# A High-Order Optimized Schwarz algorithm for Massively Parallel Climate Modeling

Amik St-Cyr

Institute for the Mathematics Applied to Geosciences (IMAGE)
Computational & Information Systems Laboratory (CISL)
National Center for Atmospheric Research

Turbulence & Dynamos at Petaspeed Boulder, CO 15-19 October 2007

# Outline

- Constraints for petascale computing
- Spectral element method
- (Very short) introduction to DDM
- Optimized Schwarz (OS) at the algebraic level
- Transforming OS into an efficient algorithm
- Application to high-order SEM: the simple case
- Application: SEM based climate modeling
- Biting the bullet: “compressible solver”
- Conclusions

# Outline

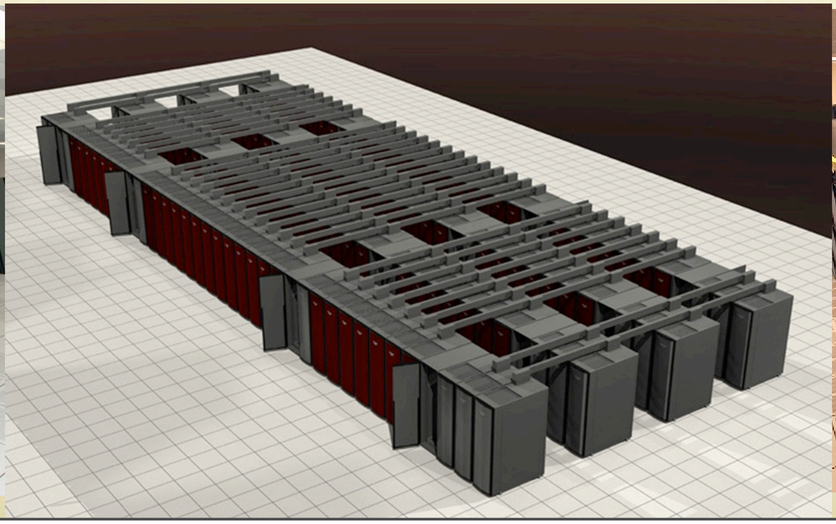
- Constraints for petascale computing
- ~~Spectral element method~~
- (Very short) introduction to DDM
- Optimized Schwarz (OS) at the algebraic level
- Transforming OS into an efficient algorithm
- Application to high-order SEM: the simple case
- Application: SEM based climate modeling
- Biting the bullet: “compressible solver”
- Conclusions

# What is a petascale machine?

- To follow Moore’s law, computers need more CPUs
- O(10000) to O(100000) processors
- Access to 10-20 times more processors: optimization not an option
- Tradeoffs: cache/heat/space/$
- Clock speed max out: burden is on parallel algorithms

# What is a petascale machine?

- To follow Moore's law, computers need more CPUs
- O(10000) to O(100000) processors
- Access to 10-20 times more processors: optimization not an option
- Tradeoffs: cache/heat/space/$
- Clock speed max out: burden is on parallel algorithms

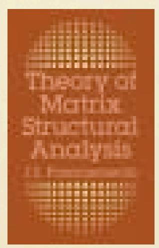


# Basic DD methods

- (Overlapping) Schwarz (1870): existence of elliptic problems on non trivial domains

- (Non-overlapping) Schur / sub-structuring methods

Kron (53)

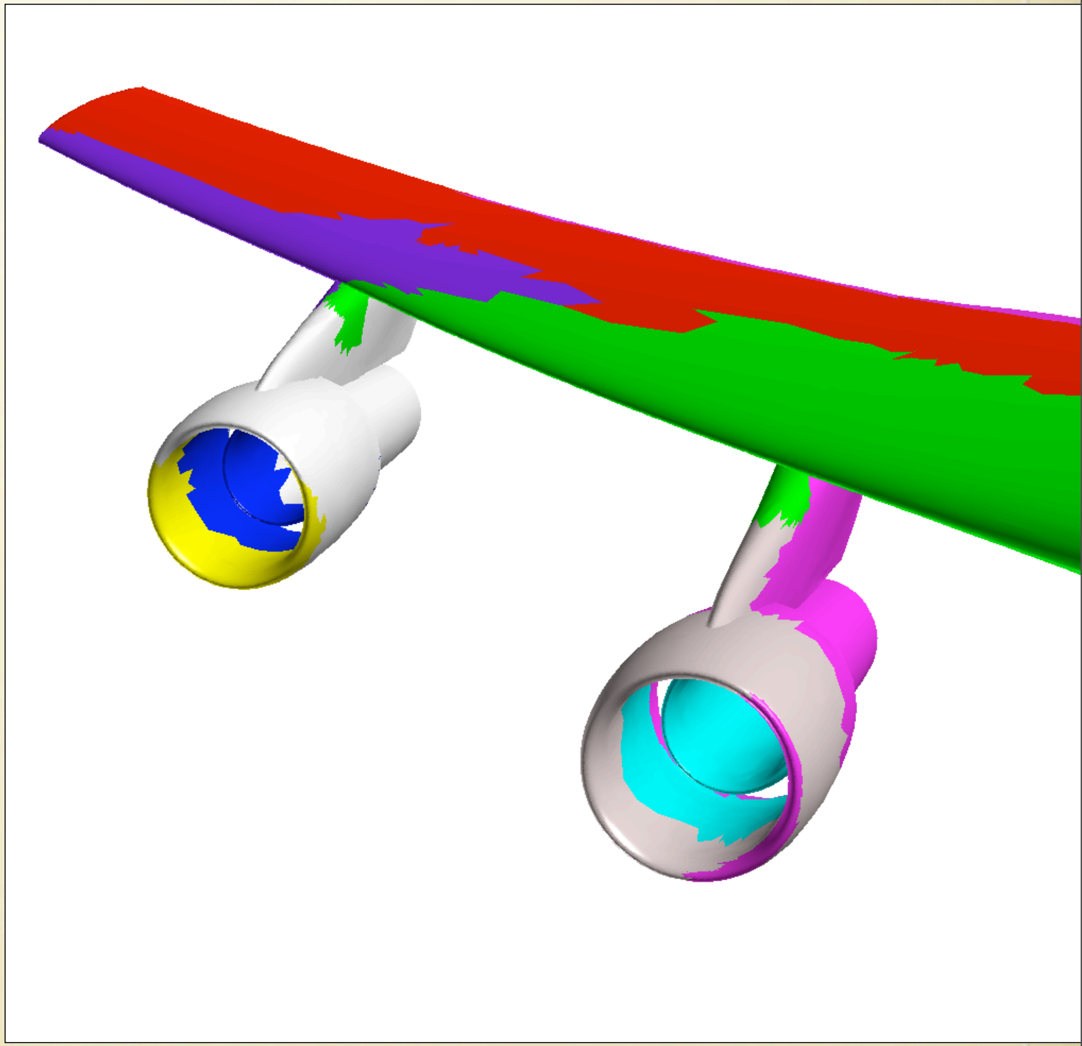



Przemieniecki (63)

