Parallel geometric multigrid solver at the reduced latitude-longitude grid

Gordey Goyman, Mikhail Tolstykh

Marchuk Institute of Numerical Mathematics RAS (INM RAS)

Hydrometeorological Center of Russia (HMCR)

PDEs on the sphere 2019

Outline

- Motivation
- Solver description
- Numerical experiments

Global semi-Lagrangian atmosphere model SL-AV

Global semi-Lagrangian atmosphere model SL-AV

Current version: SL-AV20*

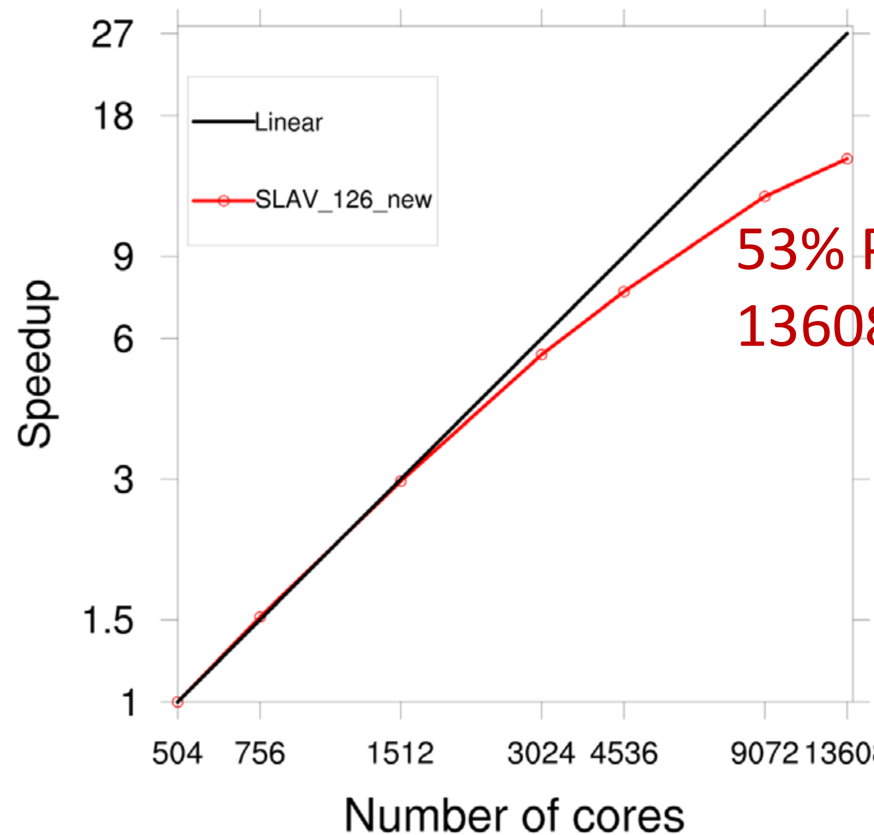
- SISL time-stepping
- **Regular** lat-lon grid, variable latitude resolution
- Vor-div formulation Z-grid
- **Fourier space in longitude**, FD in latitude
- Implicit 4th order hyper-diffusion
- **1d MPI decomposition**, OpenMP

---

*Tolstykh M. et al. Vorticity-divergence semi-Lagrangian global atmospheric model SL-AV20: dynamical core //Geoscientific Model Development. – 2017. – T. 10. – №. 5. – C. 1961-1983.

SL-AV20 strong scaling

- SL-AV20 model with the grid resolution 3024x1513x126
- Cray XC40-LC supercomputer



53% PE at 13608 cores

Global semi-Lagrangian atmosphere model SL-AV

Current version: SL-AV20*

- SISL time-stepping
- **Regular** lat-lon grid, variable latitude resolution
- Vor-div formulation Z-grid
- **Fourier space in longitude**, FD in latitude
- Implicit 4th order hyper-diffusion
- **1d MPI decomposition**, OpenMP

New version: SL-AV10

- SISL time-stepping
- **Reduced** lat-lon grid, variable latitude resolution
- Vor-div formulation Z-grid / **U-V + C-grid**?
- **Fully grid-point**
- Implicit 4th order hyper-diffusion?
- **2d MPI decomposition**, OpenMP?

