ALADIN/HIRLAM Dynamics Day
28 May 2019, Toulouse

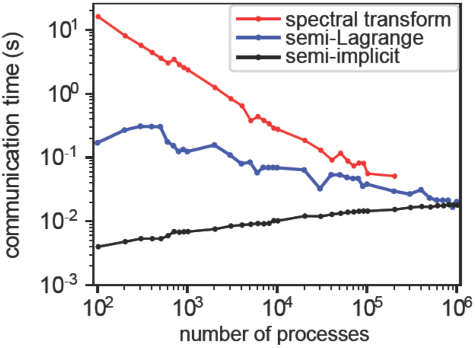
A non-spectral solver for the ALADIN-NH dynamics

Daan Degrauwe, RMI Belgium

Content

- Introduction
- The semi-implicit system under consideration
- A non-spectral solver
- Scalability
- Conclusions

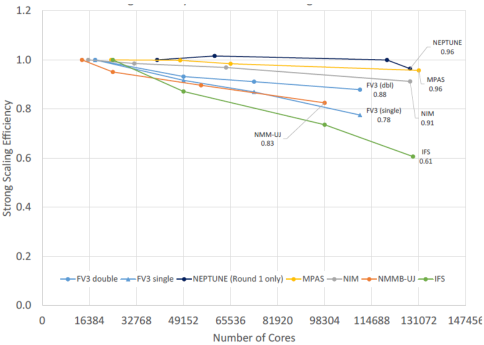
- Spectral transforms require data-rich global communications
- As long as bandwidth is maintained throughout the HPC, scalability should be really good



Zheng & Marguinaud (2018)

Scalability of the spectral transforms

- Spectral transforms require data-rich global communications
- As long as bandwidth is maintained throughout the HPC, scalability should be really good
- However, practical tests show far-from perfect scalability:



Michalakes et al. (2015)

The semi-implicit system

- (Simplifying a bit,) the semi-implicit timestepping in ALADIN-NH requires to solve the following problem:

$$
\begin{aligned}
D - \delta t R_a T^* \nabla^2 \left[(\mathbf{G}^* - 1)\hat{q} - \mathbf{G}^*(T/T^*) - (\pi_s'/\pi_s^*) \right] &= \tilde{D} \\
d - \delta t \left(-\frac{g^2}{R_a T_e^*} \mathbf{L}^* \hat{q} \right) &= \tilde{d} \\
T' + \delta t \frac{RT^*}{C_{va}} (D + d) &= \tilde{T} \\
\hat{q} - \delta t \left[\mathbf{S}^* D - \frac{C_{pa}}{C_{va}} (D + d) \right] &= \tilde{\hat{q}} \\
\pi_s' + \delta t \pi_s^* \mathbf{N}^* D &= \tilde{\pi}_s
\end{aligned}
$$

- Whether this system is solved using spectral transforms or not doesn't affect the rest of the ALADIN/HIRLAM model!

- Since all coefficients and operators are constant in space and time, this system can be reduced to a single 3D Helmholtz problem in D:

$$
\left(\mathbf{I} - \delta t^2 \nabla^2 \mathbf{B}_D^* \right) D = D^{\bullet\bullet}
$$

- The existing spectral solver takes the following steps:

  1. spectral transforms of prognostic variables D, d, T', $\hat{q}$, π'_s
  2. calculation of RHS term $D^{\bullet\bullet}$ of the Helmholtz equation in spectral space
  3. projection on eigenvectors of $\mathbf{B}^*_D$
  4. solution of 2D Helmholtz equation for each vertical eigenmode in spectral space

$$(1 - c_\ell^2 \delta t^2 \nabla^2)\psi_\ell = RHS_\ell$$

  5. inverse eigenmode projection to get D
  6. back-substitution to get d, T', $\hat{q}$, π'_s
  7. inverse spectral transforms

- A non-spectral solver takes the following steps:

1. ~~spectral transforms of prognostic variables D, d, T', $\hat{q}$, π'_s~~
2. calculation of RHS term $D^{\bullet\bullet}$ of the Helmholtz equation in ~~spectral~~ gridpoint space
3. projection on eigenvectors of $\mathbf{B}^*_D$
4. solution of 2D Helmholtz equation for each vertical eigenmode in ~~spectral~~ gridpoint space

$$(1 - c_\ell^2 \delta t^2 \nabla^2)\psi_\ell = RHS_\ell$$

5. inverse eigenmode projection to get D
6. back-substitution to get d, T', $\hat{q}$, π'_s
7. ~~inverse spectral transforms~~

