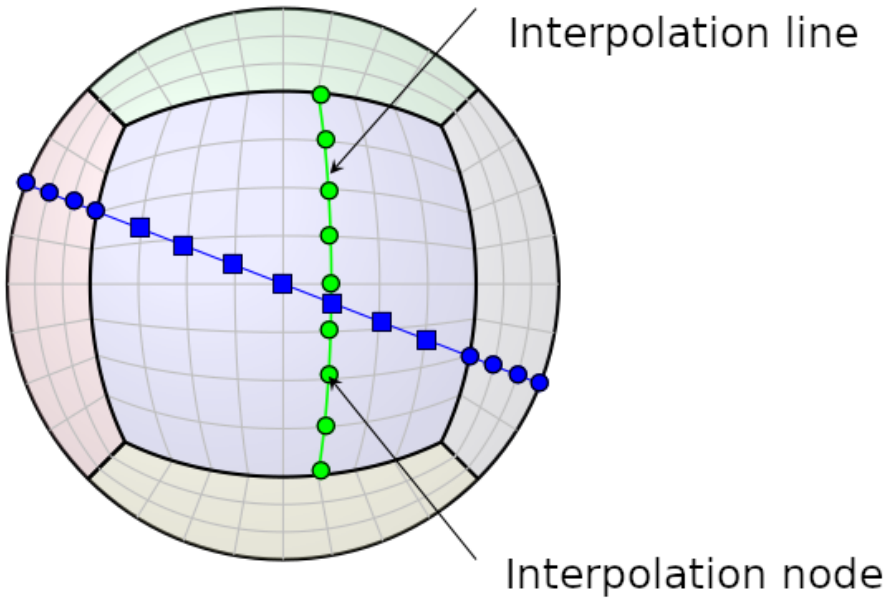
or equivalently

$$u_{x,j} = \left[(I + \frac{h^2}{6}\delta_x^2)^{-1}\delta_x\right]u_j \tag{2}$$

Truncation error:

$$u_{x,j} = u'(x_j) - \frac{1}{180}\partial_x^{(5)}u(x_j)h^4 + O(h^6) \tag{3}$$

Data along a great circle



The Hermitian discrete derivative is applied to the data located at the nodes in blue.

Discrete gradient at the nodes of the CS

- Continuous gradient $\boldsymbol{\nabla}_T(\boldsymbol{x})u$ on panels FRONT and BACK is expressed as

$$\boldsymbol{\nabla}_T u(\boldsymbol{x}) = \frac{\partial u}{\partial \xi}(\boldsymbol{x})_{|\eta}\boldsymbol{g}^{\xi}(\boldsymbol{x}) + \frac{\partial u}{\partial \eta}(\boldsymbol{x})_{|\xi}\boldsymbol{g}^{\eta}(\boldsymbol{x}) \tag{4}$$

- A discrete analog is

$$\boldsymbol{\nabla}_{T,h} u_{i,j} = \underbrace{u_{\xi,i,j}}_{\xi-\text{Hermitian deriv.}} \boldsymbol{g}^{\xi}_{i,j} + \underbrace{u_{\eta,i,j}}_{\eta-\text{Hermitian deriv.}} \boldsymbol{g}^{\eta}_{i,j} \tag{5}$$

Discrete divergence and curl

- Divergence
  - $\nabla_T \cdot \mathsf{u} = \dfrac{\partial \mathsf{u}}{\partial \xi}_{|\eta} \cdot \mathsf{g}^{\xi} + \dfrac{\partial \mathsf{u}}{\partial \eta}_{|\xi} \cdot \mathsf{g}^{\eta} \approx \nabla_{T,\Delta} \cdot \mathsf{u} = \mathsf{u}_{\xi} \cdot \mathsf{g}^{\xi} + \mathsf{u}_{\eta} \cdot \mathsf{g}^{\eta}$
- Curl
  - $\nabla_T \times \mathsf{u} = \mathsf{g}^{\xi} \times \dfrac{\partial \mathsf{u}}{\partial \xi}_{|\eta} + \mathsf{g}^{\eta} \times \dfrac{\partial \mathsf{u}}{\partial \eta}_{|\xi} \approx \nabla_{T,\Delta} \times \mathsf{u} = \mathsf{g}^{\xi} \times \mathsf{u}_{\xi} + \mathsf{g}^{\eta} \times \mathsf{u}_{\eta}$

  where $(\mathsf{g}^{\xi}, \mathsf{g}^{\eta})$ is the dual basis at (ξ_i, η_j) and h_{ξ}, h_{η}, u_{ξ} and u_{η} are the Hermitian derivatives at points (ξ_i, η_j).
- All formulas are "compact" centered but the resulting approximation is "non local".

Vector form of the spherical SW equations

$$\begin{cases} \dfrac{\partial h^\star}{\partial t} + \nabla_T \cdot (h^\star \mathrm{v}) = 0 \\ \dfrac{\partial \mathrm{v}}{\partial t} + \nabla_T \left(\dfrac{1}{2}|\mathrm{v}|^2 + gh \right) + (f + \zeta)\, \mathrm{n} \times \mathrm{v} = 0 \end{cases} \tag{6}$$

- h is the fluid thickness and v the tangential velocity,
- $h^\star = h - h_s$ with h_s the bottom topography,
- n is the normal exterior vector,
- $\zeta = (\nabla_T \times \mathrm{v}) \cdot \mathrm{n}$ is the vorticity,
- f is the Coriolis parameter (depends on the latitude).

Cubed Sphere finite difference solver

- Method of lines with:
  1. In space: center scheme of order 4 based on great circles.
  2. In time: RK4 or Rosenbrock/exponential with a minimally diffusive space filtering (10th order).
- Numerical solutions of SW after long physical physical time. Short time: 1-10 days, Medium time: 50-100 days, Long time: 500-1000 days.
- Results favourably compare to high order conservative solvers (e.g. FV, DG, SE,...) on many standard test cases.
- Few mathematical results available yet ! Conservation, convergence analysis, etc.

Time approximation and filtering

$$\frac{d}{dt}Q(t) = F(t, Q(t)) \tag{7}$$

is typically approximated by the RK4 time scheme. Let $\mathcal{F}$ be a spatial filter function.

Runge-Kutta order 4 + Spatial filtering

1. $K_1 = F(t^n, Q^n)$,
2. $K_2 = F(t^n + \frac{\Delta t}{2}, Q^n + \frac{\Delta t}{2} K_1)$,
3. $K_3 = F(t^n + \frac{\Delta t}{2}, Q^n + \frac{\Delta t}{2} K_2)$,
4. $K_4 = F(t^n + \Delta t, Q^n + \Delta t K_3)$
5. $\hat{Q}^{n+1} = Q^n + \frac{\Delta t}{6}(K_1 + 2K_2 + 2K_3 + K_4)$
6. $Q^{n+1} = \mathcal{F}(\hat{Q}^{n+1})$ (filtering step).

Conservation properties

- Physical conservation properties are a poseriori evaluated.
- Ansatz: "high order scheme + ($\exists$)high order quadrature rule $\Rightarrow$ many conserved quantities at the continuous level are well preserved up to a certain grid error".

Time invariant averaged quantities

For (h, v) a solution of the SW equations, the following quantities are conserved :

- **mass :** $\int_{\mathbb{S}_a^2} h(t,\mathrm{x}) d\sigma(\mathrm{x})$
- **energy :** $\int_{\mathbb{S}_a^2} \left(\frac{1}{2} g(h^2 - h_s^2) + \frac{1}{2} h|\mathrm{v}|^2 \right) d\sigma(\mathrm{x})$
- **potential enstrophy :** $\int_{\mathbb{S}_a^2} \frac{(f+\zeta)^2}{2gh} d\sigma(\mathrm{x})$

Quadrature over the Cubed Sphere

Discrete quadrature formula :

$$I(f) = a\Delta\xi\Delta\eta \sum_{k=(I)}^{(VI)} \sum_{i,j=-N/2}^{N/2} \sqrt{G_{i,j}^k} f_{i,j}^k \tag{8}$$

Compact schemes in Computational Fluid Dynamics

The scheme is the analog of compact schemes for wave problems in aeroacoustics or LES simulation.

- Cartesian grids
- High order accuracy
- Fully centered schemes: no upwinding

Linearized Shallow Water Equations (LSWE) (at the rest state)

$$\begin{cases} \partial_t \boldsymbol{v}(\boldsymbol{x},t) + g\boldsymbol{\nabla}_T \eta + f\boldsymbol{k}\times\boldsymbol{v} = S_{\boldsymbol{v}} \\ \partial_t \eta(\boldsymbol{x},t) + H\,\mathrm{div}_T\,\boldsymbol{v} = S_\eta \\ \boldsymbol{v}(\boldsymbol{x},0) = \boldsymbol{v}_0(\boldsymbol{x}),\quad \eta(\boldsymbol{x},0) = \eta_0(\boldsymbol{x}) \end{cases} \tag{9}$$

$$\begin{cases} \bullet\ g = \text{ gravity acceleration,} \\ \bullet\ H = \text{ mean thickness of the atmosphere,} \\ \bullet\ f = \text{Coriolis force} \end{cases} \tag{10}$$

The LSWE in climatology

- This is the reference model for linear waves in many settings: $\beta-$ plane, various "chanel" assumptions, full sphere,...
- *N. Paldor: Shallow Water Waves on the Rotating Earth, SpringerBriefs in Earth System Sciences, 2015*

Shamir-Paldor *et al.* test cases for the SW equation

- **The full spherical** LSWE is expressed as

$$\partial_t \boldsymbol{q}(t, \boldsymbol{x}) = A\boldsymbol{q} \tag{11}$$

- "Zonal" traveling wave solution

$$\boldsymbol{q}(t, \boldsymbol{x}) = \hat{\boldsymbol{q}}(\theta) \exp(ik(\lambda - Ct)) \tag{12}$$

- $\hat{\boldsymbol{q}}(\theta)$ deduced from $\psi(\theta)$ from a second order equation (Schrödinger 1D)

$$\psi''(\theta) + F_{\alpha,k,C}(\theta)\psi(\theta) = 0 \tag{13}$$

with $\alpha = gH/(2\Omega a)^2$ and with B.C. $\psi = 0$ at $\theta = \pm\pi/2$.

- The constant α determines the thickness of the atmosphere ("thick" or "thin").
- Eigensolutions are identified as
  - EIG or WIG mode (eastward or westward inertial-gravity) mode
  - Rossby mode

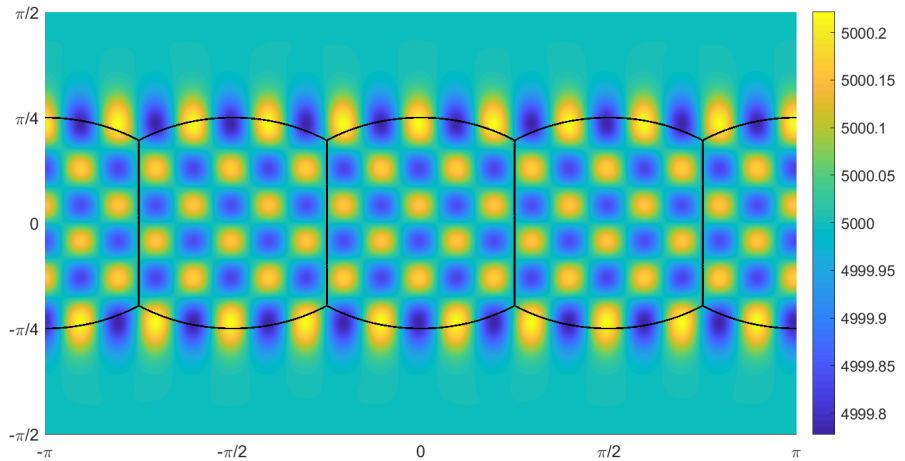
A quasi-analytic series of approximate solutions is derived in explicit form. To be compared with the ones obtained by a nonlinear solver.

A summary of the test cases

1. Test-1a: "barotropic EIG"
2. Test-1b: "barotropic Rossby"
3. Test-2a: "baroclinic EIG"
4. Test-2b: "baroclinic Rossby"

- O. Shamir and N. Paldor, A quantitative test case for global-scale dynamical cores based on analytic wave solutions of the shallow-water equations, Quart. Jour. Roy. Met. Soc., 142, 2016, 2705–2714.
- O. Shamir, I. Yacoby, S.Z. Ziv and N. Paldor, The Matsuno baroclinic wave test case, Geo. Model. Dev., 12, 2019, 2181–2193.