2 classes of methods: overlapping and non-overlapping

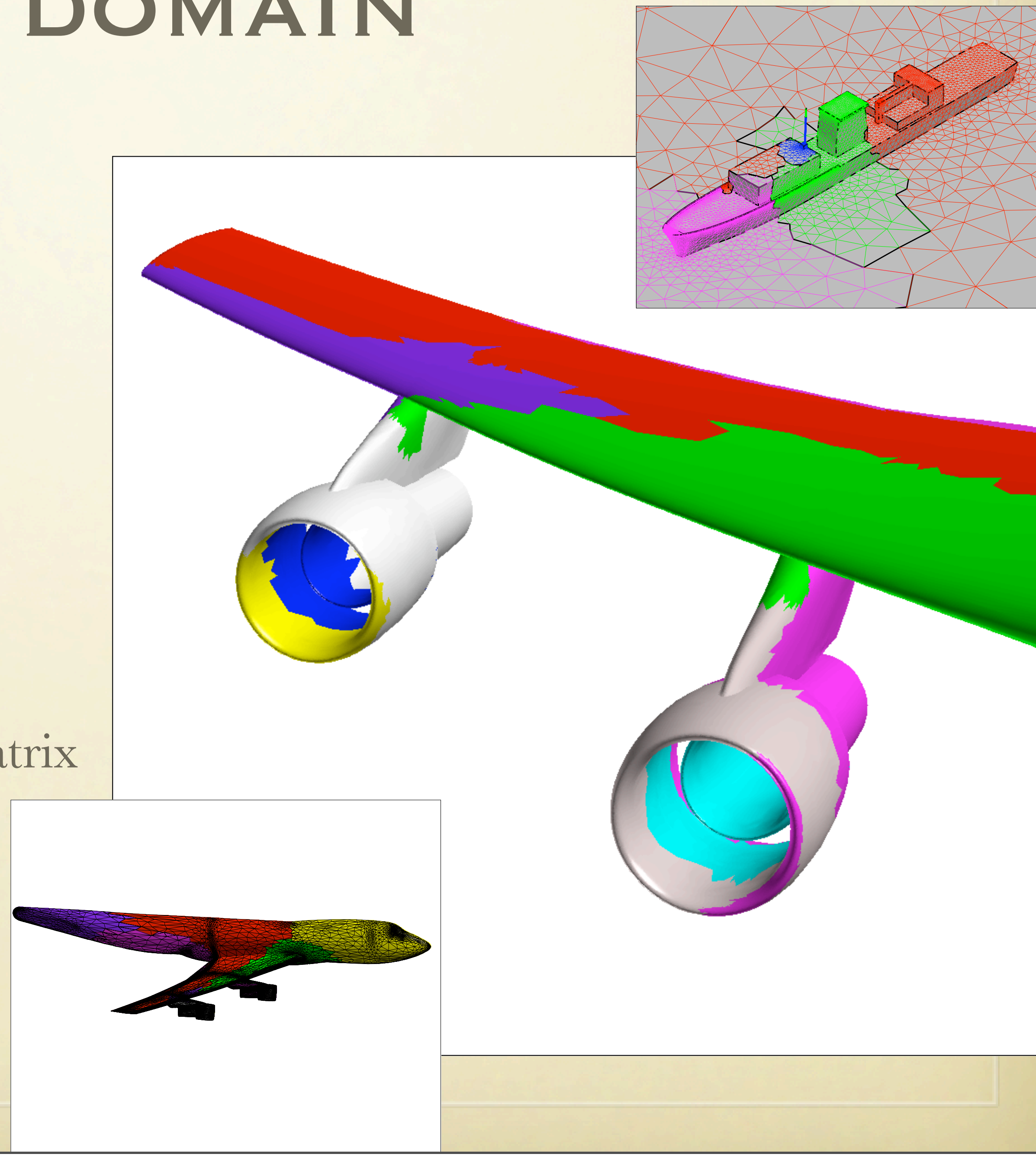
# Mesh partitioning: decompose the domain

- Geometric Based Algorithms
  - Coordinate bisection
  - Inertia bisection
- Graph Theory Based Algorithms
  - Graph bisection
  - Greedy algorithm
  - Spectral bisection
  - K-L algorithm
- Other Partitioning Algorithms
  - Global optimization algorithms
  - Reducing the bandwidth of the matrix
  - Index based algorithms
- The State of the Art
  - Hybrid approach
  - Multilevel approach
  - Parallel partitioning algorithms



# Mesh partitioning: decompose the domain

- Geometric Based Algorithms
  - Coordinate bisection
  - Inertia bisection
- Graph Theory Based Algorithms
  - Graph bisection
  - Greedy algorithm
  - Spectral bisection
  - K-L algorithm
- Other Partitioning Algorithms
  - Global optimization algorithms
  - Reducing the bandwidth of the matrix
  - Index based algorithms
- The State of the Art
  - Hybrid approach
  - Multilevel approach
  - Parallel partitioning algorithms



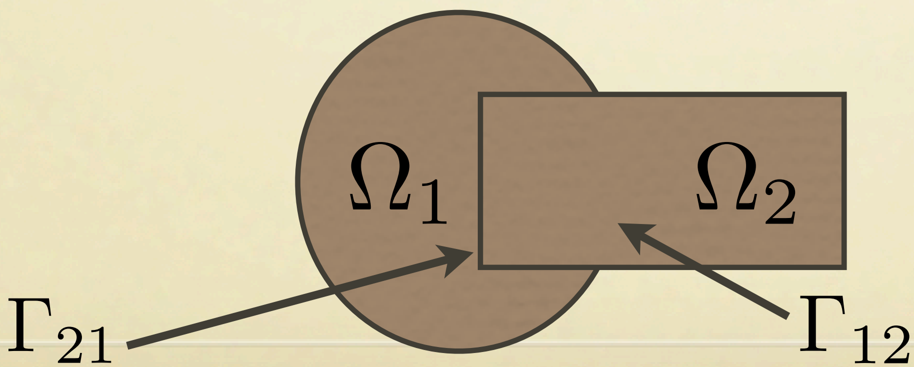
# Classical Schwarz

Suppose we need to solve:

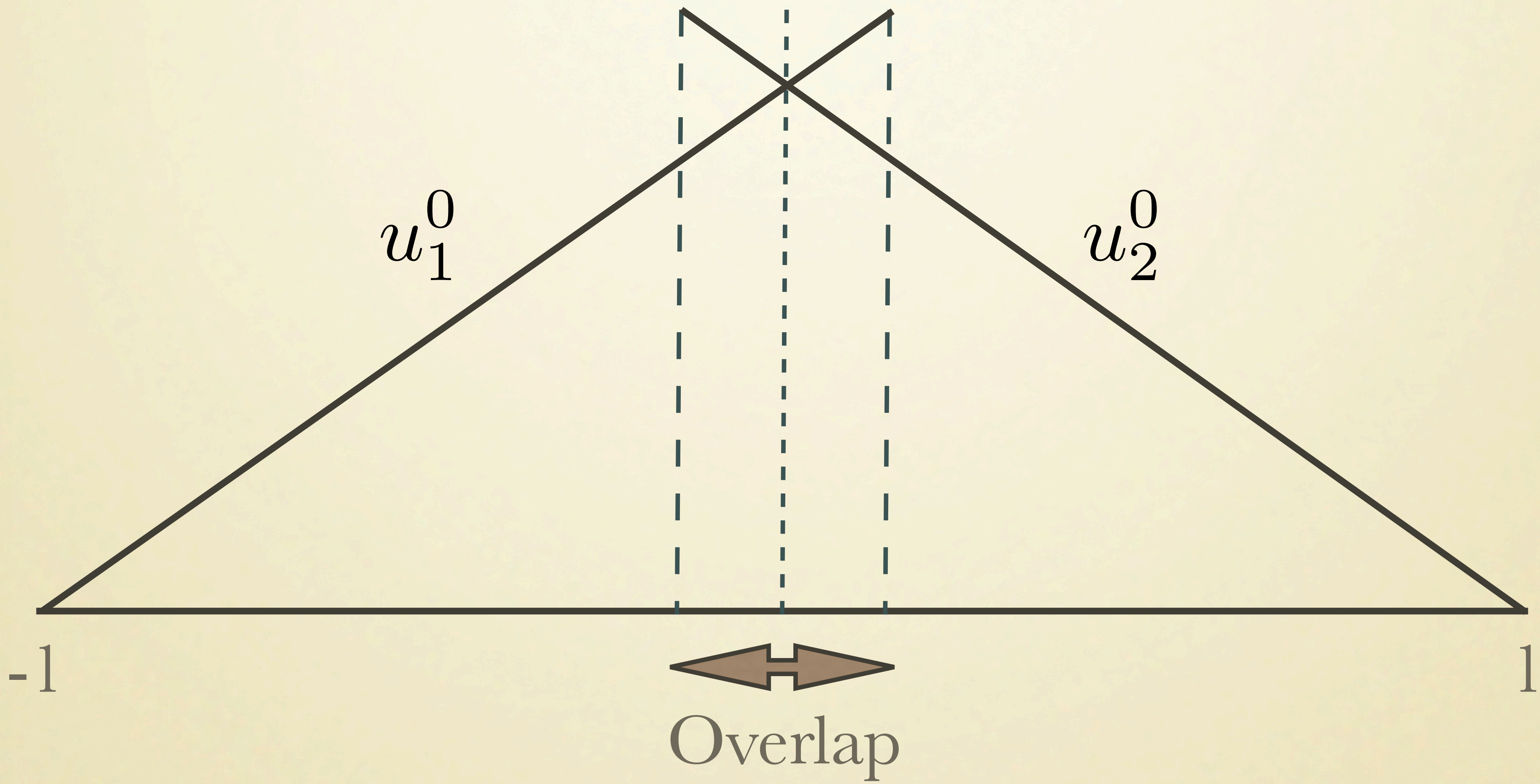
$$\mathcal{L}u = f \quad \text{in } \Omega, \qquad \mathcal{B}u = g \quad \text{on } \partial\Omega$$

Partition the original domain into 2 domains:

$$\begin{aligned} \mathcal{L}u_1^{n+1} &= f && \text{in } \Omega_1, & \mathcal{L}u_2^{n+1} &= f && \text{in } \Omega_2, \\ \mathcal{B}(u_1^{n+1}) &= g && \text{on } \partial\Omega_1, & \mathcal{B}(u_2^{n+1}) &= g && \text{on } \partial\Omega_2, \\ u_1^{n+1} &= u_2^n && \text{on } \Gamma_{12}, & u_2^{n+1} &= u_1^n && \text{on } \Gamma_{21}. \end{aligned}$$