---

*Tolstykh M. et al. Vorticity-divergence semi-Lagrangian global atmospheric model SL-AV20: dynamical core //Geoscientific Model Development. – 2017. – T. 10. – №. 5. – C. 1961-1983.

SL-AV model: elliptic problems

SISL time-stepping	$(\omega^2 - \Delta)G = R$
Vor-Div formulation	$\begin{cases} \Delta\psi = \xi \\ \Delta\chi = D \end{cases}$
Implicit hyper-diffusion	$\psi_f = \psi - \nabla \cdot \left(\boldsymbol{K}\nabla^3\psi_f\right)$

Set of 2d elliptic equations at each vertical level of the model

SL-AV20 elliptic solver:

- FFT in longitude direction ⇒ set of 1d independent problems
- Global forward data transposition
- Block Thomas algorithm
- Global backward data transposition

SL-AV10 elliptic solver:

?

Fast

Robust

Scalable

Multigrid solvers

Basic idea:

- Eliminate high frequency part of the error (smoothing operator)
- Move to the coarse grid (restriction)
- Repeat recursively
- Interpolate error correction back to the fine grid (prolongation)

proven efficiency in various applications, including NWP*

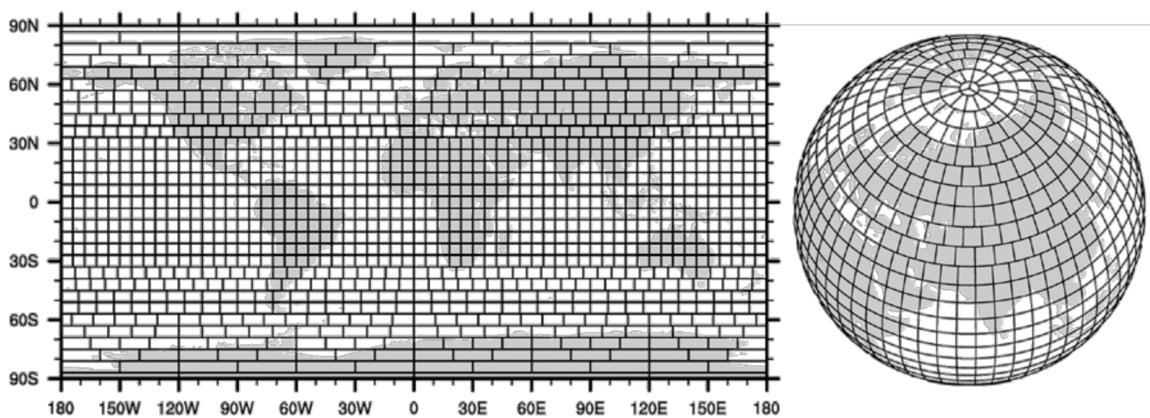
BUT!