Barotropic EIG wave



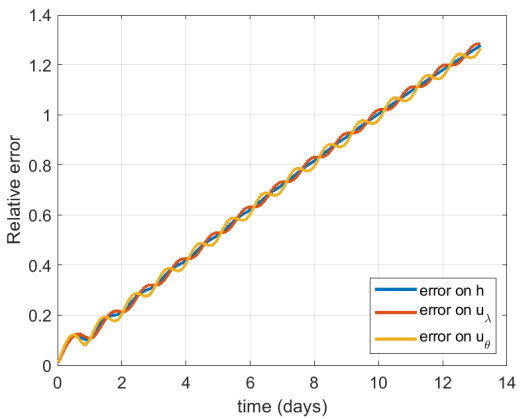


Figure: Test 1-a (barotropic EIG wave). The final time is $t = 13.5$ days (100 periods). Left: total height at final time. Right : History of relative errors. Exponential ERK2 time scheme, $\mathrm{CFL} = 4$. Resolution $6 \times 64 \times 64$.

Barotropic EIG wave, (cont.)

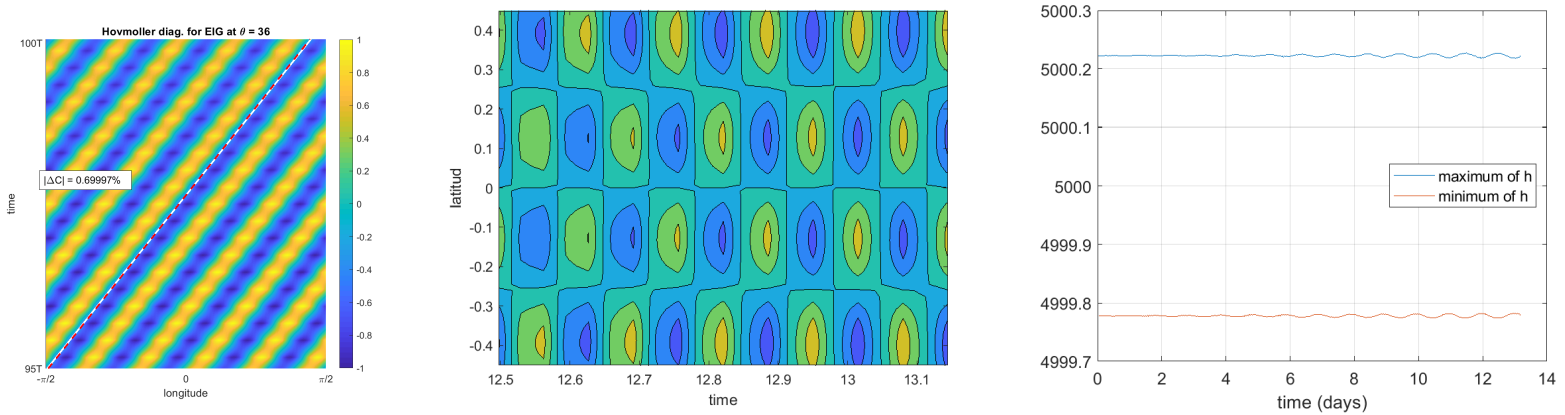


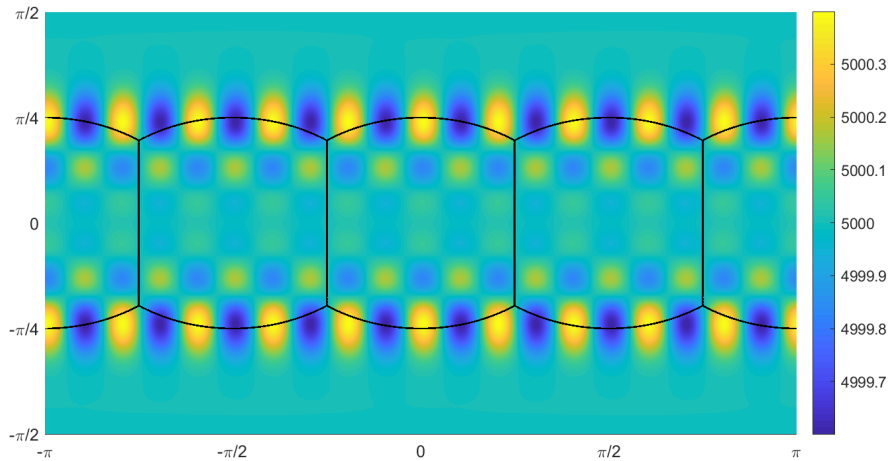
Figure: Test 1-a (barotropic EIG wave). Final time is $t = 13.5$ days (100 periods), ERK2 scheme, $\mathrm{CFL} = 4$. Left: Time-longitude Hovmöller diagram by intersecting zonal velocity at latitude $\theta = 36\,\mathrm{deg}$. Dashed line= analytic solution. Center: Latitude-time Hovmöller diagram by intersecting the zonal velocity at longitude $\lambda = -18\mathrm{deg}$. Right: Max/Min history of the total height h. Resolution $6 \times 64 \times 64$.

Barotropic EIG wave, (cont.)

Time scheme	**Courant Number** CFL	**barotrop. EIG wave**
ERK2	1	0.6389%
	4	0.6999%
	8	0.4405%
RK4	1	unstable
	0.9	0.1468%

Table: Test 1-a (barotropic EIG wave). Dispersion analysis at final time $t = 13.5$ days (100 periods). The relative velocity errors $|\Delta C|$ are reported for various values of CFL. The relative error on the velocity is smaller than 1% in all cases after 100 periods. The RK4 scheme with $\mathrm{CFL} = 0.9$ corresponds to 1808 time iterations and the ERK2 time scheme with $\mathrm{CFL} = 4$ corresponds to 407 time iterations.

Barotropic Rossby wave



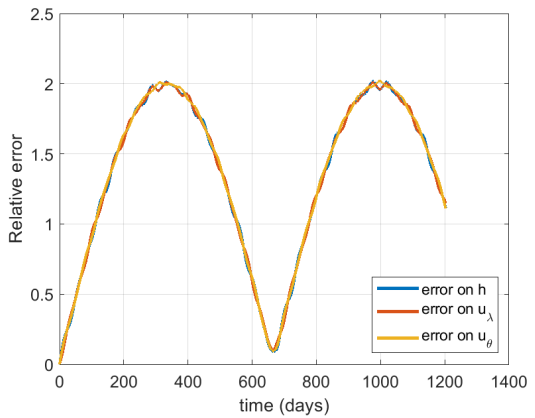


Figure: Test 1-b (barotropic Rossby wave). Final time is $t = 1203$ days (100 periods), ERK2 scheme, 9289 time iterations with $\mathrm{CFL} = 16$. Left: total h at final time. Right : history of the relative errors. Resolution $6 \times 64 \times 64$.

Barotropic Rossby wave, (cont.)

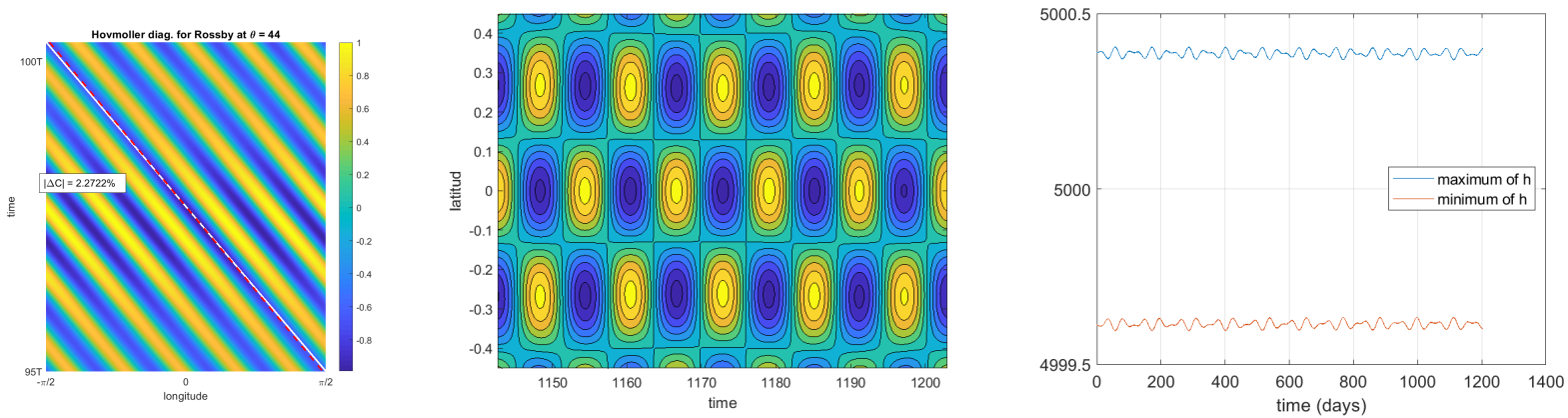


Figure: Test 1-b (barotropic Rossby wave). Final time is 1203 days, 100 periods. ERK2 time scheme at $\mathrm{CFL} = 16$ and 9289 time iterations. Left: time-longitude Hovmöller diagram by intersecting the zonal velocity at $\theta = 44\,\mathrm{deg}$. Center: latitude-time Hovmöller diagram by intersecting the zonal velocity at $\lambda = -18\,\mathrm{deg}$. Right: Maximum and minimum values of the total h over the full simulation. Resolution is is $6 \times 64 \times 64$.

Baroclinic EIG and Rossby waves

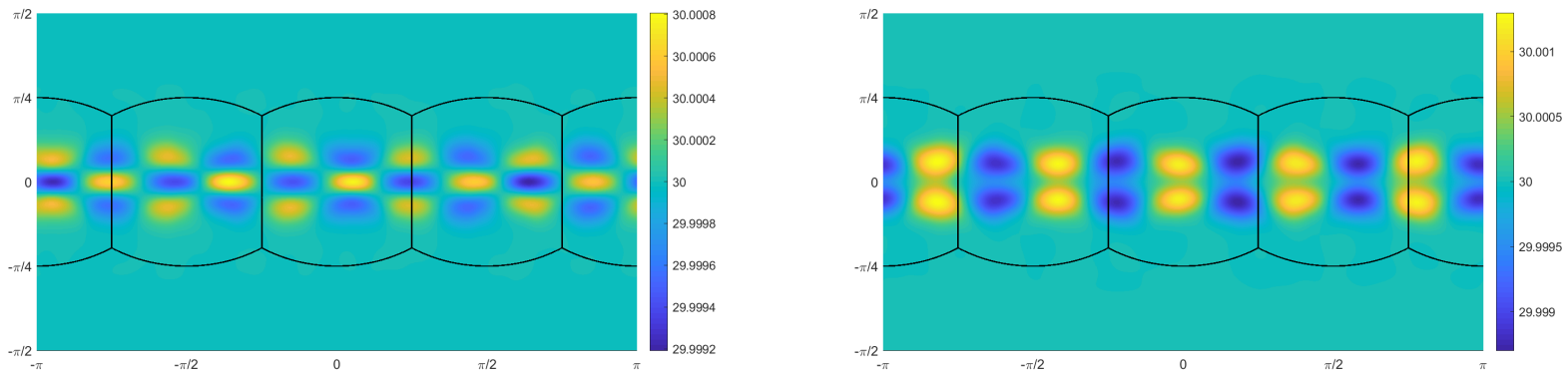


Figure: Left: Test 2-a (baroclinic EIG wave). Total height h at final time $t = 190$ days (100 periods) with $\mathrm{CFL} = 5$ (ERK 2 scheme) and 431 time iterations. Right: Test 2-b (baroclinic Rossby wave). Total height at final time $t = 1850$ days (100 periods) with $\mathrm{CFL} = 5$ and 4313 time iterations. Resolution: $6 \times 64 \times 64$.

Baroclinic EIG and Rossby waves (cont.)

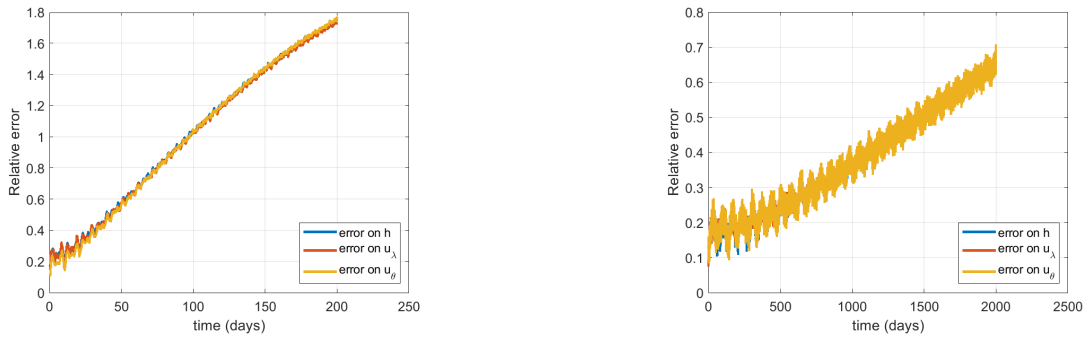


Figure: Left : Test 2-a (baroclinic EIG wave). Final time= $t = 190$ days (100 periods), with $\mathrm{CFL} = 5$ and 431 time iterations. Right : Test 2-b (baroclinic Rossby wave), final time $t = 1850$ days (100 periods) with $\mathrm{CFL} = 5$ and 4313 time iterations. ERK2 time scheme. Resolution=$6 \times 64 \times 64$.