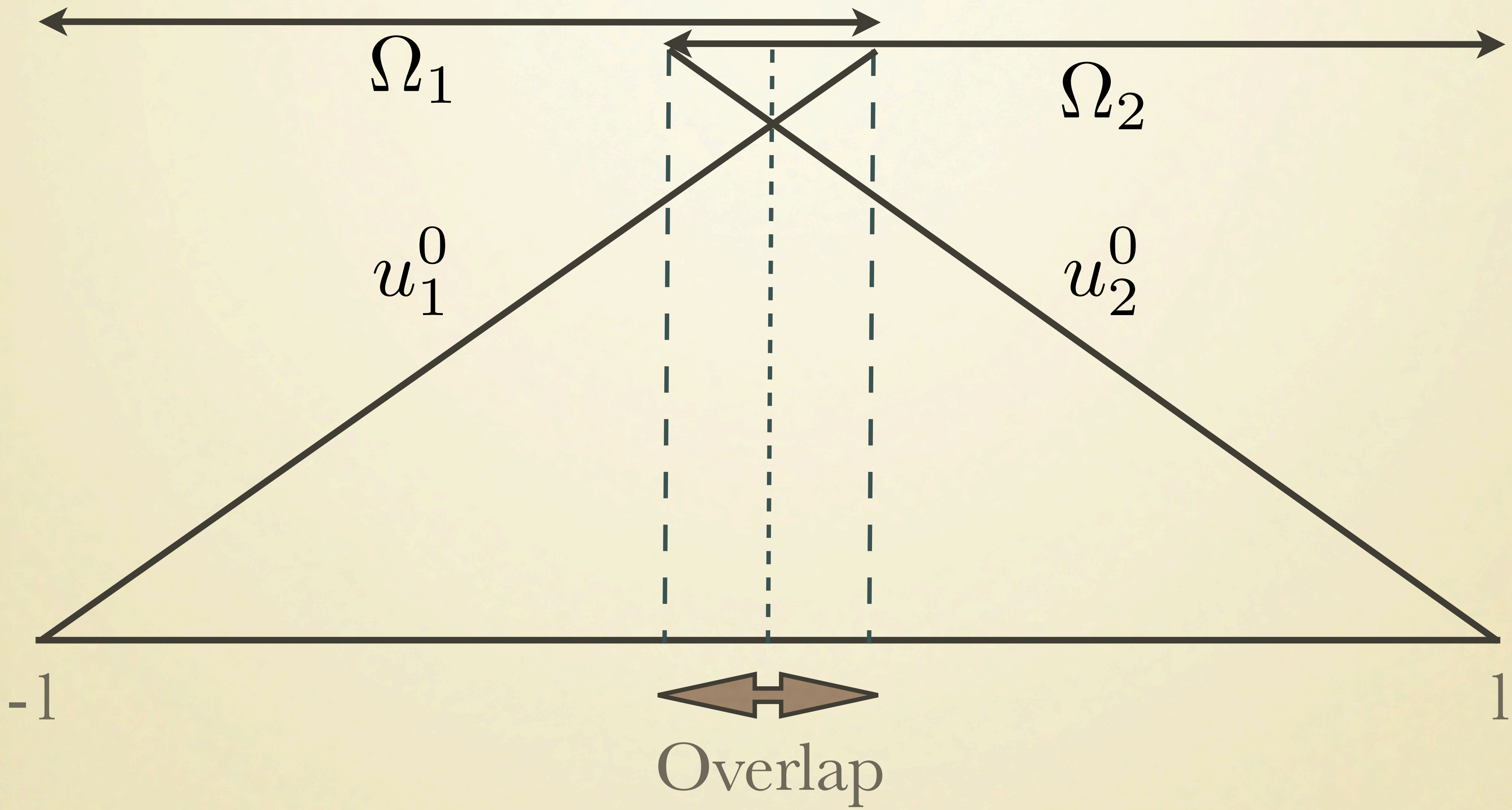


# Schwarz with large overlap



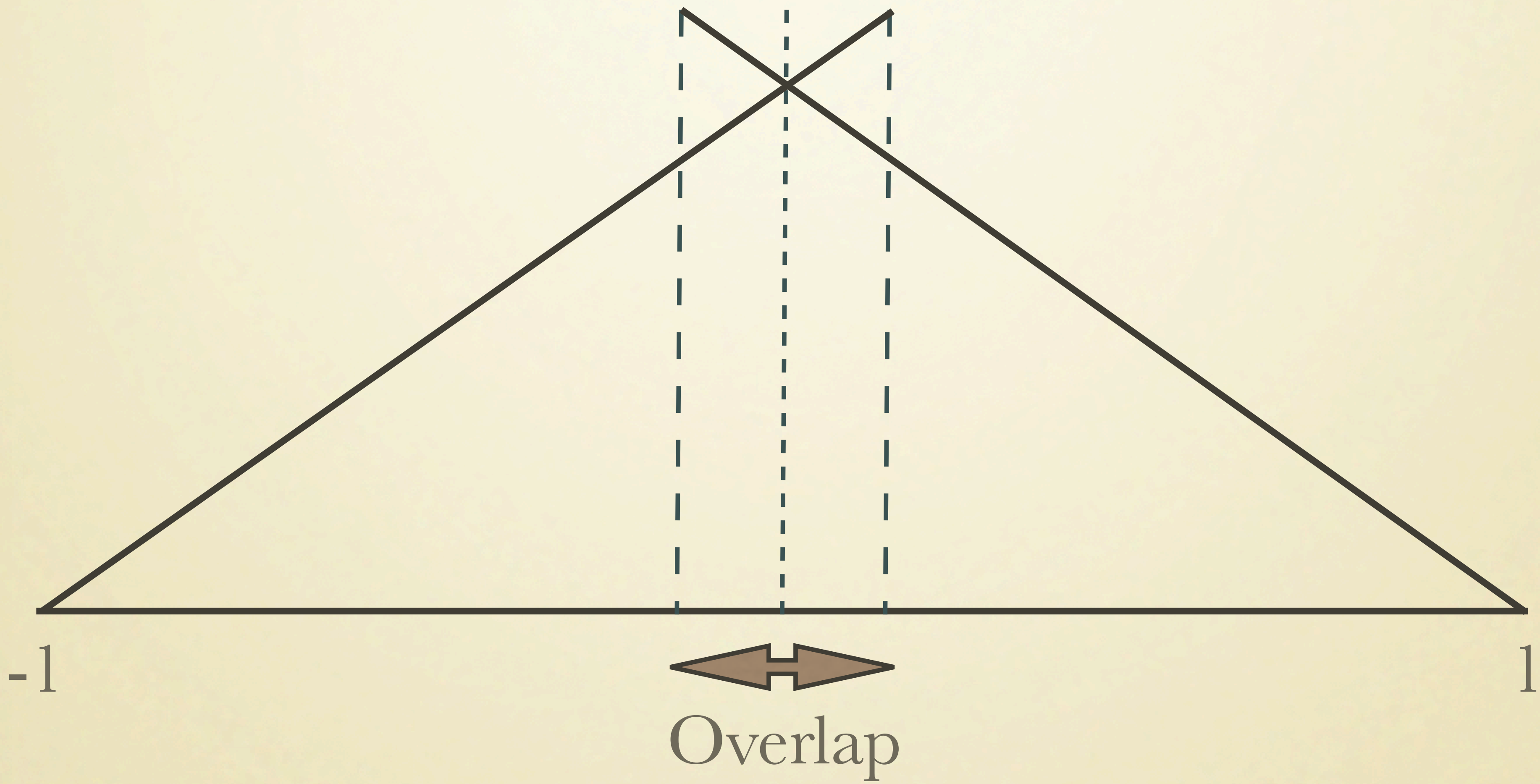

$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



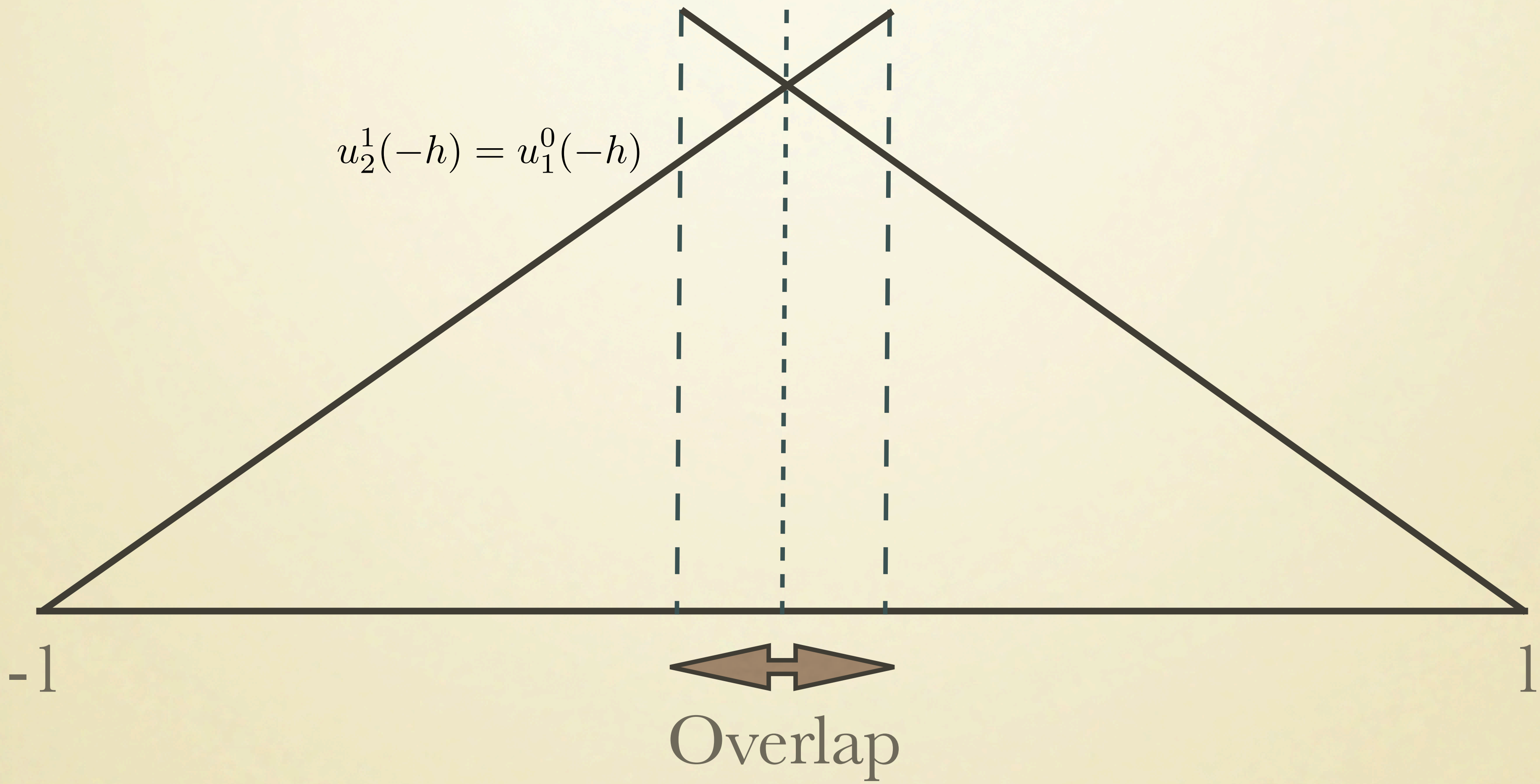