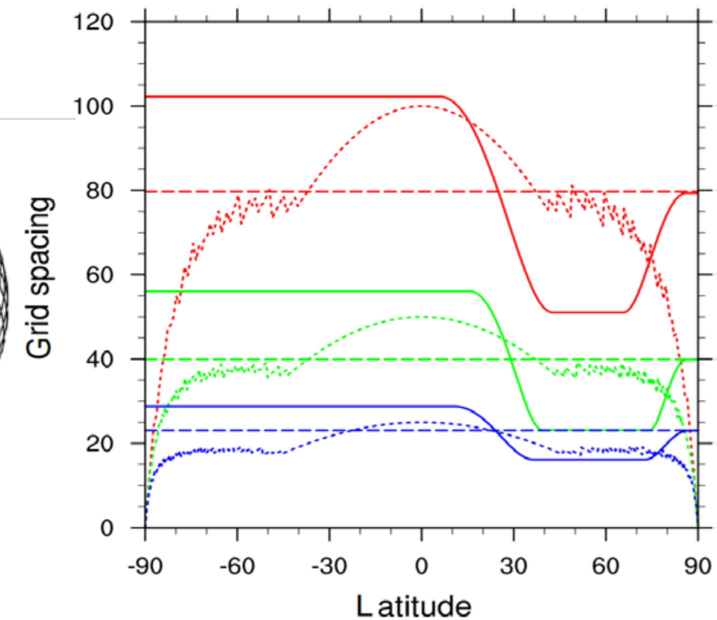
Special care is needed in the presence of anisotropy in the equation coefficients

*Müller E. H., Scheichl R. Massively parallel solvers for elliptic partial differential equations in numerical weather and climate prediction //Quarterly Journal of the Royal Meteorological Society. – 2014.

Reduced grid

- Points located at the latitudes φ_j with longitude spacing $\Delta\lambda_j$
- First and last grid latitudes half-step shifted from the poles
- Anisotropy ratio relaxed, but still noticeable





Multigrid solvers. Anisotropic case

- Pointwise technics (Jacobi-like methods) failed to provide error smoothing in both directions

- 2 ways to overcome problems with anisotropy
  - Special smoothing technics
  - Special coarse-grid construction technics

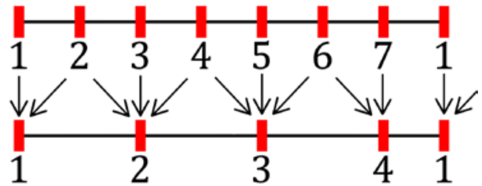
Conditional semi-coarsening

- Firstly mentioned in (Larsson 2008)
- Regular lat-lon grid solver presented in (Buckeridge 2010)
- Closely related to Algebraic MG concept
- Always coarsen in lon-direction, coarsen in lat-direction only in the areas with low anisotropy ratio

Longitude coarsening

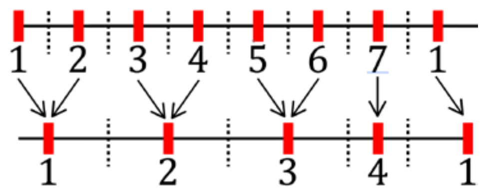
- **Vertex-centered strategy**

coarse points are the subset of the fine grid points



- **Cell-centered strategy**

coarse points are the centers of the joint fine grid cells



Non-uniform grid spacing could be introduced

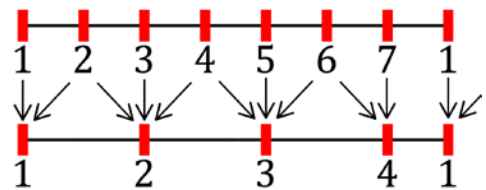
Complicates the structure of the code

Requires equation discretization for the non uniform case

Longitude coarsening

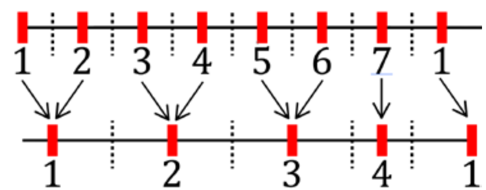
- **Vertex-centered strategy**

coarse points are the subset of the fine grid points



- **Cell-centered strategy**

coarse points are the centers of the joint fine grid cells



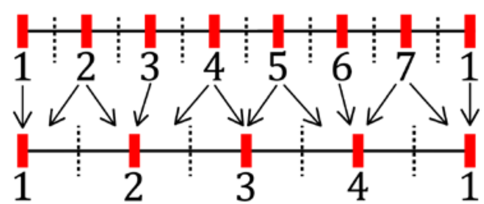
Non-uniform grid spacing could be introduced