Baroclinic EIG wave

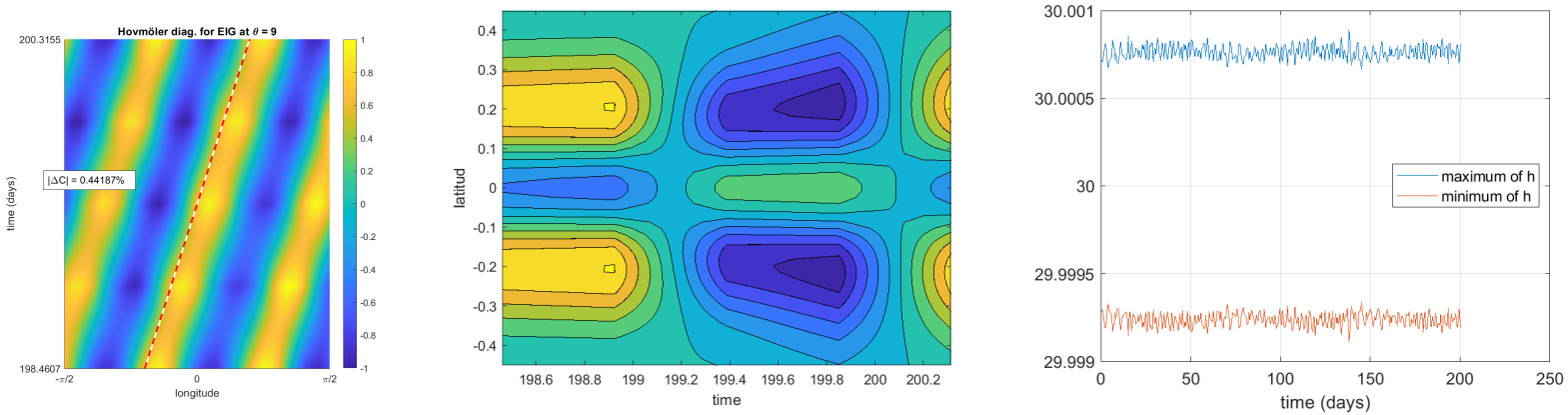


Figure: Test 2-a (baroclinic EIG wave). ERK2 time scheme with $\mathrm{CFL} = 5$. Left: time-longitude Hovmöller diagram by intersecting the zonal velocity at latitude $\theta = 9\,\mathrm{deg}$. Center: Latitude-time Hovmöller diagram by intersecting the zonal velocity at $\lambda = -18\,\mathrm{deg}$. Right: max/min of the total height h over the full simulation (final time: 190 days). Resolution: $6 \times 64 \times 64$.

Baroclinic Rossby wave

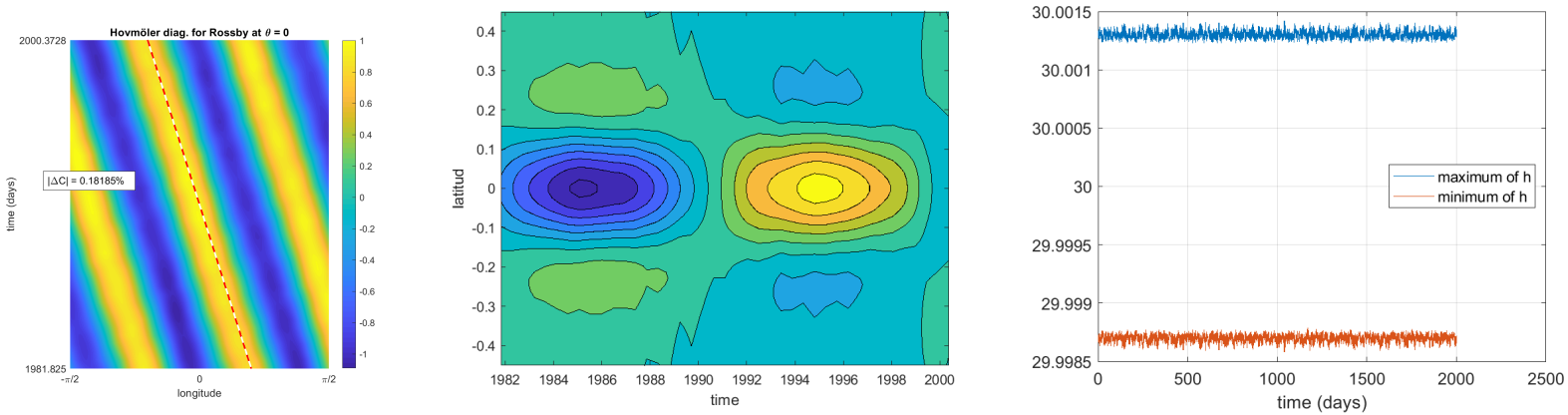


Figure: Test 2-b (baroclinic Rossby wave). ERK2 time scheme with $\mathrm{CFL} = 5$. Left: Hovmöller diagram by intersecting the zonal velocity at $\theta = 0\,\mathrm{deg}$. Center: Latitude-time Hovmöller diagram by intersecting the zonal velocity at longitude $\lambda = -18\,\mathrm{deg}$. Right: Max/min of h over 1850 days (100 periods). Resolution: $6 \times 64 \times 64$.

Numerical conservation

	Test case 1		Test case 2	
	EIG wave	*Rossby wave*	*EIG wave*	*Rossby wave*
CFL	4	16	5	5
Relative Mass Error	1.7666×10^{-14}	2.4476×10^{-8}	2.8297×10^{-15}	1.6564×10^{-12}
Relative Energy Error	2.6838×10^{-14}	4.8947×10^{-8}	2.2746×10^{-16}	3.2573×10^{-12}
Relative Enstrophy Error	2.5541×10^{-14}	3.0250×10^{-8}	2.0419×10^{-14}	2.9696×10^{-13}
Mean value divergence	1.5840×10^{-17}	3.6990×10^{-16}	3.6539×10^{-20}	6.8451×10^{-20}
Mean value vorticity	2.9723×10^{-19}	5.7501×10^{-25}	9.3416×10^{-21}	8.3526×10^{-21}

Table: Test 1-a, Test 1-b, Test 2-a and Test 2-b: conservation of invariant quantities: mass (relative), energy (relative), enstrophy (relative), mean divergence and mean vorticity. ERK 2 scheme Test 1-a (barotropic EIG wave): $\mathrm{CFL} = 4$, $t = 13.5$ days, (100 periods), 407 time iterations. Test 1-b (barotropic Rossby wave): $\mathrm{CFL} = 16$, $t = 1203$ days, (100 periods), 9289 time iterations. Test 2-a (baroclinic EIG wave): $\mathrm{CFL} = 5$, $t = 190$ days, (100 periods), 431 time iterations. Test 2-b (baroclinic Rossby wave): $\mathrm{CFL} = 5$, $t = 1850$ days, (100 periods), 4313 time iterations.

The 1D Munk equation

1D model for wind driven flow

$$(\text{Munk}) \begin{cases} -\beta \partial_x u(x) + \varepsilon \partial_x^4 u(x) = f(x), \quad x \in (a,b) \\ u(a) = u(b) = u'(a) = u'(b) = 0 \end{cases} \tag{14}$$

Length: $\gamma = (\varepsilon/\beta)^{1/3}$

FD scheme

$$(\text{Munk})_h \begin{cases} -\beta \tilde{\delta}_x u_j + \varepsilon \delta_x^4 u_j = f_j^*, \quad 1 \leq j \leq N-1. \\ u_0 = u_N = \widetilde{\delta}_x u_0 = \widetilde{\delta}_x u_N = 0 \end{cases} \tag{15}$$

Compact finite difference operators

HD compact for $\partial_x u$

$$\tilde{\delta}_x = (I + \frac{h^2}{6}\delta_x^2)^{-1}\delta_x \tag{16}$$

DBO compact for ∂_x^4

$$\delta_x^4 == \frac{12}{h^2}\left(\delta_x \circ \tilde{\delta}_x - \delta_x^2\right) \simeq (I + \frac{h^2}{6}\delta_x^2)^{-1}(\delta_x^2)^2 \tag{17}$$

Accuracy

$$\delta_x^4 u^* - (\partial_x^{(4)} u)^* = \frac{1}{720} h^4 (\partial_x^{(8)} u)^* + O(h^6). \tag{18}$$

$$\tilde{\delta}_x u^* - (\partial_x u)^* = -\frac{1}{180} h^4 (\partial_x^{(5)} u)^* + O(h^6). \tag{19}$$

Chekroun Hong Temam test case

Hand manufactured solution

$$u(x) = \Big(1 - \frac{1}{\sqrt{3}} Q_1(x,\gamma) - Q_2(x,\gamma)\Big)(1-x)^2, \tag{20}$$

with

$$\begin{cases} Q_1(x,\gamma) = \exp(-\dfrac{x+1}{2\gamma}) \sin(-\dfrac{\sqrt{3}(x+1)}{2\gamma}) \\ Q_2(x,\gamma) = \exp(-\dfrac{x+1}{2\gamma}) \cos(-\dfrac{\sqrt{3}(x+1)}{2\gamma}) \end{cases} \tag{21}$$

- sequence of solutions with $\beta = 10^{2p}$, $\varepsilon = 10^{-p}$ and $\gamma = (\varepsilon/\gamma)^{1/3} = 10^{-p}$, $p = 0, 1, 2, ...$
- Easy for $p = 0, 1, 2$ and difficult for $p > 2$.
- M.D. Chekroun, Y. Hong and R.M. Temam, Enriched numerical scheme for singularly perturbed quasi-geostrophic equations, J. Comp. Phys.,416, (2020), 109493.

Underresolved boundary layer

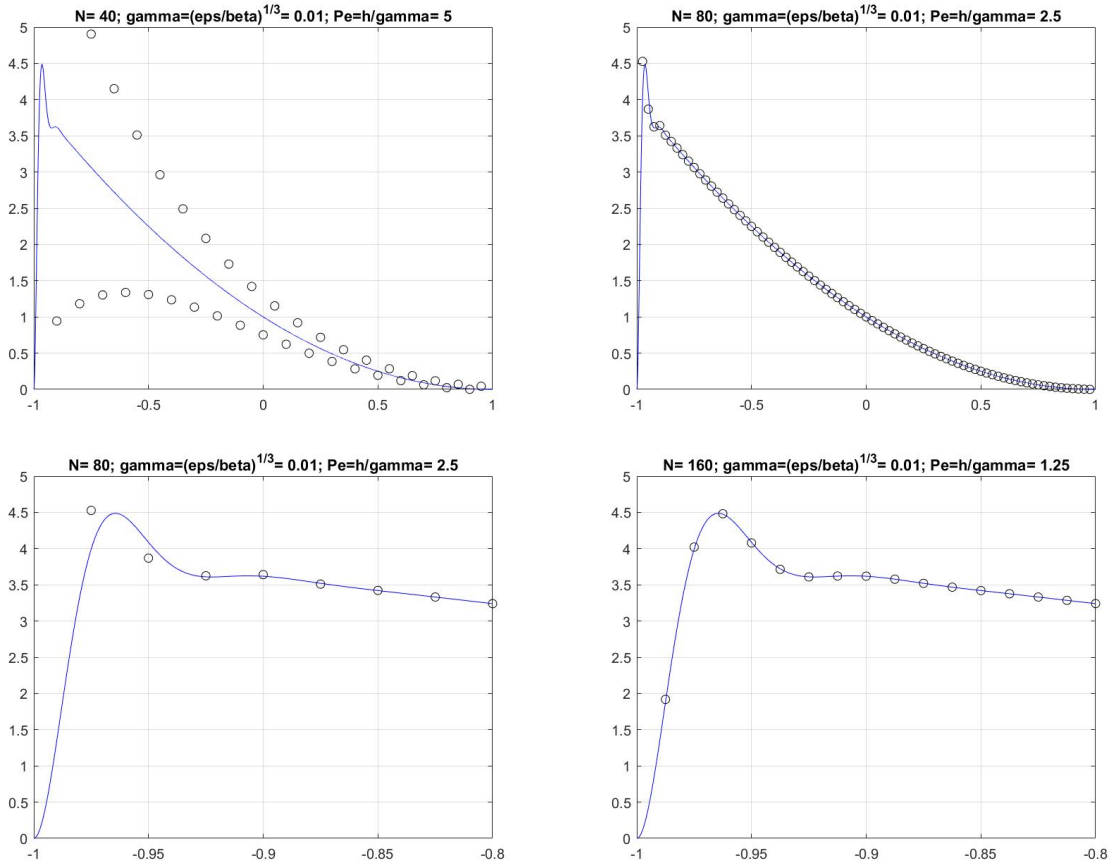


Figure: Left: u with $N = 40$ and $N = 80$. Right: $\partial_x u$ with $N = 40$ and $N = 80$

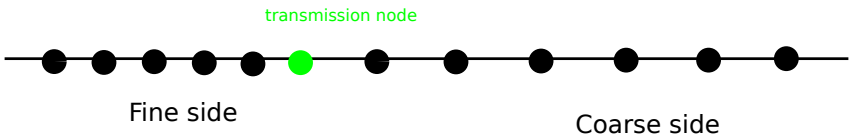
A multiscale scheme

- Stepsize h in the boundary layer zone. This is the fine zone in the boundary layer.
- Stepsize Rh in the ocean zone, with $R > 1$.
- At the transmission node: h on the left, and Rh on the right: a new scheme is required here.