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



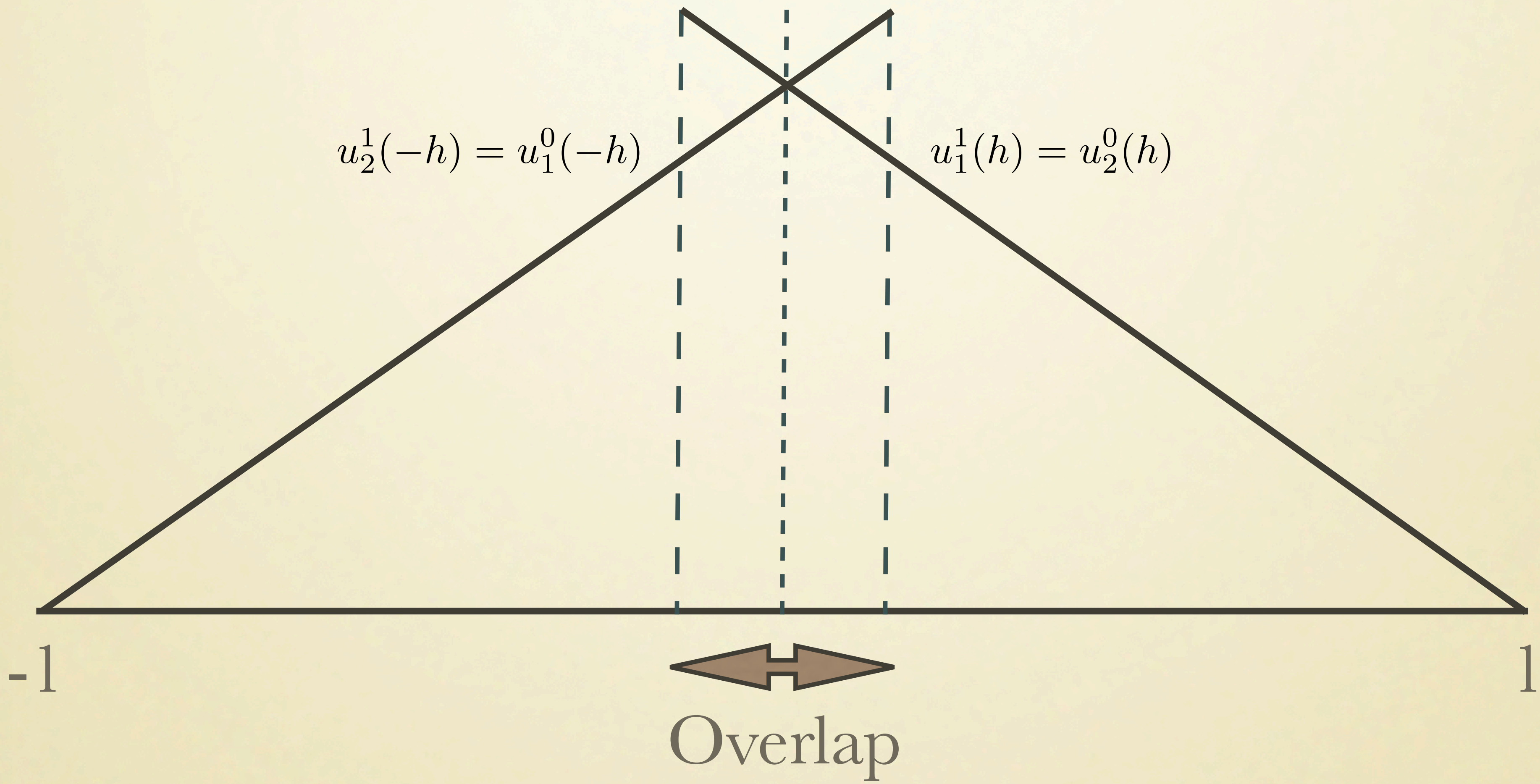

$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



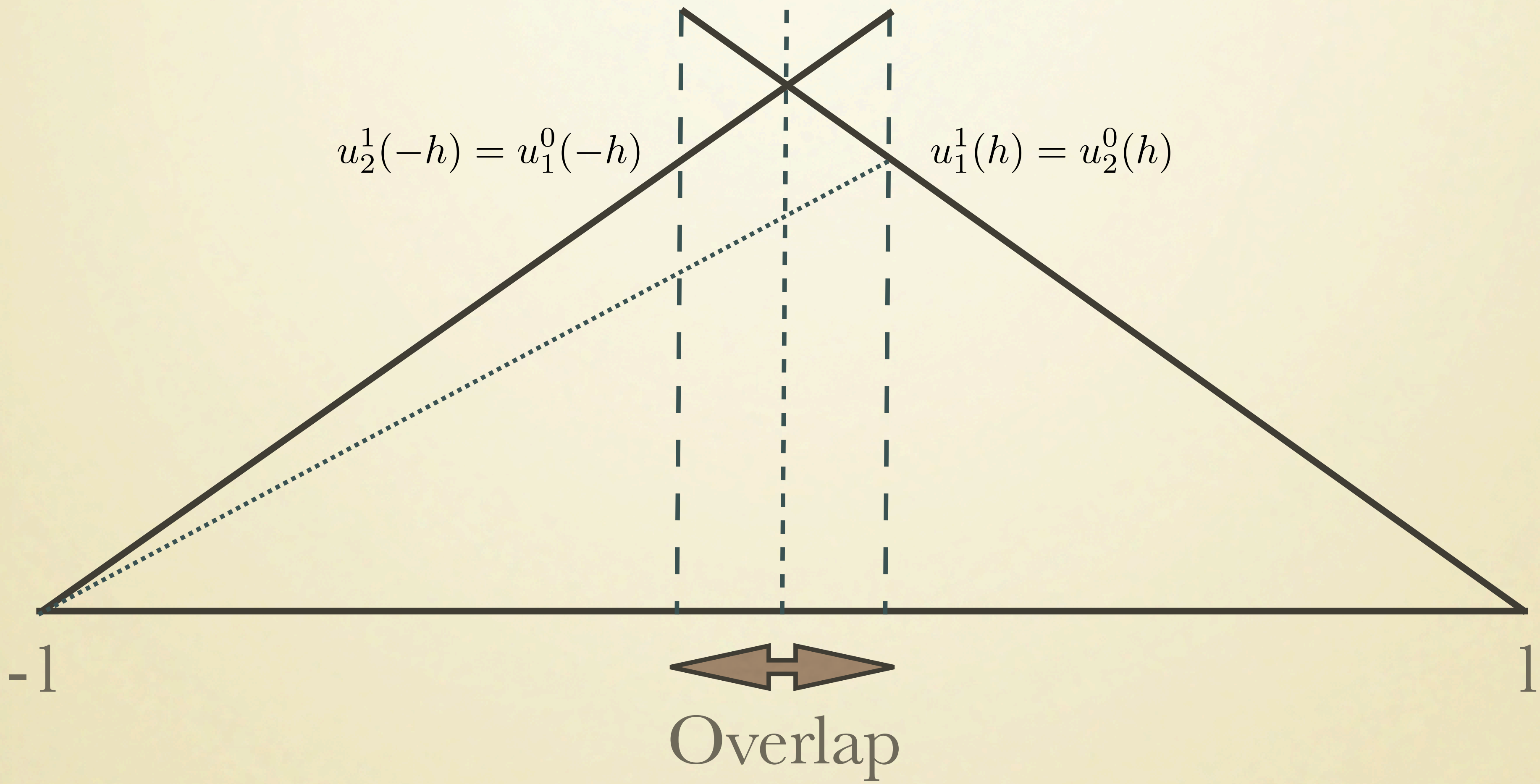