- So, in order to provide a non-spectral alternative for the spectral Helmholtz solver in ALADIN/HIRLAM, all we need is a solver for a 2D Helmholtz problem.
- This is a pretty common numerical problem, which is often solved by iterative solvers.
- What makes our application special are the *tight operational constraints*: we can't risk to have delayed forecasts due to unpredictable convergence speed.

A non-spectral solver

- So, in order to provide a non-spectral alternative for the spectral Helmholtz solver in ALADIN/HIRLAM, all we need is a solver for a 2D Helmholtz problem.
- This is a pretty common numerical problem, which is often solved by iterative solvers.
- What makes our application special are the *tight operational constraints*: we can't risk to have delayed forecasts due to unpredictable convergence speed.
- However, for our constant-coefficient problem and a given solver/preconditioner, one can show that the *convergence speed is guaranteed*! Moreover, it can be determined (semi-)analytically.

Convergence speed of iterative methods

- The Richardson solver solves a general linear problem $\mathbf{Ax} = \mathbf{b}$ as follows:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{r}^{(k)}$$

with $r^{(k)} = \mathbf{b} - \mathbf{Ax}^{(k)}$ the residual vector.

- One can easily show that

$$\mathbf{r}^{(k+1)} = (\mathbf{I} - \mathbf{A})\mathbf{r}^{(k)}$$

so the convergence is determined by the maximum absolute eigenvalue $|\lambda|_{max}$ of $\mathbf{I} - \mathbf{A}$.

Convergence speed of iterative methods

- The Richardson solver solves a general linear problem $\mathbf{Ax} = \mathbf{b}$ as follows:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{r}^{(k)}$$

with $r^{(k)} = \mathbf{b} - \mathbf{Ax}^{(k)}$ the residual vector.

- One can easily show that

$$\mathbf{r}^{(k+1)} = (\mathbf{I} - \mathbf{A})\mathbf{r}^{(k)}$$

so the convergence is determined by the maximum absolute eigenvalue $|\lambda|_{max}$ of $\mathbf{I} - \mathbf{A}$.

- When using a preconditioner to transform the problem into $\mathbf{P}^{-1}\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{b}$, the convergence speed is determined by the maximum absolute eigenvalue of $\mathbf{I} - \mathbf{AP}^{-1}$.

Convergence speed of iterative methods

- The Richardson solver solves a general linear problem $\mathbf{Ax} = \mathbf{b}$ as follows:

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{r}^{(k)}$$

  with $r^{(k)} = \mathbf{b} - \mathbf{Ax}^{(k)}$ the residual vector.

- One can easily show that

$$\mathbf{r}^{(k+1)} = (\mathbf{I} - \mathbf{A})\mathbf{r}^{(k)}$$

  so the convergence is determined by the maximum absolute eigenvalue $|\lambda|_{max}$ of $\mathbf{I} - \mathbf{A}$.

- When using a preconditioner to transform the problem into $\mathbf{P}^{-1}\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{b}$, the convergence speed is determined by the maximum absolute eigenvalue of $\mathbf{I} - \mathbf{AP}^{-1}$.

- For Krylov methods, the convergence speed is determined by the spectral radius $\lambda_{max}/\lambda_{min}$.

Advantage of a constant-coefficient system

- For our constant-coefficient Helmholtz problem, the matrix $\mathbf{A} = 1 - c_\ell^2 \delta t^2 \nabla^2$ does not depend on the weather situation! So its eigenvalues (and the convergence speed of an iterative solver) are predictable!

- The Helmholtz problem is entirely determined by a single parameter: the wave Courant number $c_\ell \delta t / \Delta x$.