A multiscale scheme (cont.)

$$\begin{cases} -\beta\tilde{\delta}'_x u_j + \varepsilon\delta'^4_x u_j = f^*_j, \quad 1 \le j \le N' - 1, \quad \text{(fine equispaced scheme using } h \\ -\beta\hat{\delta}_x u_{N'} + \varepsilon\widehat{\delta}^4_x u_{N'} = f^*_{N'}, \quad \text{(transmission node scheme: } h \text{ and } Rh) \\ -\beta\tilde{\delta}_x u_j + \varepsilon\delta^4_x u_j = f^*_j, \quad N' + 1 \le j \le N + N' - 1, \quad \text{coarse scheme.} \end{cases} \tag{22}$$

Design of the scheme at the transmission node c



Consider the relation

$$\delta_x^4 u_j = \frac{6}{h^4}(u_{j+1} - 4u_j + 6u_{j-1} - 4u_{j-2} + u_{j-3}) - \delta_x^4 u_{j-2} - 4\delta_x^4 u_{j-1} \tag{23}$$

If j is the transmission node, it is enough to interpolate $u_{j+1} \simeq u(x_j + h)$ on the coarse side.

$$\widetilde{u}_{j+1} = r(u_{j-4}, u_{j-3}, u_{j-2}, u_{j-1}, u_j, u_{j+1}, u_{j+2}, u_{j+3})(x_j + h) \tag{24}$$

with r a $7-$ order Lagrange polynomial. Finally

$$\widehat{\delta}_x^4 u_j \triangleq \frac{6}{h^4}(\widetilde{u}_{j+1} - 4u_j + 6u_{j-1} - 4u_{j-2} + u_{j-3}) - \delta_x^4 u_{j-2} - 4\delta_x^4 u_{j-1} \tag{25}$$

Sharp boundary layer: Case 1: $(\varepsilon/\beta)^{1/3} = 10^{-4}$

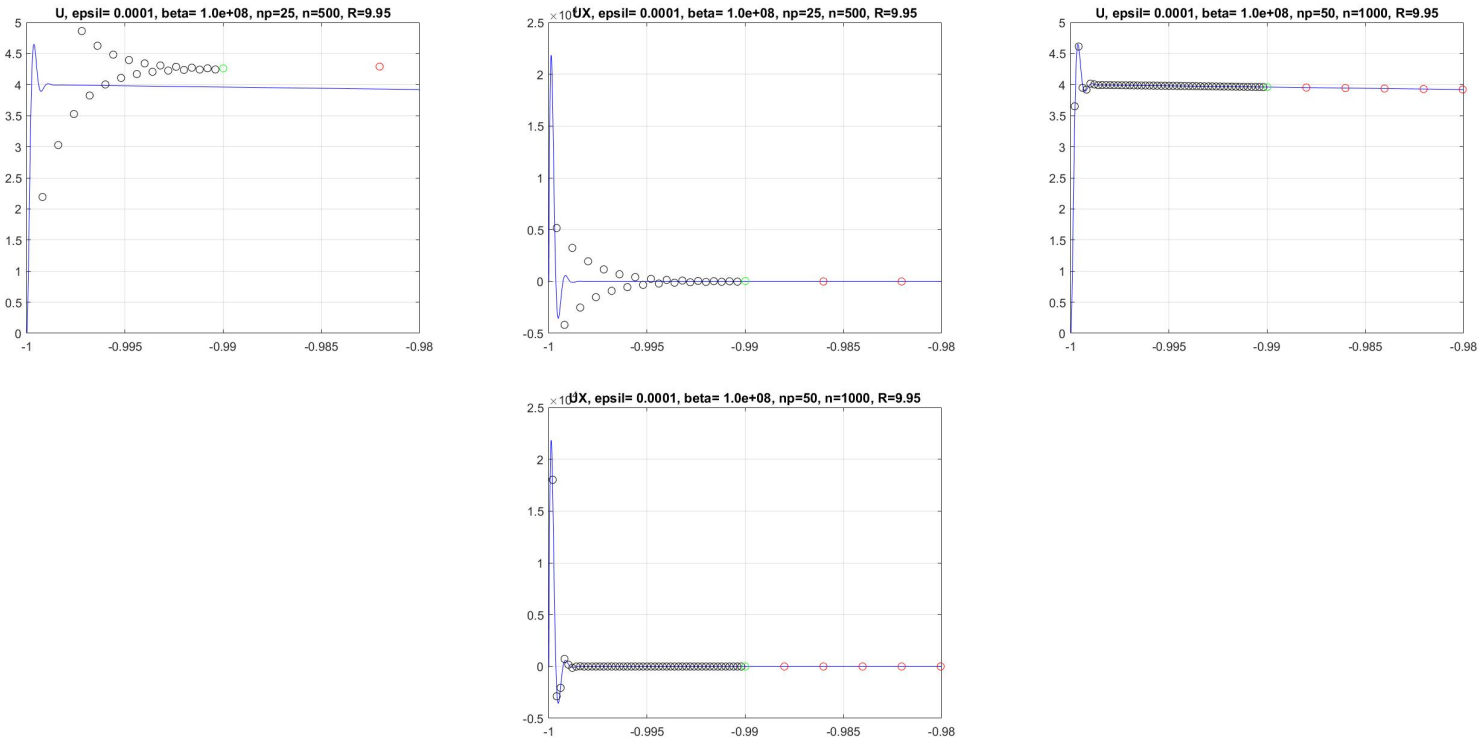


Figure: $p = 4$, $\varepsilon = 10^{-4}$, $\beta = 10^{8}$, $\gamma = 10^{-4}$. Grid 1: $N' = 25$, $N = 500$. Transmission node at $c = -0.99$. The problem is not resolved. Grid 2: $N' = 50$, $N = 1000$: the problem is resolved. $R \simeq 10$.

Sharp boundary layer: Case 1: $(\varepsilon/\beta)^{1/3} = 10^{-4}$ (cont. 1)

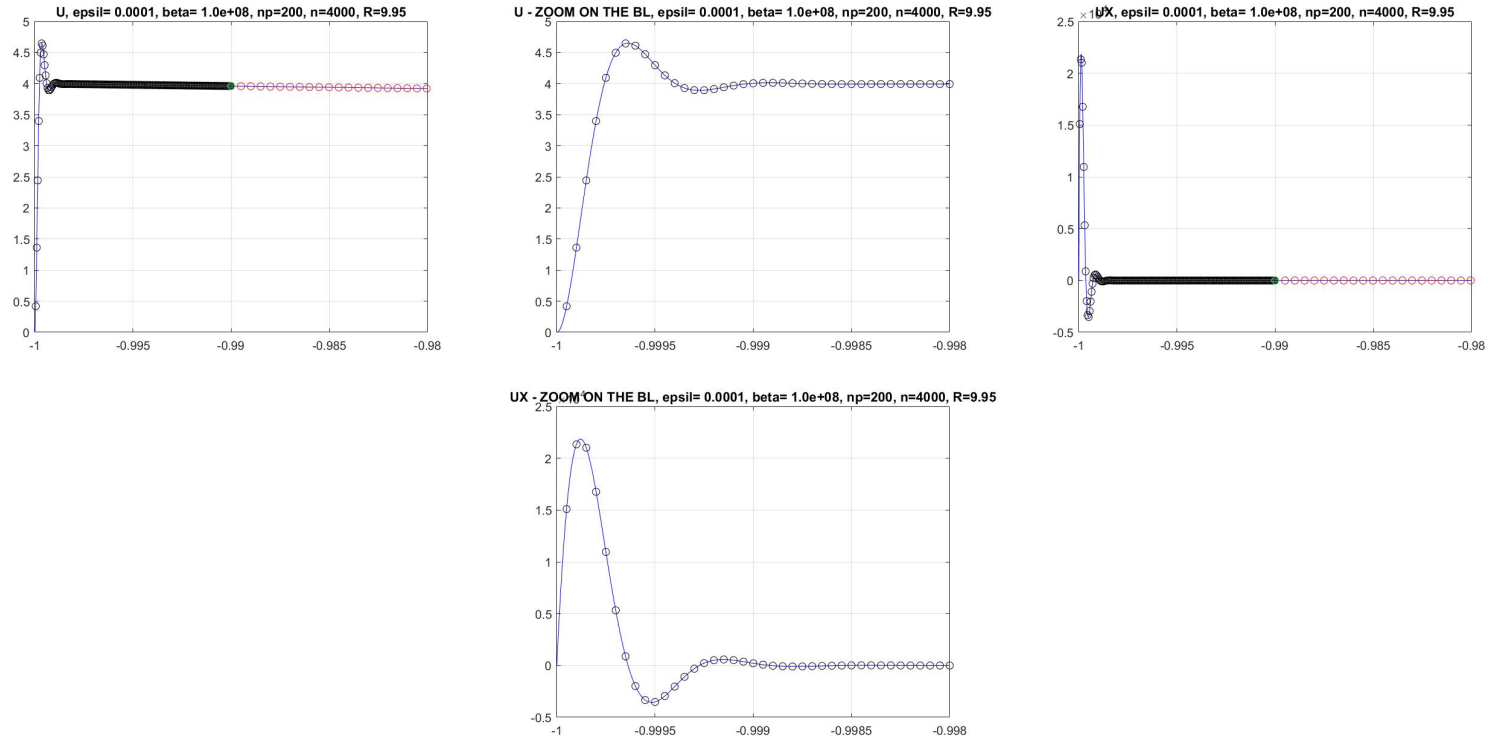
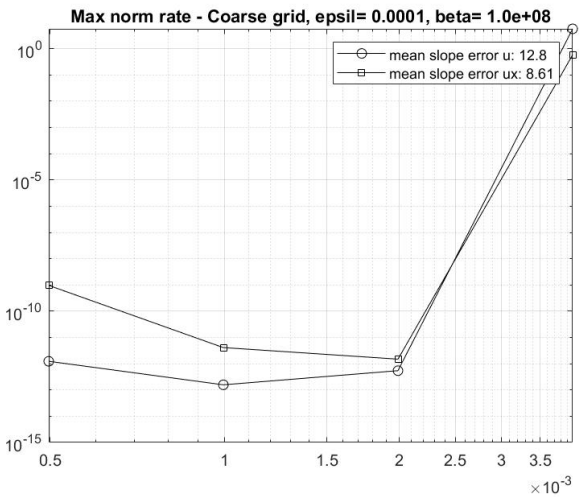


Figure: $p = 4$, $\varepsilon = 10^{-4}$, $\beta = 10^{8}$, $\gamma = 10^{-4}$. Grid 4: $N' = 200$, $N = 4000$: the problem is well resolved.

Sharp boundary layer: Case 1: $(\varepsilon/\beta)^{1/3} = 10^{-4}$ (cont. 2)



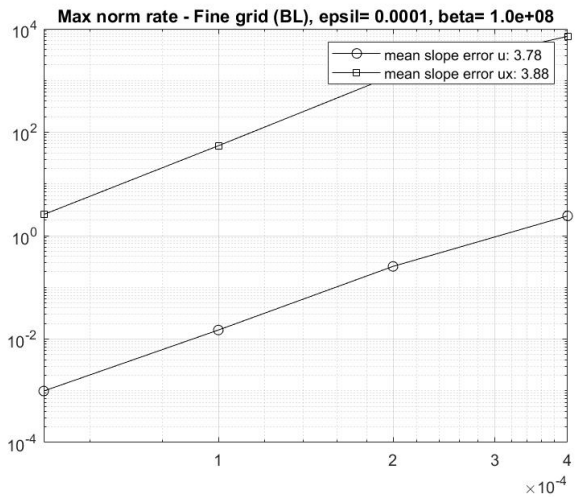


Figure: $p = 4$, $\varepsilon = 10^{-4}$, $\beta = 10^{8}$, $\gamma = 10^{-4}$. Convergence rates.