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



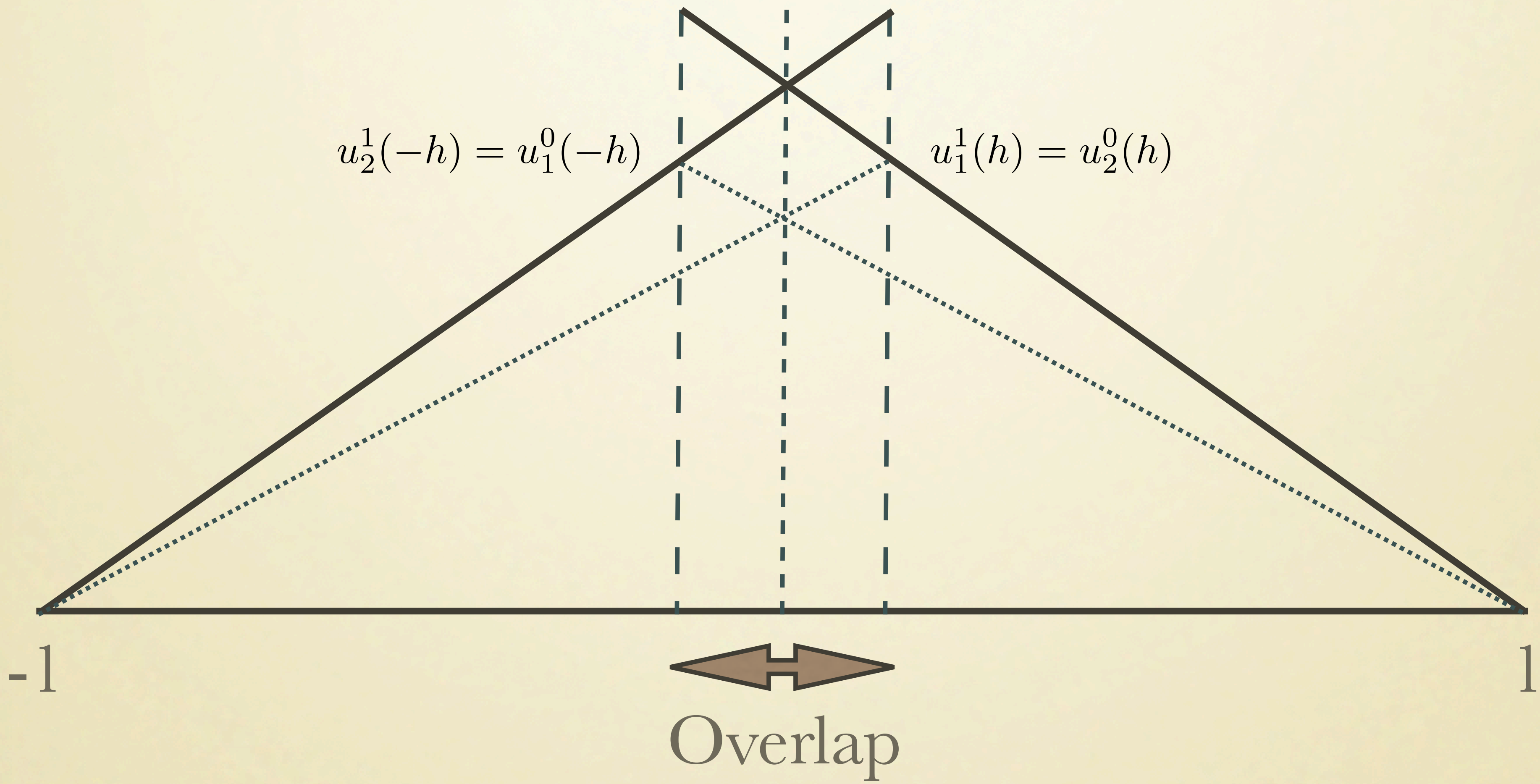

$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



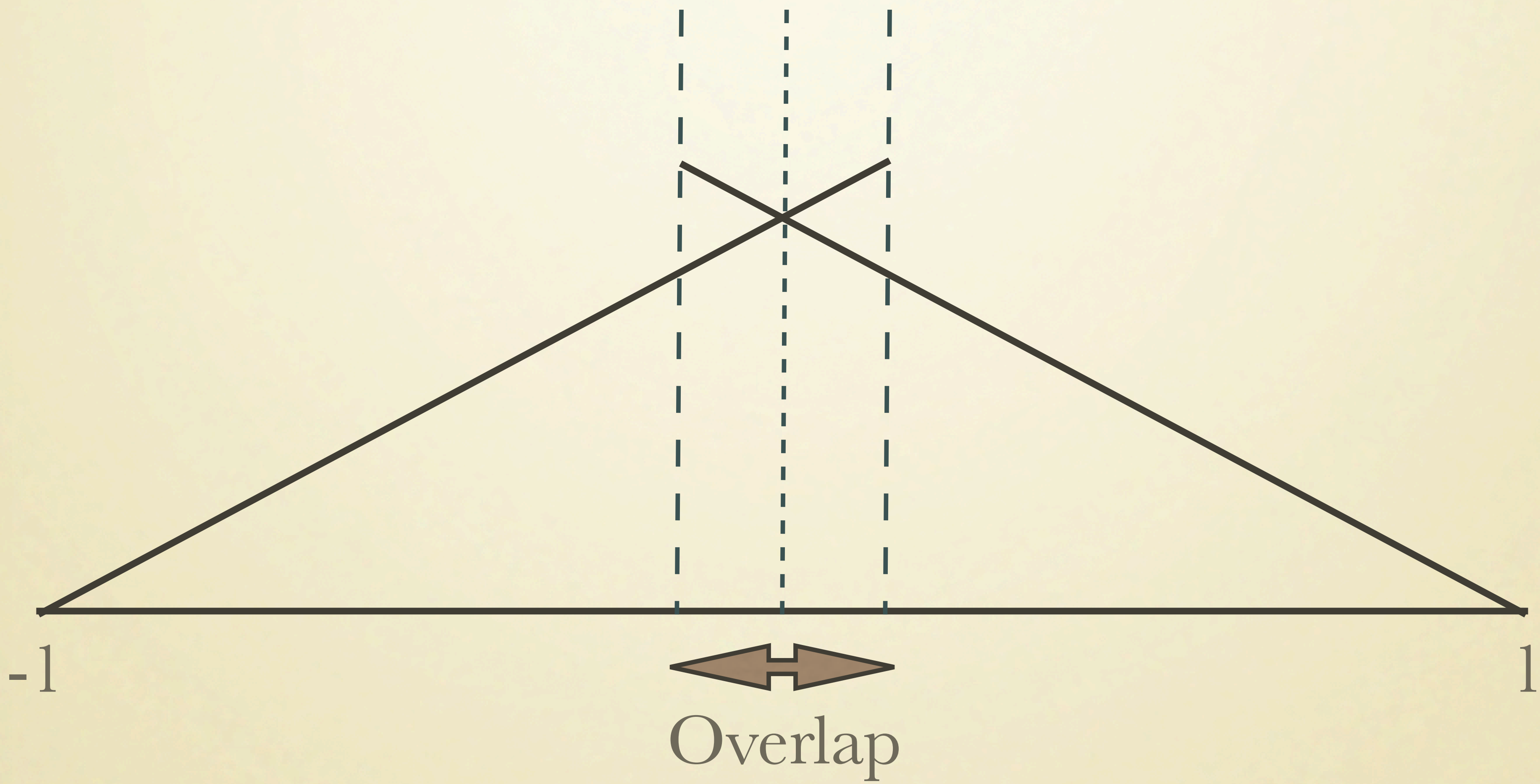