Advantage of a constant-coefficient system

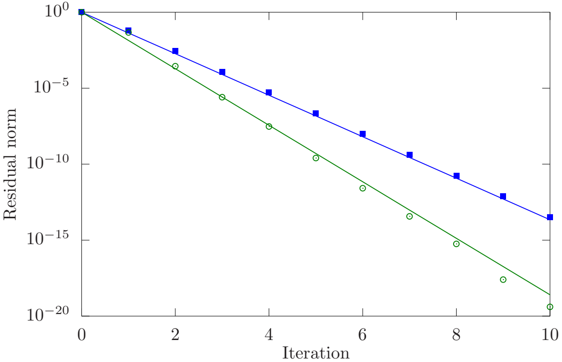
- For our constant-coefficient Helmholtz problem, the matrix $\mathbf{A} = 1 - c_\ell^2 \delta t^2 \nabla^2$ does not depend on the weather situation! So its eigenvalues (and the convergence speed of an iterative solver) are predictable!

- The Helmholtz problem is entirely determined by a single parameter: the wave Courant number $c_\ell \delta t / \Delta x$.

- For a **LAM geometry** and a multigrid preconditioner, the eigenvalues of $\mathbf{AP}^{-1}$ do not even depend on the grid dimensions, and can be determined semi-analytically with a low-dimensional Rayleigh-Ritz method

Predicted convergence speed

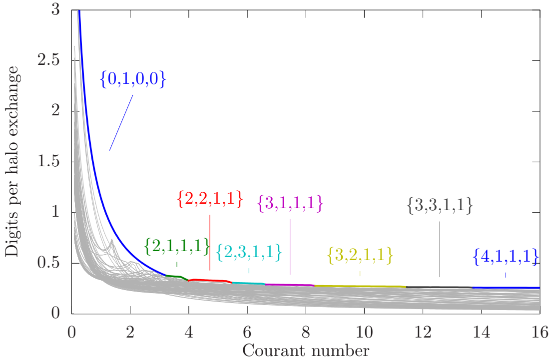
- Comparison of predicted and measured convergence speed:



$$\mu_\ell = 1.44,\ d = 1,\ \nu_0 = 2,\ \nu_1 = 1,\ \nu_2 = 1$$

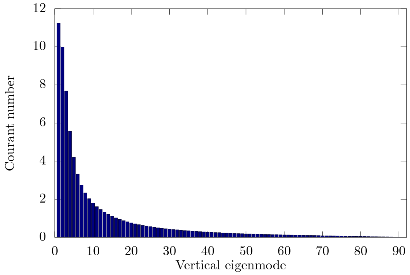
Choice of preconditioner

- The choice of a preconditioner is commonly regarded to be difficult task
- Thanks to the predictable convergence rates, it is possible to pick the optimal preconditioner parameters



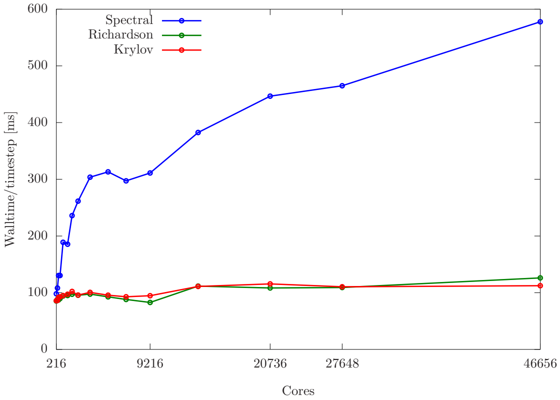
Choice of preconditioner

- The choice of a preconditioner is commonly regarded to be difficult task
- Thanks to the predictable convergence rates, it is possible to pick the optimal preconditioner parameters
- It's even possible to use optimal parameters for each vertical eigenmode separately.



This allows for a reduction in communication volume by a factor 5!

- Weak scalability tests on ECMWF's Cray:



Important note: only scalability of Helmholtz solver!

Conclusions and future work

- Scalability of the spectral transforms seems problematic, albeit not in the immediate future for our domains.

- The development of an alternative non-spectral solver seems feasible. As it turns out, specific properties of our NH dynamics can be used to greatly improve the performance of iterative solvers:
  - constant-coefficient semi-implicit

    $\Rightarrow$ predictable convergence = **robustness**
  - vertical decoupling

    $\Rightarrow$ optimal preconditioner parameters = **efficiency**

- The scalability of the preconditioned iterative solvers is really good!

RMI

Thank you