Sharp boundary layer: Case 2: $(\varepsilon/\beta)^{1/3} = 10^{-5}$

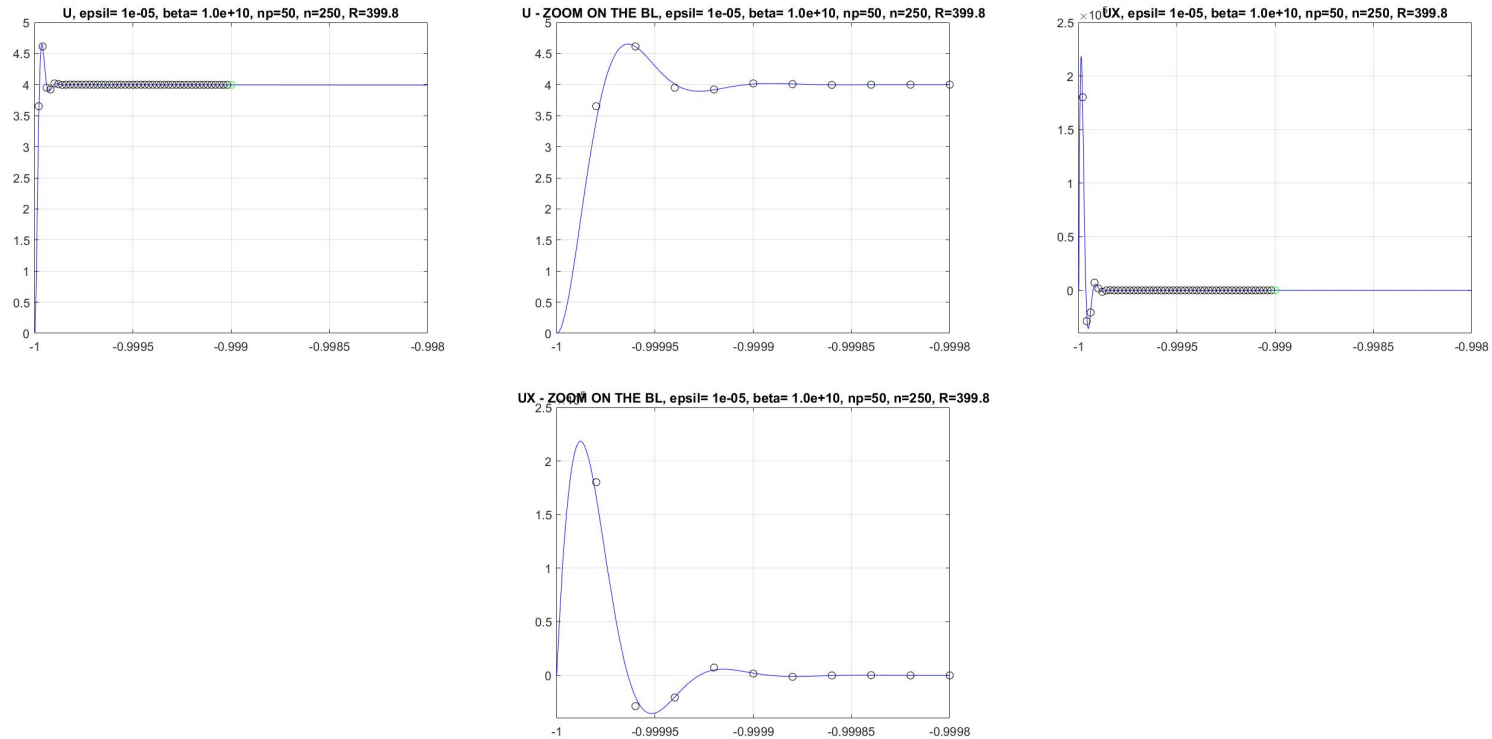


Figure: $p = 5$, $\varepsilon = 10^{-5}$, $\beta = 10^{10}$, $\gamma = 10^{-5}$. Transmission node at $x = -0.999$. Grid 1: $N' = 50$, $N = 250$: the problem is already well resolved.

Sharp boundary layer: Case 2: $(\varepsilon/\beta)^{1/3} = 10^{-5}$, (cont. 1)

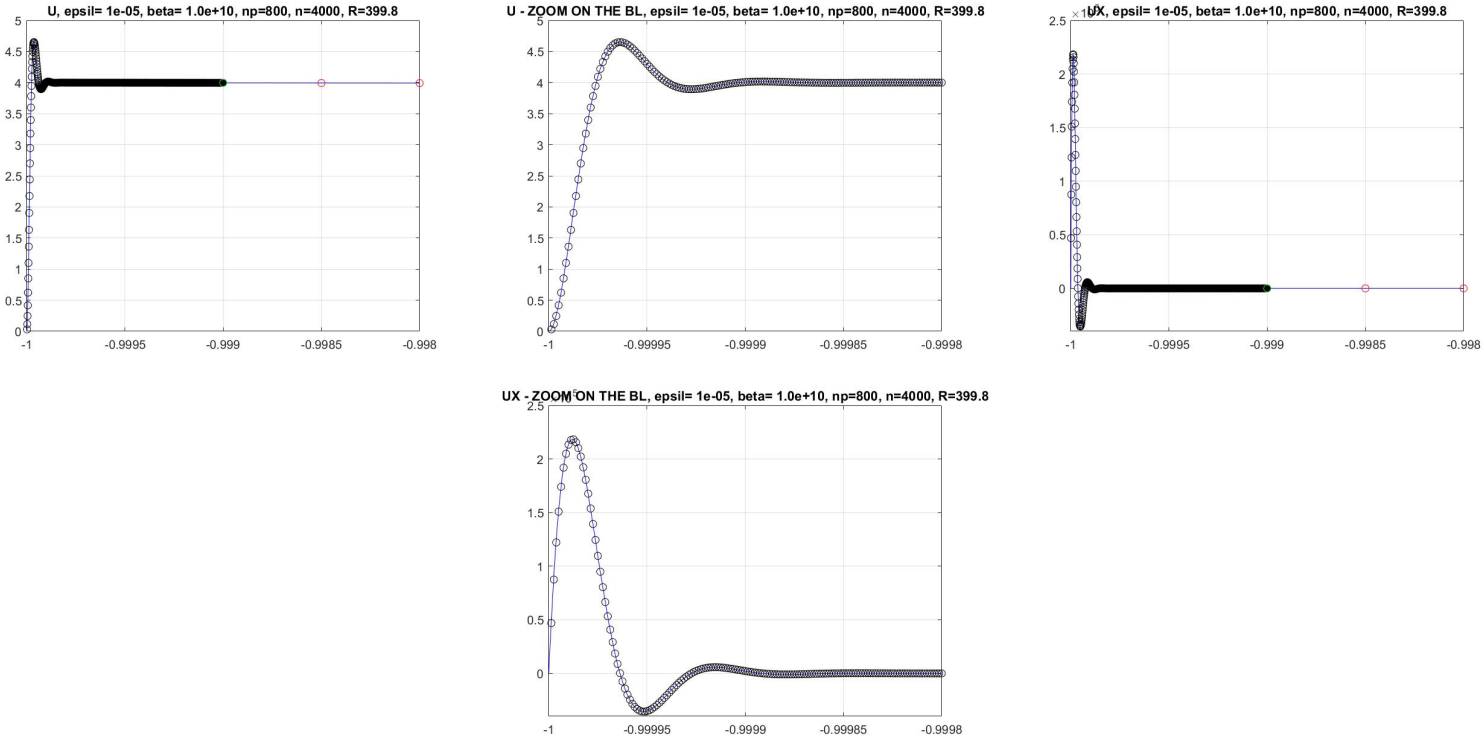
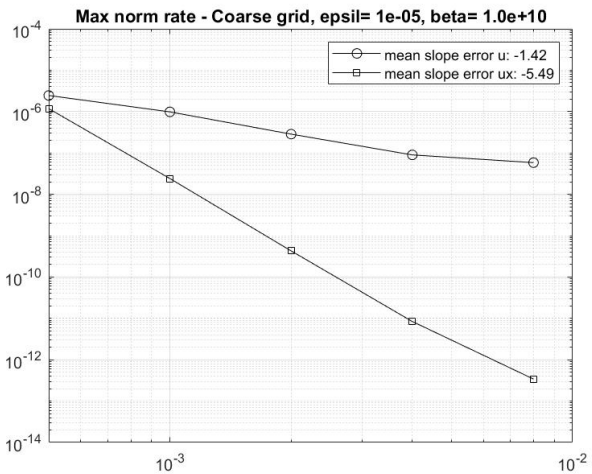


Figure: $p = 5$, $\varepsilon = 10^{-5}$, $\beta = 10^{10}$, $\gamma = 10^{-5}$. Grid 4: $N' = 800$, $N = 4000$: the problem is accurately resolved.

Sharp boundary layer: Case 2: $(\varepsilon/\beta)^{1/3} = 10^{-5}$ (cont. 2)



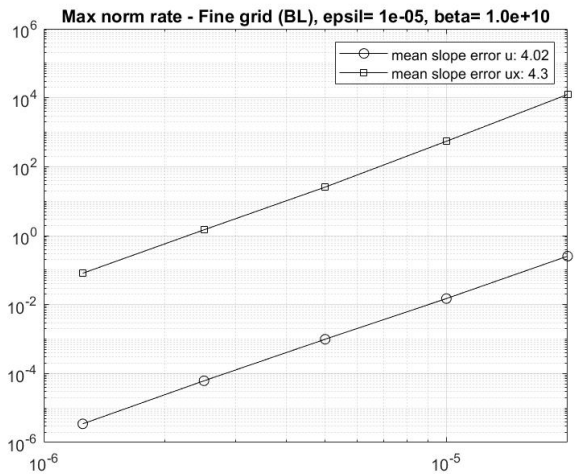


Figure: $p = 5$, $\varepsilon = 10^{-5}$, $\beta = 10^{10}$, $\gamma = 10^{-5}$. Convergence rates.

Empirical rules

- The transmission node must be located sufficiently far from the BL. A factor of at least 10 times the width of the BL seems appropriate.
- The accuracy in the coarse zone must be as high as possible. We try to reach the computer accuracy in the coarse zone. This is a good indicator that the accuracy in the BL will be also good.

The Cubed Sphere: why is it useful ?

What is known

- The Cubed Sphere has the symmetry group of the Cube (J.-B. Bellet). No pole problem.
- The Cubed Sphere is observed to have good approximation properties when combined with PDE's numerical methods: FEM (finite element method), FV (finite volume), DG (discontinuous Galerkin), FD (finite differences).

What is needed

- Need of a canonical Spherical Harmonics (SH) basis analog of the sine/cosine basis on $\mathbb{S}^1$.
- Need of a well conditioned SH basis for the interpolation problem, in which pseudospectral calculus for PDE's is accurate and easily implementable.
- Need of a least squares well conditioned framework.

Spherical Harmonics

- They are the harmonic polynomials in R^3: $p \in \mathrm{R}[x, y, z]$ with $\Delta p = 0$.
- Hilbert basis Y_n^m, $0 \leq |m| \leq n \leq +\infty$ of $L^2(\mathbb{S}^2)$

$$Y_n^m(\boldsymbol{x}(\theta, \phi)) = C_n^m \cdot \cos^{|m|}(\theta) \cdot Q_n^{|m|}(\sin\theta) \cdot \begin{cases} \sin m\phi, & -n \leq m < 0, \\ \cos m\phi, & 0 \leq m \leq n, \end{cases}$$

with

  - $\boldsymbol{x}(\theta, \phi) = (\cos\theta\cos\phi, \cos\theta\sin\phi, \sin\theta)$, $\phi \in [0, 2\pi)$: longitude and $\theta \in [-\pi/2, \pi/2]$: latitude.
  - normalizing constant $C_n^m \in \mathbb{R}$,
  - Legendre polynomial $Q_n^{|m|}(t) = \frac{\mathrm{d}^{|m|+n}}{\mathrm{d}t^{|m|+n}} \frac{1}{2^n n!}(t^2-1)^n$, degree $n - |m|$

- Finite dimensional subspace of the Spherical Harmonics with degree $\leq D$

$$\mathcal{Y}_D = \mathrm{Span}\{Y_n^m,\ |m| \leq n \leq D\}$$

The restriction to any great circle of any $f \in \mathcal{Y}_D$ is a trigonometric polynomial in the angular variable with degree $\leq D$.

Structure of $\mathcal{Y}_D$

- Space $\mathbb{Y}_k$: restriction to $\mathbb{S}^2$ of the harmonic polynomials of degree k in (x, y, z).

$$\dim \mathbb{Y}_k = 2k + 1.$$

- Space $\mathcal{Y}_D = \oplus_{k=0}^{D} \mathbb{Y}_k$, $\dim \mathcal{Y}_D = (D+1)^2$.
- Expansion of a function $u \in L^2(\mathbb{S}^2)$ in $\mathcal{Y}_D$: analog of a trigonometric polynomial expansion of $f \in L^2(\mathbb{S}^1)$ in $\cos kx$, $\sin kx$.