$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



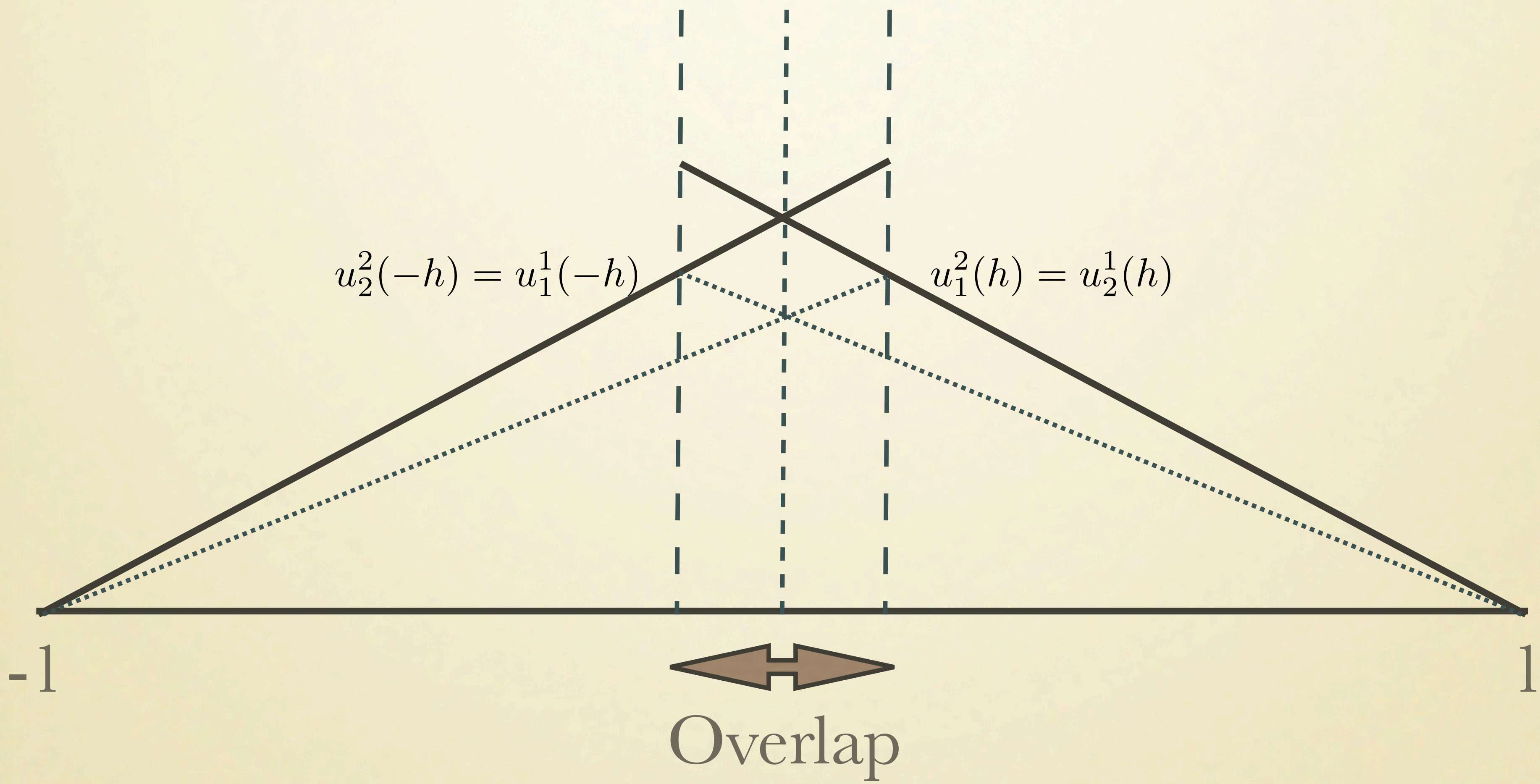

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



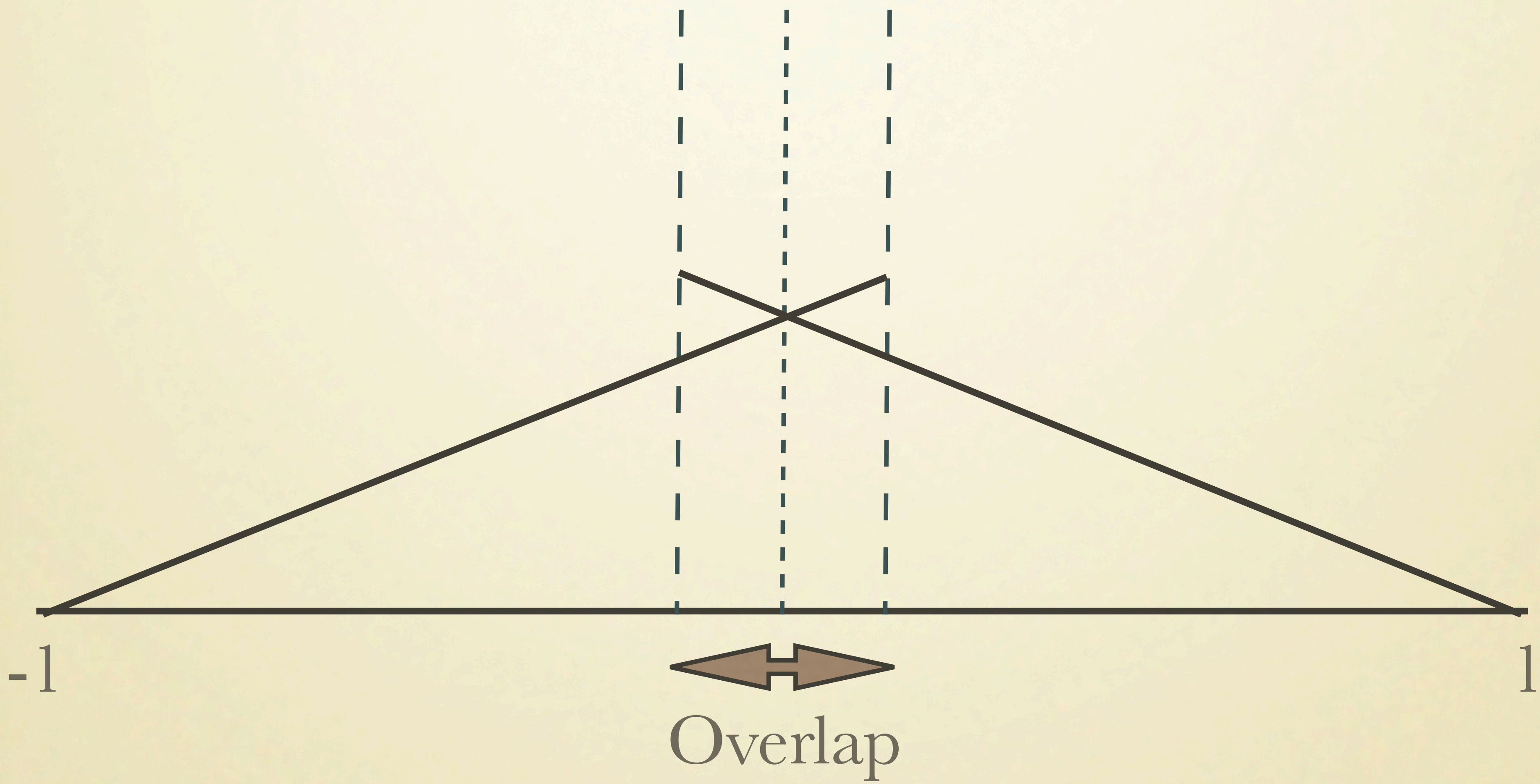

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