Complicates the structure of the code

Requires equation discretization for the non uniform case

- **Uniform coarsening strategy**

coarse points distributed uniformly



Uniform spacing

Coarse grids have the same structure as the initial one

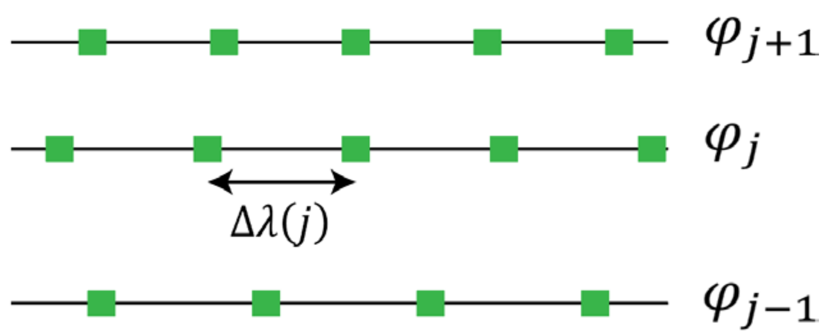
Latitude coarsening

Grid anisotropy ratio is function of latitude only

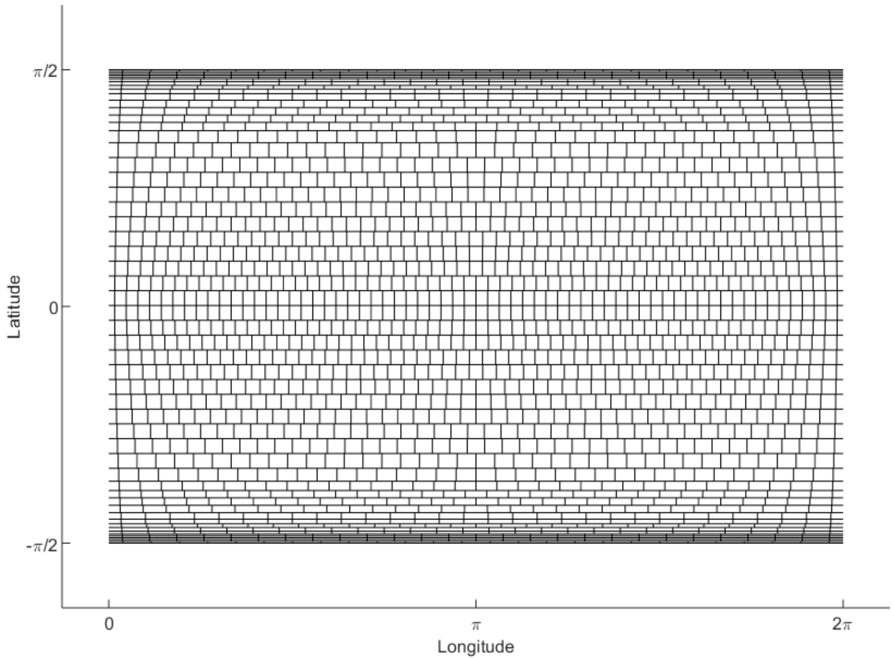
$$\alpha(\varphi_j) = \left(\frac{2\Delta\lambda(j)\cos\varphi_j}{\varphi_{j+1} - \varphi_{j-1}}\right)^2$$