Least squares approximation on the Cubed Sphere with Spherical Harmonics

- Problem Setup:
  1. Fix a maximal degree D (maximal angular frequency). This gives the space $\mathcal{Y}_D$. One has $D \leq \bar{N}$.
  2. For a given set of values $y_j \in \mathbb{R}$ at the Cubed Sphere nodes $\boldsymbol{x}_j$, compute the least squares Spherical Harmonics $f \in \mathcal{Y}_D$ solution of

$$\inf_{f \in \mathcal{Y}_D} \sum_{j=1}^{\bar{N}} |f(\boldsymbol{x}_j) - y_j|^2. \tag{LS}$$

Vandermonde matrix

- Data: $y = [y_j]_{1 \le j \le \bar{N}}$, ($y_j$ =value at $x_j \in \mathrm{CS}_N$)
- Unknown: $\hat{f} = [\hat{f}_n^m]_{|m| \le n \le D}$, $f = \sum_{|m| \le n \le D} \hat{f}_n^m Y_n^m \in \mathcal{Y}_D$
- Associated Vandermonde matrix:

$$A_D = [Y_n^m(\boldsymbol{x}_j)]_{1 \le j \le \bar{N}, |m| \le n \le D} \in \mathbb{R}^{\bar{N} \times (D+1)^2}$$

- Least squares problem

$$\inf_{\hat{f} \in \mathbb{R}^{(D+1)^2}} \|A_D \hat{f} - y\|_2^2, \qquad \text{(LS)}$$

Resolution of the least squares problem

Standard solution

- (LS) admits a unique solution only in the case where A_D is full column-rank or equivalently if $A_D^T A_D$ is positive definite.
- In this case the solution $\hat{f}$ satisfies

$$A_D^\intercal A_D \ \hat{f} = A_D^\intercal y \iff \hat{f} = (A_D^\intercal A_D)^{-1} A_D^\intercal y$$

- Otherwise, some regularization/selection is required to define a solution $\hat{f}$.

Open theoretical problem

Find the largest degree D such that (LS) satisfies:

- A_D is full column-rank.
- Condition number of $A_D^\intercal A_D$ is small.

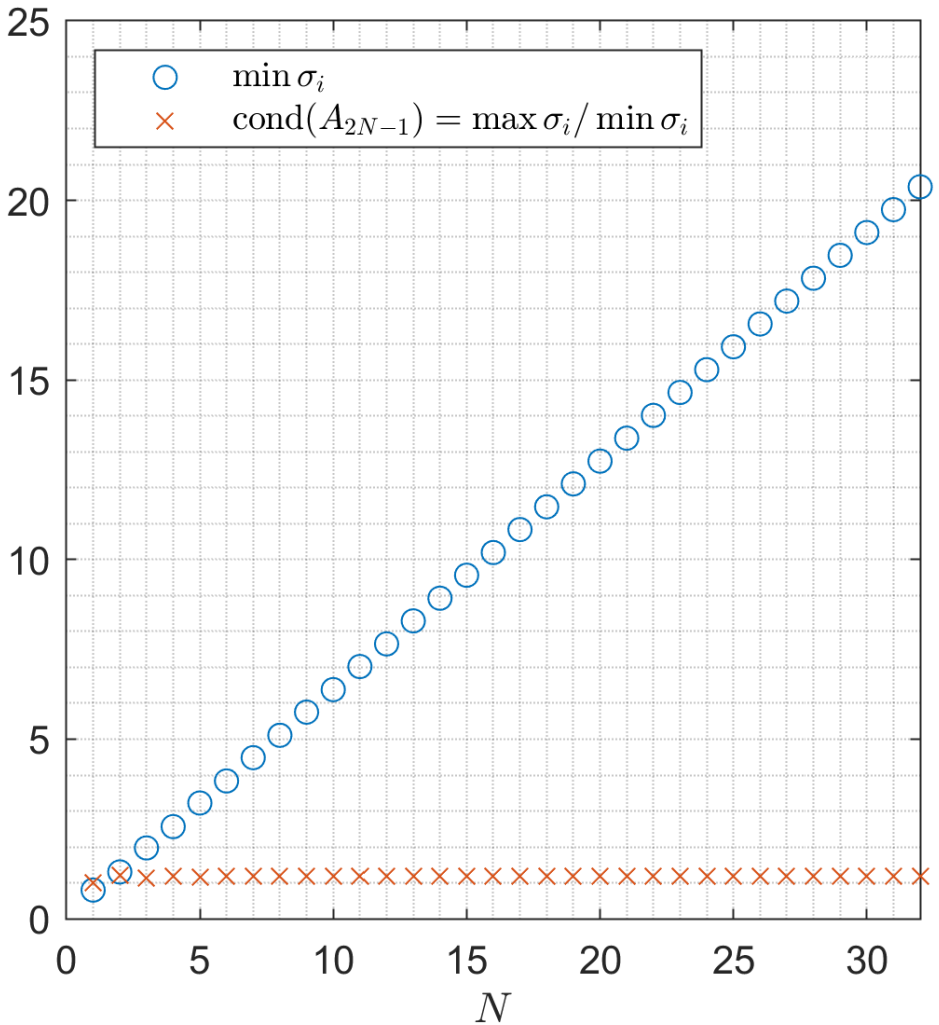
Partial results

The VDM matrix A_D is

$$A_D = [Y_n^m(\boldsymbol{x}_j)]_{1\le j\le \bar{N},\, |m|\le n\le D} \in \mathbb{R}^{\bar{N}\times(D+1)^2}, \quad \bar{N} = 6N^2+2$$

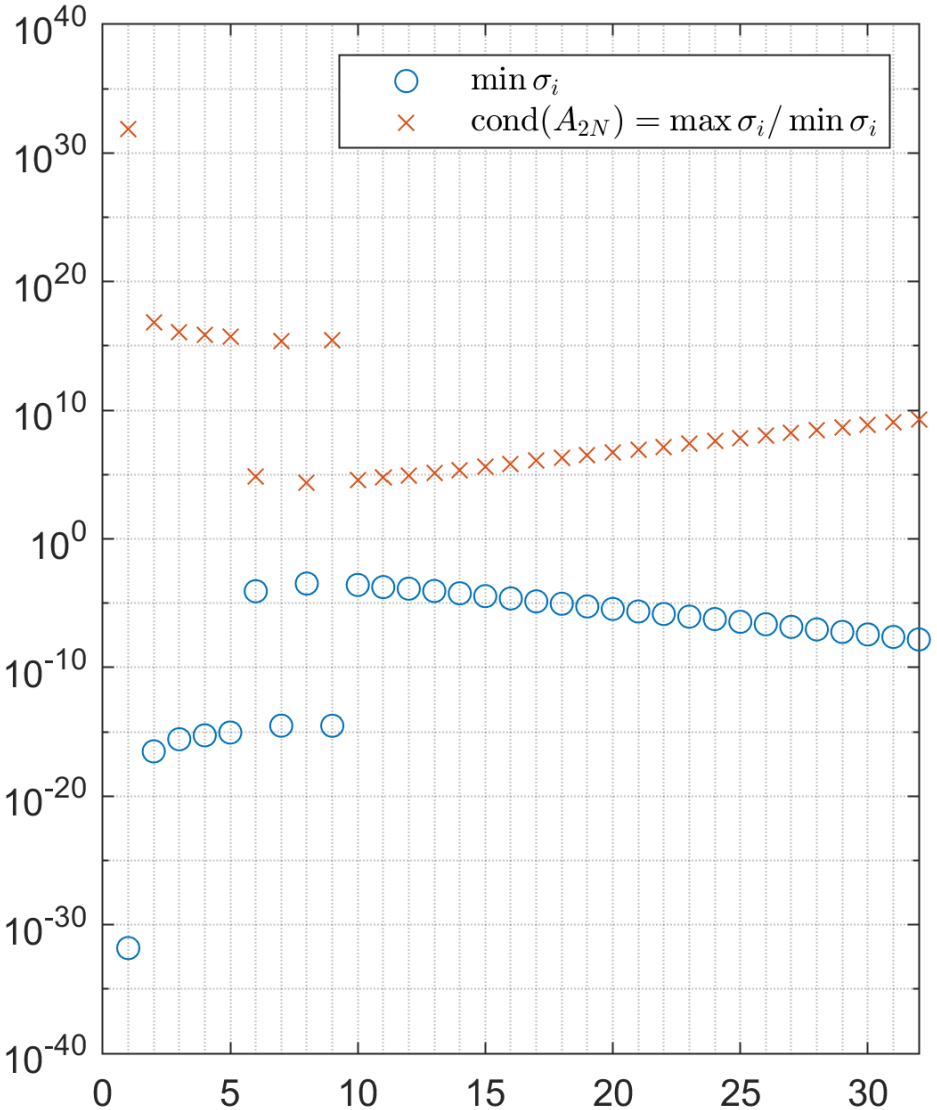
1. $A_0 \in \mathbb{R}^{\bar{N}} \setminus \{0\}$ is full rank (rank 1).
2. A_D full rank $\Rightarrow (D+1)^2 \le \bar{N} \Rightarrow D \le \bar{N}^{1/2} - 1$,
   - Upper bound: $\bar{N}^{1/2} - 1 = \sqrt{6}N - 1 + o(1) \approx 2.45N$
3. For $1 \le N \le 4$, A_D full rank $\Rightarrow D \le 2N-1$:
   $Y_{2N}^{-2N}(\theta, \phi - \frac{\pi}{4}) \in \mathcal{Y}_{2N}$ is null on CS_N
   - $Y_{2N}^{-2N}(\theta, \phi - \frac{\pi}{4}) = \star \cdot \sin 2N(\phi - \frac{\pi}{4})$
   - $\mathrm{CS}_N \subset \{\boldsymbol{x}(\theta,\phi) : \phi \equiv \frac{\pi}{4}\,[\frac{\pi}{2N}]\}$

Statistics of the singular values of A_{2N-1}, $1 \leq N \leq 32$



- A_{2N-1} is numerically observed as full rank ($\min \sigma_i(A_{2N-1}) \gg 0$),
- $\mathrm{cond}(A_{2N-1}) \approx 1.19$
- (LS) is well-conditioned for $D = 2N - 1$

Statistics of the singular values of A_{2N}, $1 \leq N \leq 32$



- A_{2N} is not full rank for small values of N $(\min \sigma_i(A_{2N}) \approx 0)$
- Whenever A_{2N} is full rank, one observes numerically that $\mathrm{cond}(A_{2N}) > 10^4$
- (LS) is numerically ill-conditioned for $D = 2N$ $(\mathrm{cond}(A_{2N}^{\mathsf{T}} A_{2N}) > 10^8)$

Numerical evidence

- The maximal degree for (LS) to be well posed with a full rank VDM matrix is $D = 2N - 1$.
- In this case, (LS) is well-conditioned (condition number $\simeq 1$).
- Any $f \in \mathcal{Y}_{2N-1}$ is faithfully represented (sampled) on CS_N. The (LS) problem reconstructs exactly f from the restriction $f|_{\mathrm{CS}_N}$.
- The Cubed Sphere represents the equatorial line with a step-size $\delta = \frac{\pi}{2N}$.
- For trigonometric polynomials sampled with a step-size δ, the *Nyquist Shannon* angular frequency is defined as $\frac{\pi}{\delta} = 2N$.
- Main observation: the critical frequency for sampling along the equator coincides with the critical frequency for sampling on the Cubed Sphere!