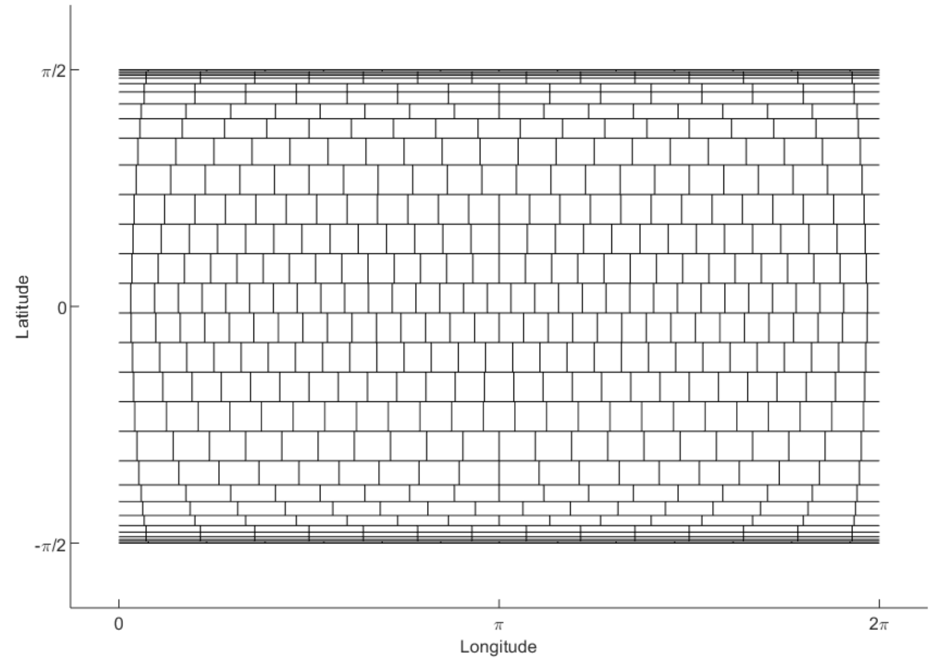
Pointwise smoother effectively dump high frequency errors in both directions at latitudes with

$$\alpha(\varphi_j) \geq 0.5$$



Coarse grids construction example





MG components. Intergrid operators

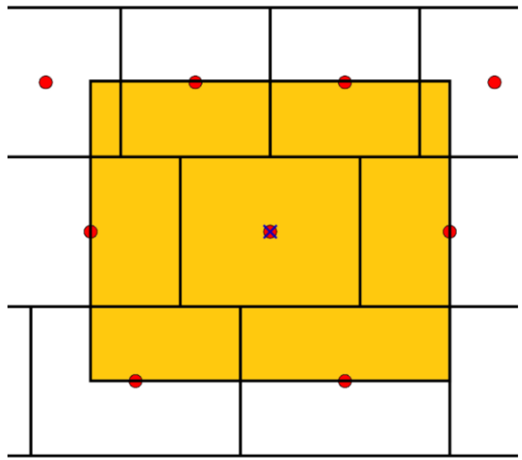
- **Restriction** – volume-weighted average:

$$\psi_i^c = \frac{1}{V(\Omega_i^c)} \sum_k \psi_k^f \, V\left(\Omega_i^c \cap \Omega_k^f\right)$$

- **Prolongation** – bilinear interpolation

- Other choices are possible

$$(Px, y)_{\Omega_f} = (x, Ry)_{\Omega_c}$$



MG components. Smoothing operator

- Smoother must effectively eliminate high frequency error modes

- All standard approaches are applicable (Jacobi, Gauss-Seidel, SPAI, block ILU(0), …)

- Our choice – hybrid Gauss-Seidel with red-black ordering

MG components. Coarse grid solver

- Common practice is to use several iterations of the smoother (assuming coarse problem well-conditioned)

- Laplacian discretization matrix spectrum:

$$\Delta\psi \approx L\psi, \;\; \mu_{min}(L) = O(1), \;\; \mu_{max}(L) \sim \max_j \frac{1}{\Delta\lambda_j^2 \Delta\varphi_j^2}$$

- Pointwise smoothing is not effective coarse grid solver

- Several iterations (3-8) of BICGSTAB/CG method is used

Algorithm convergence analysis

- Given the problem $Ax = b$
- rhs – known solution (random vector) multiplied by the system matrix
- Looking for the average number of iterations N to reach relative error tolerance $\frac{\|\xi^N\|}{\|x^*\|} \leq 10^{-7}$
- 2 iterations of pre- and post-smoothing

Grids setup

- Grids with initial resolution:

  512x256, 1024x512, 2048x1024, 4096x2048

- Linear reduction towards poles from N_λ to αN_λ, $\alpha \in \{1, 0.1, 0.01\}$

Initial grid	α	Dof	N_λ at poles
512x256	1	$1.3 \cdot 10^5$	512
	0.1	$7.2 \cdot 10^4$	54
	0.01	$6.6 \cdot 10^4$	10
1024x512	1	$5.2 \cdot 10^5$	1024
	0.1	$2.8 \cdot 10^5$	106
	0.01	$2.6 \cdot 10^5$	14

Initial grid	α	Dof	N_λ at poles
2048x1024	1	$2.1 \cdot 10^6$	512
	0.1	$1.1 \cdot 10^6$	208
	0.01	$1.0 \cdot 10^6$	24
4096x2048	1	$8.3 \cdot 10^6$	4096
	0.1	$4.6 \cdot 10^6$	414
	0.01	$4.2 \cdot 10^6$	44

Algorithm convergence. Poisson problem

$$\Delta G = R$$

Initial grid	α	Coarse grids	N_{iter}
512x256	1	6	**6**
	0.1	5	**7**
	0.01	5	**7**

Initial grid	α	Coarse grids	N_{iter}
1024x512	1	7	**6**
	0.1	6	**7**
	0.01	6	**8**

Initial grid	α	Coarse grids	N_{iter}
2048x1024	1	8	**6**
	0.1	7	**7**
	0.01	7	**7**

Initial grid	α	Coarse grids	N_{iter}
4096x2048	1	9	**6**
	0.1	8	**7**
	0.01	8	**8**

Algorithm convergence. Helmholtz problem

$$(\omega^2 - \Delta)G = R$$

$\omega^2 = \{300, 1200, 4800, 19200\}$ (scaling proportional to $(\Delta t)^{-2}$)

Initial grid	α	Coarse grids	N_{iter}	Coarsest level Dof
512x256	1	6	**7**	120
	0.1	3	**6**	1200
	0.01	2	**6**	4185

Initial grid	α	Coarse grids	N_{iter}	Coarsest level Dof
1024x512	1	7	**7**	136
	0.1	4	**7**	1205
	0.01	2	**6**	16654

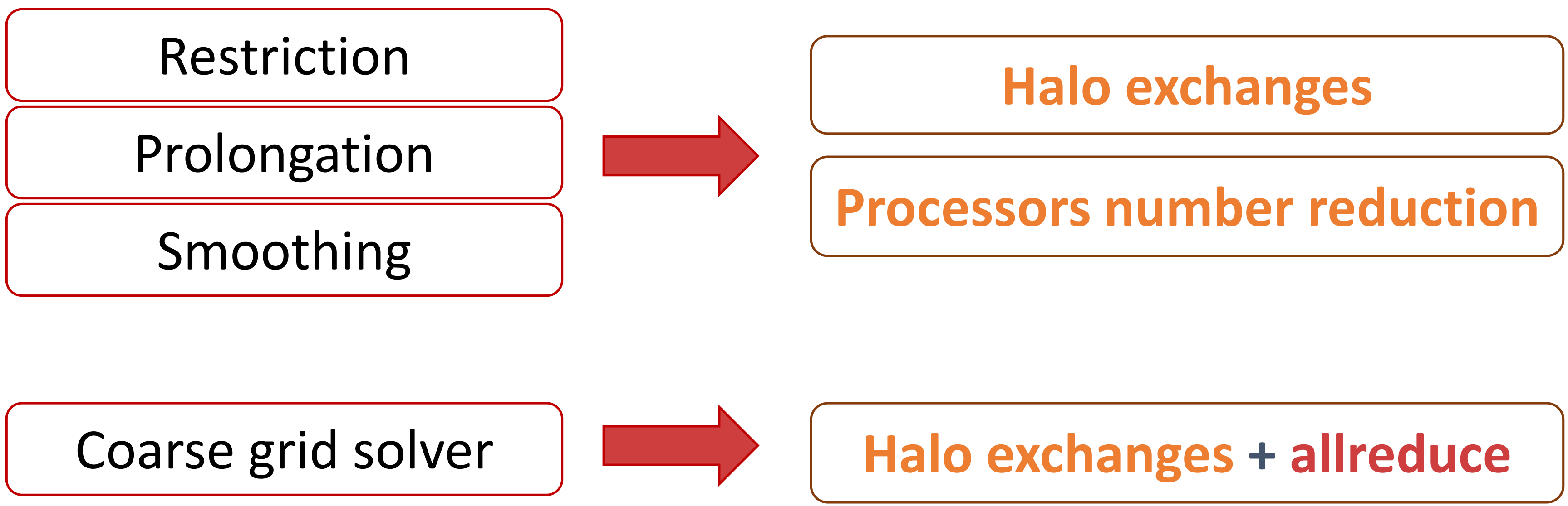
Initial grid	α	Coarse grids	N_{iter}	Coarsest level Dof
2048x1024	1	8	**6**	144
	0.1	5	**7**	1212
	0.01	2	**7**	66413