Error evaluation in the least squares approximation of a function

- Function f evaluated on $\mathrm{CS}_N = \{\boldsymbol{x}_j,\ 1 \le j \le \bar{N}\}$ provides the data y_j

$$y_j = f(\boldsymbol{x}_j), \quad 1 \le j \le \bar{N}$$

- Compute $\tilde{f} \in \mathcal{Y}_{2N-1}$, solution to (LS) with degree $D = 2N-1$
- Relative ℓ^2-error on CS_M (for some large M)

$$\epsilon_M(f) := \left(\frac{\sum_{\boldsymbol{x} \in \mathrm{CS}_M} |f(\boldsymbol{x}) - \tilde{f}(\boldsymbol{x})|^2}{\sum_{\boldsymbol{x} \in \mathrm{CS}_M} |f(\boldsymbol{x})|^2} \right)^{1/2}$$

Test functions

- Very smooth: $f_1(x,y,z) = \exp(x)$
- Smooth

$$\begin{aligned} f_2(x,y,z) = & \tfrac{3}{4}\exp[-\tfrac{(9x-2)^2}{4} - \tfrac{(9y-2)^2}{4} - \tfrac{(9z-2)^2}{4}] \\ & + \tfrac{3}{4}\exp[-\tfrac{(9x+1)^2}{49} - \tfrac{9y+1}{10} - \tfrac{9z+1}{10}] \\ & + \tfrac{1}{2}\exp[-\tfrac{(9x-7)^2}{4} - \tfrac{(9y-3)^2}{4} - \tfrac{(9z-5)^2}{4}] \\ & - \tfrac{1}{5}\exp[-(9x-4)^2 - (9y-7)^2 - (9z-5)^2] \end{aligned}$$

- Spike: $f_3(x,y,z) = \frac{1}{10}\frac{\exp(x+2y+3z)}{(x^2+y^2+(z+1)^2)^{1/2}} 1(z > -1)$
- Not differentiable: $f_4(x,y,z) = \cos(3\arccos z) 1(3\arccos z \leq \frac{\pi}{2})$
- Discontinuous spherical cap: $f_5(x,y,z) = 1(z \geq \frac{1}{2})$

Fast convergence for smooth functions

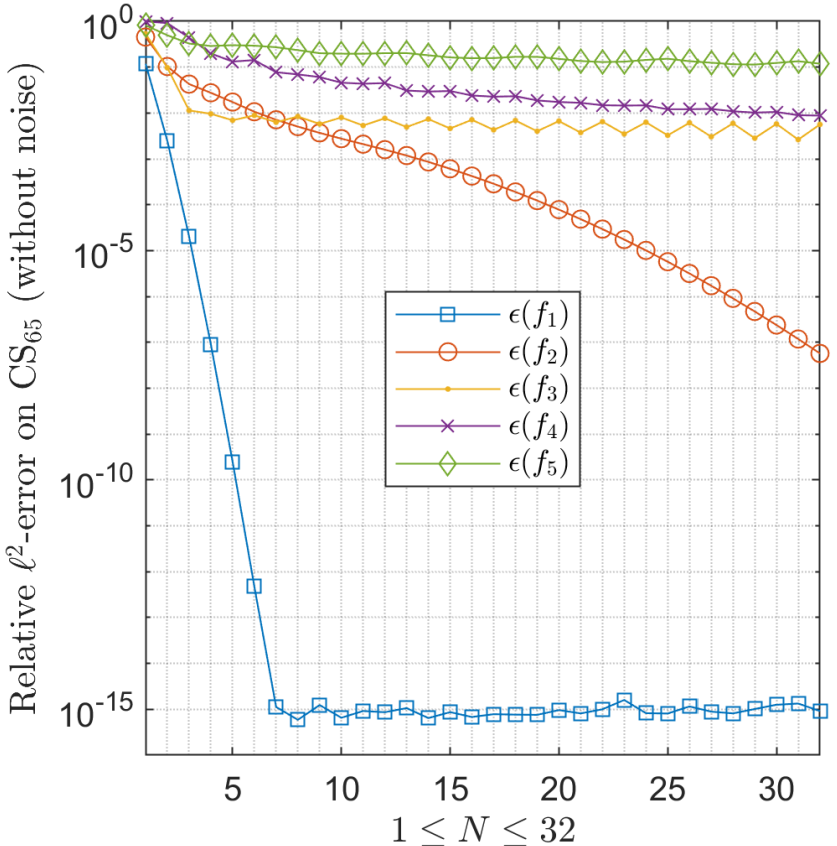


Figure: Relative l_2 Error in function of the resolution N

Stability (noisy data)

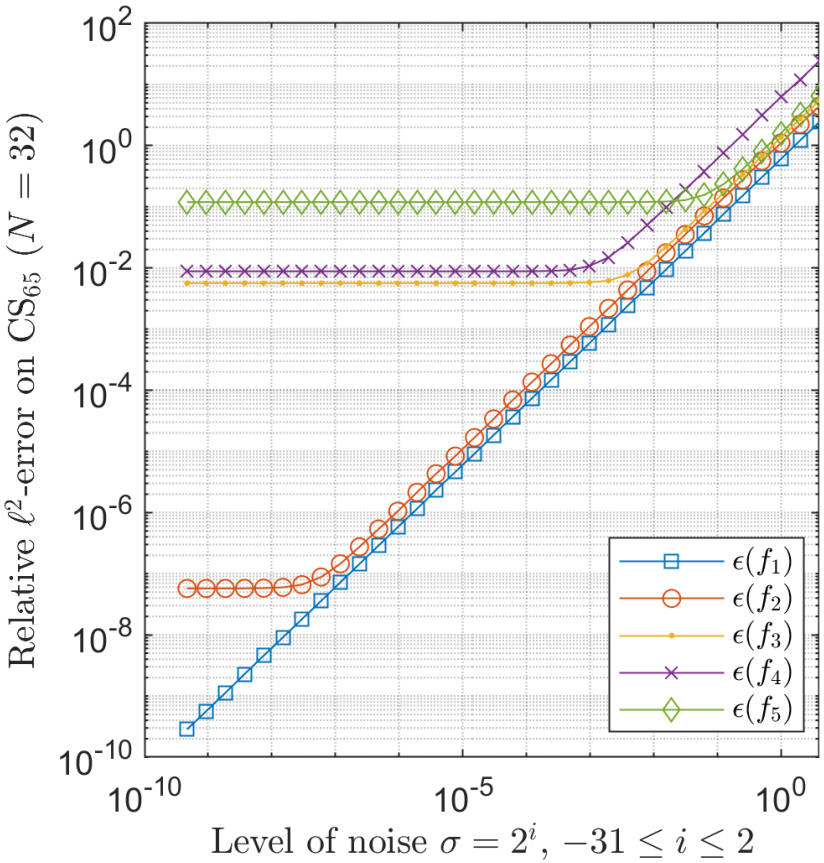


Figure: Evolution of the l_2 error in function of the noise magnitude $(N = 32)$

Noisy data

$y_j = f(\boldsymbol{x}_j) + \sigma\mathcal{N}(0,1),\ 1 \leq j \leq \bar{N}$

Numerical observation

$\epsilon(f)(\sigma) \approx \epsilon(f)(0) + \sigma$

Conclusion

Good stability !

Convergence for derivatives

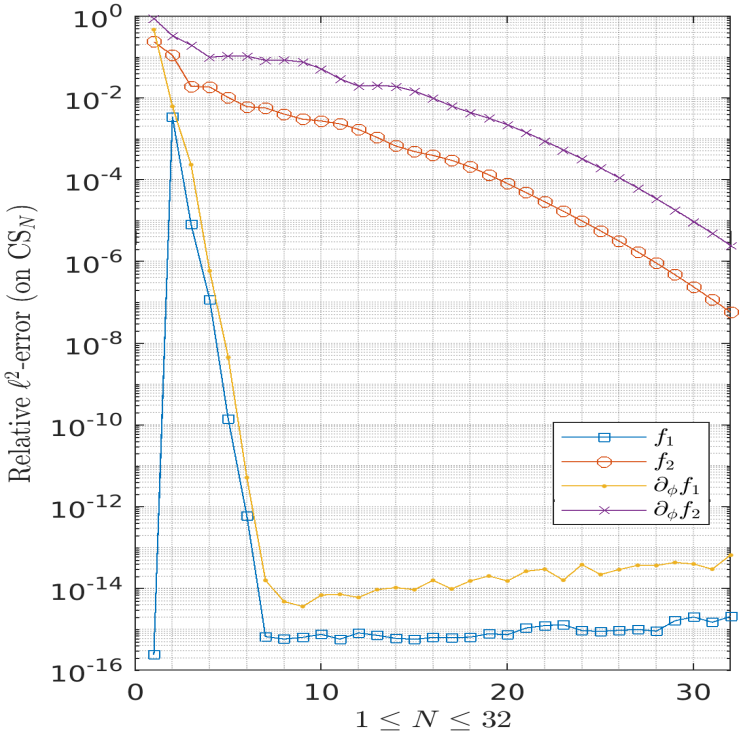


Figure: Functions f_1 and f_2: relative l_2 Error of the derivative in ϕ (longitude) in function of the resolution N.

Numerical observation

The error on the derivative behaves as the error on the function.

Interpolation problem

Data on the sphere

- Values $y_j \in \mathbb{R}$ given at the nodes $\boldsymbol{x}_j$, $1 \leq j \leq \bar{N}$ on the Cubed Sphere.

Problem

Find an integer $N' = N'(N)$ and a subspace $\mathcal{Y}'_{N'} \subset \mathcal{Y}_{N'}$, such that the problem

$$p(\boldsymbol{x}_j) = y_j, \forall 1 \leq j \leq \bar{N}.$$

with $p \in \mathcal{Y}'_{N'}$ has a unique solution (**unisolvence**) and is **well conditioned**.

Vandermonde matrix (collocation matrix)

- For k fixed, the rectangular matrix A_k is the VDM matrix associated to the basis $(Y_k^m)_{-k\leq m\leq k}$ of the SH space $\mathbb{Y}_k$, and to the nodes $\boldsymbol{x}_j \in \mathrm{CS}_N$,

$$A_k \triangleq [Y_k^m(\boldsymbol{x}_j)]_{1\leq j\leq \bar{N}, -k\leq m\leq k} \in \mathbb{R}^{\bar{N}\times(2k+1)}. \tag{26}$$

- For n fixed, the matrix $\boldsymbol{A}_n$ is the VDM matrix associated to the basis $(Y_k^m)_{|m|\leq k\leq n}$ of the SH space $\mathcal{Y}_n$,

$$\boldsymbol{A}_n \triangleq \begin{bmatrix} A_0 & \cdots & A_n \end{bmatrix} \in \mathbb{R}^{\bar{N}\times(n+1)^2}. \tag{27}$$

Rank increment

- For all $n \geq 0$, the rank of $\boldsymbol{A}_n$ is denoted by r_n and the rank increment between $\boldsymbol{A}_{n-1}$ and $\boldsymbol{A}_n$ is denoted by g_n:

$$\begin{cases} r_n \triangleq \operatorname{rank} \boldsymbol{A}_n, \ n \geq 0, & (a) \\ g_n \triangleq r_n - r_{n-1}, \ n \geq 0, & (b) \end{cases} \tag{28}$$

  (with $r_{-1} \triangleq 0, \quad g_0 \triangleq r_0$).

- How the sequence of rank increments g_n behaves with increasing n ?

Rank increment (cont.)

Facts:

- For n large enough, we have $\text{rank}(\boldsymbol{A}_n) = \bar{N}$.
- We call $N'(N)$ the smallest integer n such that $\boldsymbol{A}_n$ has full row rank $\bar{N}$.
- Numerical ansatz: $N'(N) = 3N$.

Echelon form and incremental algorithm

Main idea: perform an incremental Row Echelon factorization of the VDM matrix $\boldsymbol{A}_n$.

Theorem (Structure of $\boldsymbol{A}_n$)

Let $n \geq 0$.
The matrix $\boldsymbol{A}_n$ can be factorized in the form

$$\boldsymbol{A}_n = \boldsymbol{V}_n \boldsymbol{E}_n \boldsymbol{U}_n^{\mathsf{T}}, \tag{29}$$

where

- *the matrices $\boldsymbol{U}_n \in \mathbb{R}^{(n+1)^2 \times (n+1)^2}$, $\boldsymbol{V}_n \in \mathbb{R}^{\bar{N} \times \bar{N}}$ are orthogonal;*
- *the matrix $\boldsymbol{E}_n \in \mathbb{R}^{\bar{N} \times (n+1)^2}$ is in Row Echelon form.*

In particular, $\mathrm{rank}(\boldsymbol{E}_n) = \mathrm{rank}(\boldsymbol{A}_n) = r_n$.

Row Echelon Form of $\boldsymbol{A}_n$

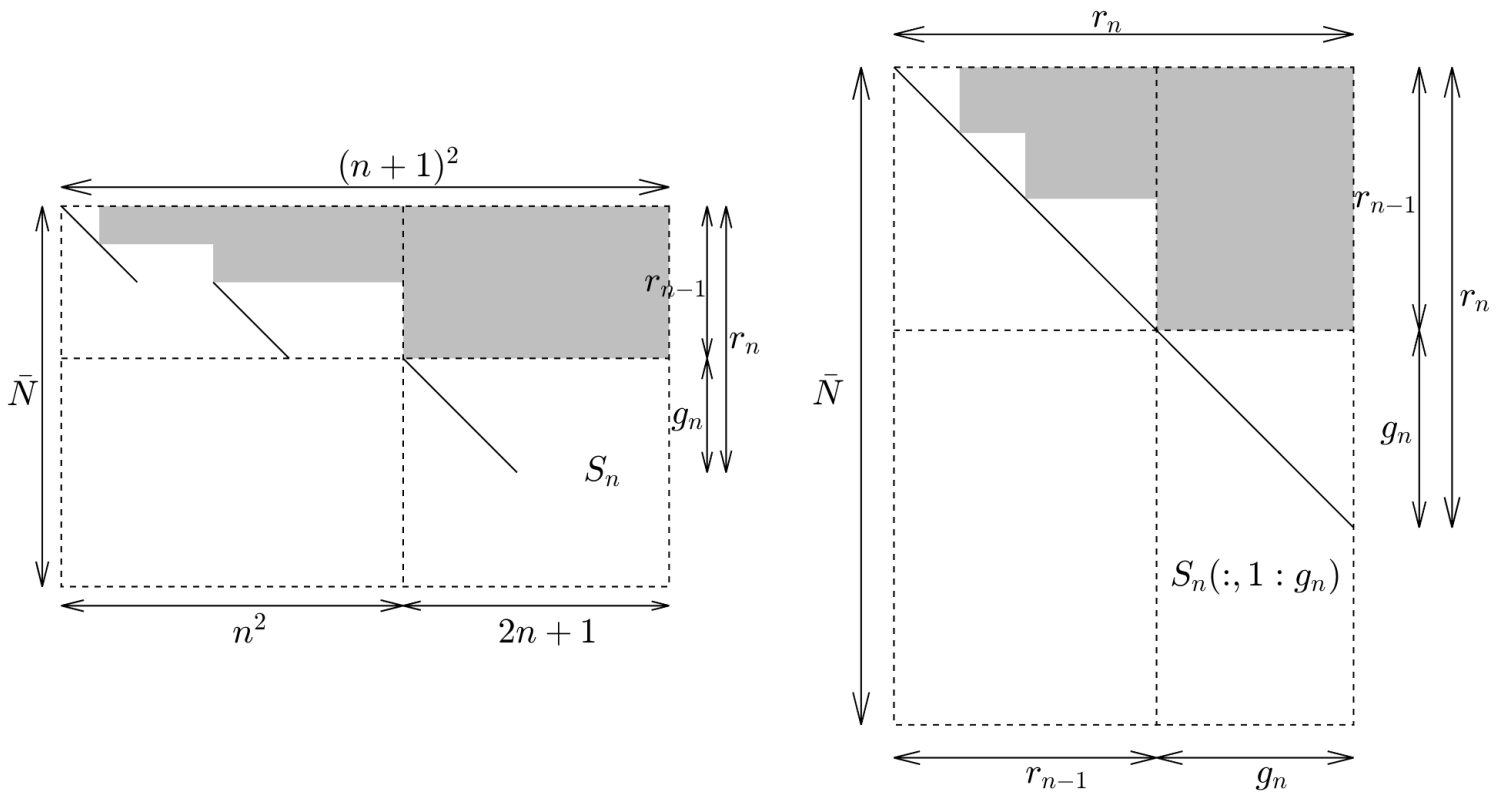


Figure: Left panel: the VDM $\boldsymbol{A}_n$ is equivalent to the row echelon matrix $\boldsymbol{E}_n$, whose shape is represented in gray. Right panel: elimination of redundant columns in $\boldsymbol{E}_n$ results in the upper triangular matrix $\boldsymbol{R}_n$, displayed in gray. (The white color represents zero entries.)

Row Echelon form (cont.)

- The algorithm consists in a sequence of SVD factorizations of matrices of size $\bar{N} \times (2k+1)$, $0 \leq k$,

$$V_n S_n U_n^\intercal \triangleq \boldsymbol{V}_{n-1}([\![1 : \bar{N}]\!], [\![r_{n-1}+1 : \bar{N}]\!])^\intercal A_n$$

- The matrices $\boldsymbol{V}_n$, $\boldsymbol{E}_n$, $\boldsymbol{U}_n$ are assembled iteratively.

Proposition (Orthonormal bases of ranges and null spaces)

(i) For every $n \geq 0$, the columns of $\boldsymbol{V}_n([\![1 : \bar{N}]\!], [\![1 : r_n]\!])$ are an orthonormal basis of $\operatorname{Ran} \boldsymbol{A}_n$.
(ii) For every $n \geq 0$, the columns of $\boldsymbol{V}_n([\![1 : \bar{N}]\!], [\![r_n+1 : \bar{N}]\!])$ are an orthonormal basis $\operatorname{Ker} \boldsymbol{A}_n^\intercal$.
(iii) For every $n \geq 1$, the columns of $U_n([\![1 : \bar{N}]\!], [\![1 : g_n]\!])$ are an orthonormal basis of $\operatorname{Ran} A_n^\intercal \boldsymbol{V}_{n-1}([\![1 : \bar{N}]\!], [\![r_{n-1}+1 : \bar{N}]\!])$. *(iv) For every $n \geq 1$, the columns of $U_n([\![1 : \bar{N}]\!], [\![g_n+1 : 2n+1]\!])$, are an orthonormal basis of* $\operatorname{Ker} \boldsymbol{V}_{n-1}([\![1 : \bar{N}]\!], [\![r_{n-1}+1 : \bar{N}]\!])^\intercal A_n$.

- Algorithm exits when $r_n = \bar{N}$.

Structure of the interpolation space

- For each n, one has a decomposition of $\mathbb{Y}_n$ as

$$\mathbb{Y}_n = \mathbb{Y}'_n \oplus \mathbb{Y}''_n$$

with $\mathbb{Y}'_n =$ the SH of degree n not already represented by the space $\mathcal{U} = \oplus_{k=0}^{n-1} \mathbb{Y}'_k$.

- A relevant SH subspace given by the algorithm is

$$\mathcal{U}_{3N} = \oplus_{k=0}^{3N} \mathbb{Y}'_k$$

Poisson problem (continuous level)

- for $g : \boldsymbol{x} \in \mathbb{S}^2 \mapsto g(\boldsymbol{x})$. Consider the null mean Poisson equation on the sphere:

$$\begin{cases} \Delta u = g \text{ on } \mathbb{S}^2, \\ \int_{\mathbb{S}^2} u \mathrm{d}\sigma = 0. \end{cases}$$

- Expansion in SH

$$g = \sum_{n \geq 0} \sum_{|m| \leq n} g_{n,m} Y_n^m. \tag{30}$$

- Using

$$\Delta Y_n^m = -n(n+1) Y_n^m, \tag{31}$$

the solution is expressed as

$$g = -\sum_{n \geq 1} \sum_{|m| \leq n} \frac{g_{n,m}}{n(n+1)} Y_n^m. \tag{32}$$

- Approximation: do the same in the finite dimensional space $\mathcal{U}_{3N}$.

Approximate Poisson problem in $\mathcal{U}_{3N}$

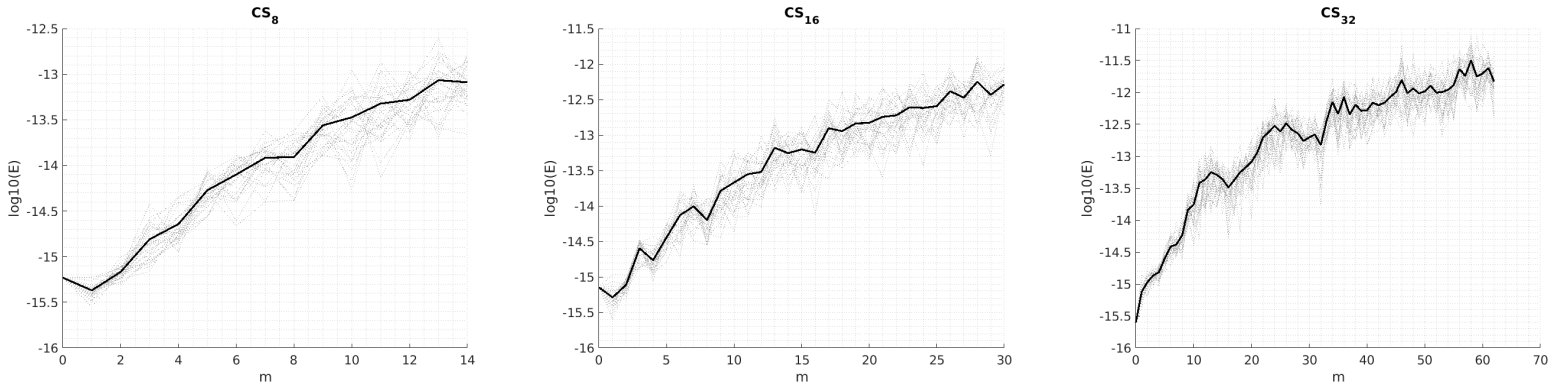


Figure: Poisson equation solver error on CS_N for $N \in \{8, 16, 32\}$. The relative error is plotted related to the value m for 30 random values e_m and d_m in $[0, 2\pi]$. Each gray dotted line corresponds to a random value. The black line is the mean error.

Quadrature on the sphere

Theorem (Quadrature rule)

Let $u : \mathbb{S}^2 \to \mathrm{R}$ be a given function. The quadrature rule $\mathcal{Q}_N$ is defined by

$$\mathcal{Q}_N u := \int_{\mathbb{S}^2} \mathcal{I}_N[u|_{\mathrm{CS}_N}](x) d\sigma.$$

(i) The formula $\mathcal{Q}_N$ can be expressed as follows:

$$\mathcal{Q}_N u = \sum_{j=1}^{\bar{N}} \omega_N(x_j) u(x_j), \tag{33}$$

where the weight function $\omega_N \in \mathcal{F}(CS_N)$ is defined by

$$[\omega_N(x_j)] = (\boldsymbol{A}^{\mathsf{T}})^{-1} [\sqrt{4\pi}\ 0\ \cdots\ 0]^{\mathsf{T}}, \tag{34}$$

with $\boldsymbol{A}$ the Vandermonde matrix.
(ii) The formula $\mathcal{Q}_N$ is exact on the space $\mathcal{U}_{3N}$,

$$\forall u \in \mathcal{U}_{3N}, \quad \mathcal{Q}_N u = \int_{\mathbb{S}^2} u(x) d\sigma.$$

In addition, the rule $\mathcal{Q}_N$ and the weight function ω_N are invariant under the group $\mathcal{G}$ of CS_N.

Quadrature on the sphere

i	$f_i(x,y,z)$	$\int_{\mathbb{S}^2} f_i(x,y,z)d\sigma$	Ref.
1	$\exp(x)$	$14.7680137457653\cdots$	[?, ?]
2	$\frac{3}{4}\exp[-\frac{(9x-2)^2}{4}-\frac{(9y-2)^2}{4}-\frac{(9z-2)^2}{4}]$ $+\frac{3}{4}\exp[-\frac{(9x+1)^2}{49}-\frac{9y+1}{10}-\frac{9z+1}{10}]$ $+\frac{1}{2}\exp[-\frac{(9x-7)^2}{4}-\frac{(9y-3)^2}{4}-\frac{(9z-5)^2}{4}]$ $-\frac{1}{5}\exp[-(9x-4)^2-(9y-7)^2-(9z-5)^2]$	$6.6961822200736179523\cdots$	[?, ?, ?, ?, ?]
3	$\frac{1}{10}\frac{\exp(x+2y+3z)}{(x^2+y^2+(z+1)^2)^{1/2}}1(z>-1)$	$4.090220018862976\cdots$	[?]
4	$\cos(3\arccos z)1(3\arccos z\leq\frac{\pi}{2})$	$\frac{\pi}{8}$	inspired from [?]
5	$1(z\geq\frac{1}{2})$	π	
6	$\frac{1}{9}[1+\mathrm{sign}(-9x-9y+9z)]$	$\frac{4\pi}{9}$	[?, ?, ?, ?]

Table: Several test functions and exact mean values.

References

1. M. Brachet, JPC, Spherical Shallow Water simulation by a cubed sphere finite difference solver, *Quart. Jour. Roy. Met. Soc.*, 147 (735), 2021, 786–800.
2. M. Ben-Artzi, JPC, D. Fishelov, A high order finite difference scheme for high order quasi-geostrophic equations, in preparation, 2023.
3. J.-B. Bellet: Symmetry group of the equiangular cubed sphere, *Quart. App. Math., 80, 114142, 2022*
4. J.-B. Bellet, JPC: Least squares spherical harmonics approximation on the cubed sphere, *J. Comp. App. Math, 429, 115213, 2023*

Summary and questions

- CS is a set of nodes on the sphere with a good "approximation power". In which cases is the icosahedron grid proved to be better ?
- Cartesian local structure, great circles coordinate lines make CS an interesting mathematical object. Finite differencing along these circles is effectively high order. No significant interpanel problems. Computational efficiency ?
- CS supports a natural discrete harmonic analysis setup. Full mathematical analysis remains to be done. Fast global solver on the CS ?
- Munk-Stommel equations with the problems of oceanic gyres and western boundary layers are very attractive. What is the main message concerning this problem for theoretical and numerical analysis coming from the oceanographic community