Initial grid	α	Coarse grids	N_{iter}	Coarsest level Dof
4096x2048	1	9	**6**	160
	0.1	6	**6**	1212
	0.01	3	**6**	66378

Implicit hyper-diffusion equation

- $\psi_f = \psi - \nabla \cdot \left(\boldsymbol{K}\nabla^3\psi_f\right) \Rightarrow \begin{cases} \psi_f = \psi - \nabla \cdot (\boldsymbol{K}\nabla z) \\ z = \nabla^2\psi_f \end{cases}$
- Smoothing operator – collective relaxation
- All other components remain the same
- Convergence results quite similar to the Helmholtz problem case

Parallel implementation



- Pure MPI implementation
- Scaling tests – yet to be done

Summary

- Geometric multigrid algorithm at the reduced lat-lon grid is developed
- Algorithm is robust with respect to the problem size and grid reduction ratio
- The use of sufficient grid reduction allows using fewer coarse levels (in case of Helmholtz and implicit hyper-diffusion problems), which is likely to have a positive effect on the algorithm parallel efficiency
- Algorithm could be extended to the 3d case
- Scalability tests and test within the SL-AV model framework have to be performed

Reduced grid discretization

$$(\nabla^2\psi)_{i,j} = \frac{1}{V_j}\left(dS_j^{\varphi}\frac{\psi_{i+1,j} - 2\psi_{i,j} + \psi_{i-1,j}}{\Delta\lambda_j \cos\varphi_j} + dS^{\lambda}_{j+1/2}\frac{\widetilde{\psi}_{i,j+1} - \psi_{i,j}}{\varphi_{j+1} - \varphi_j} - dS^{\lambda}_{j-1/2}\frac{\psi_{i,j} - \widetilde{\psi}_{i,j-1}}{\varphi_j - \varphi_{j-1}}\right),$$

$$dS_j^{\varphi} = \frac{\varphi_{j+1} - \varphi_{j-1}}{2}, \quad dS^{\lambda}_{j\pm 1/2} = \Delta\lambda_j \cos\frac{\varphi_{j\pm1} + \varphi_j}{2}, \quad V_j = a^2\frac{dS^{\lambda}_{j+1/2} + dS^{\lambda}_{j-1/2}}{2} dS_j^{\varphi},$$

$\widetilde{\psi}_{i,j\pm1}$ is the value interpolated to the point $((i-1)\Delta\lambda_j, \varphi_{j\pm1})$

Coarse grid matrix condition number

- Coarse grid with resolution 4x8