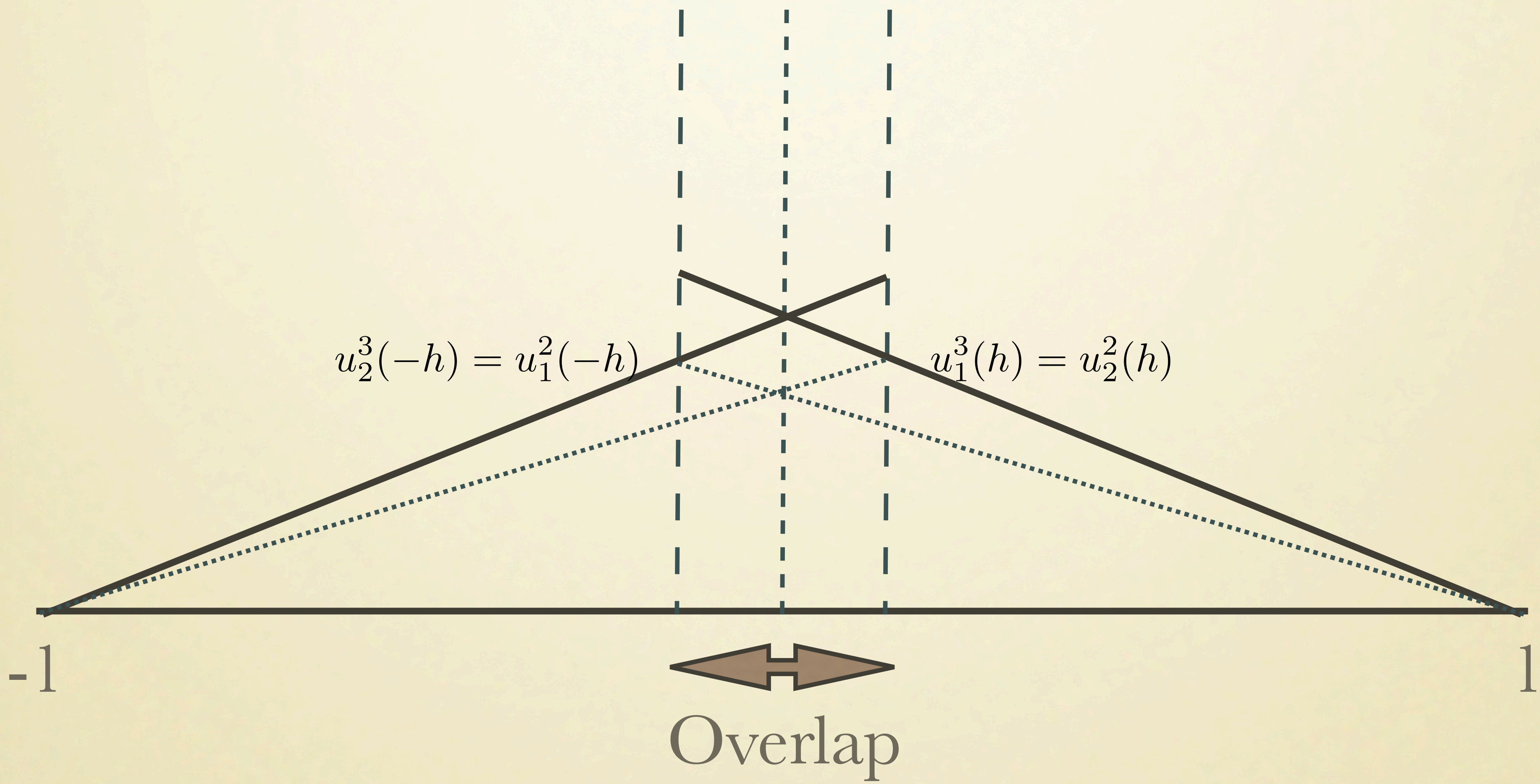

$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



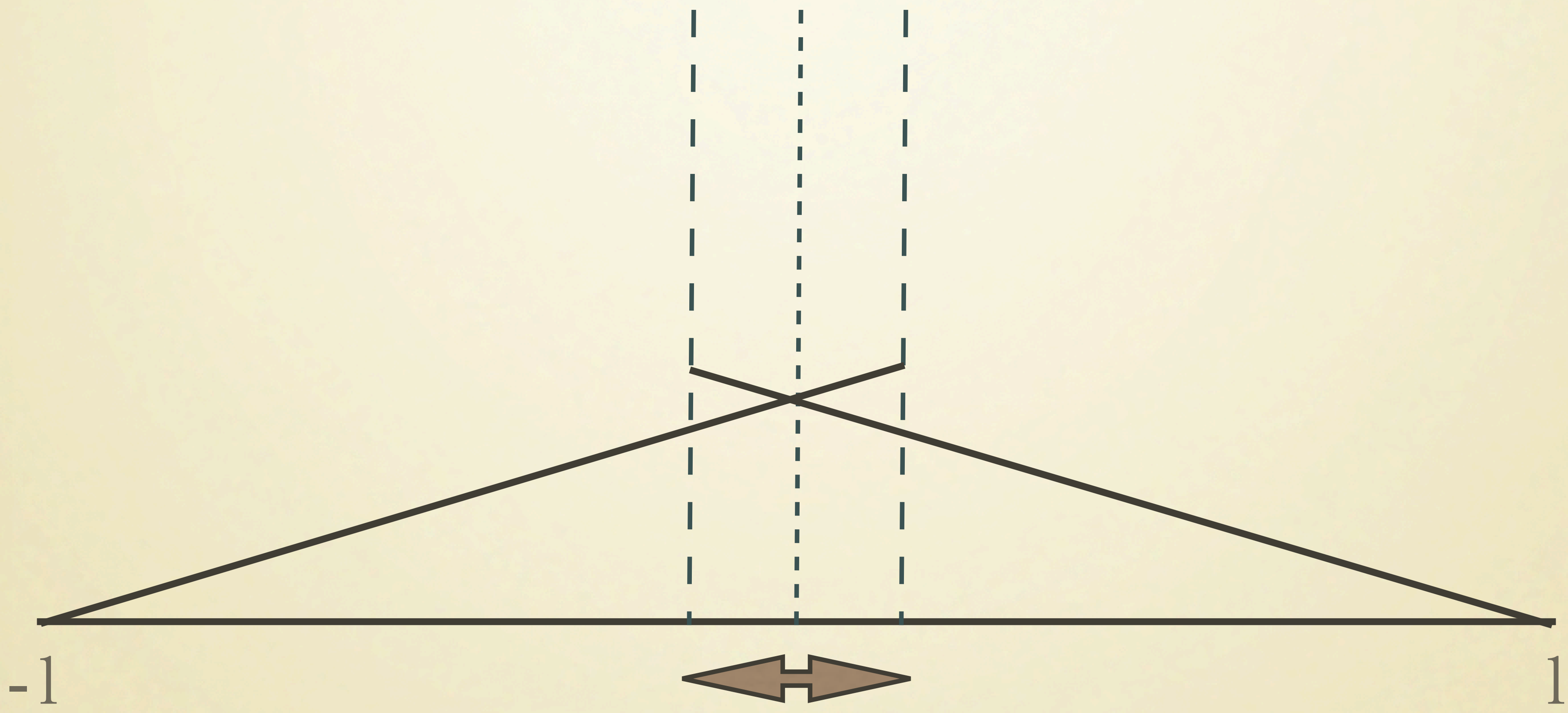

$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



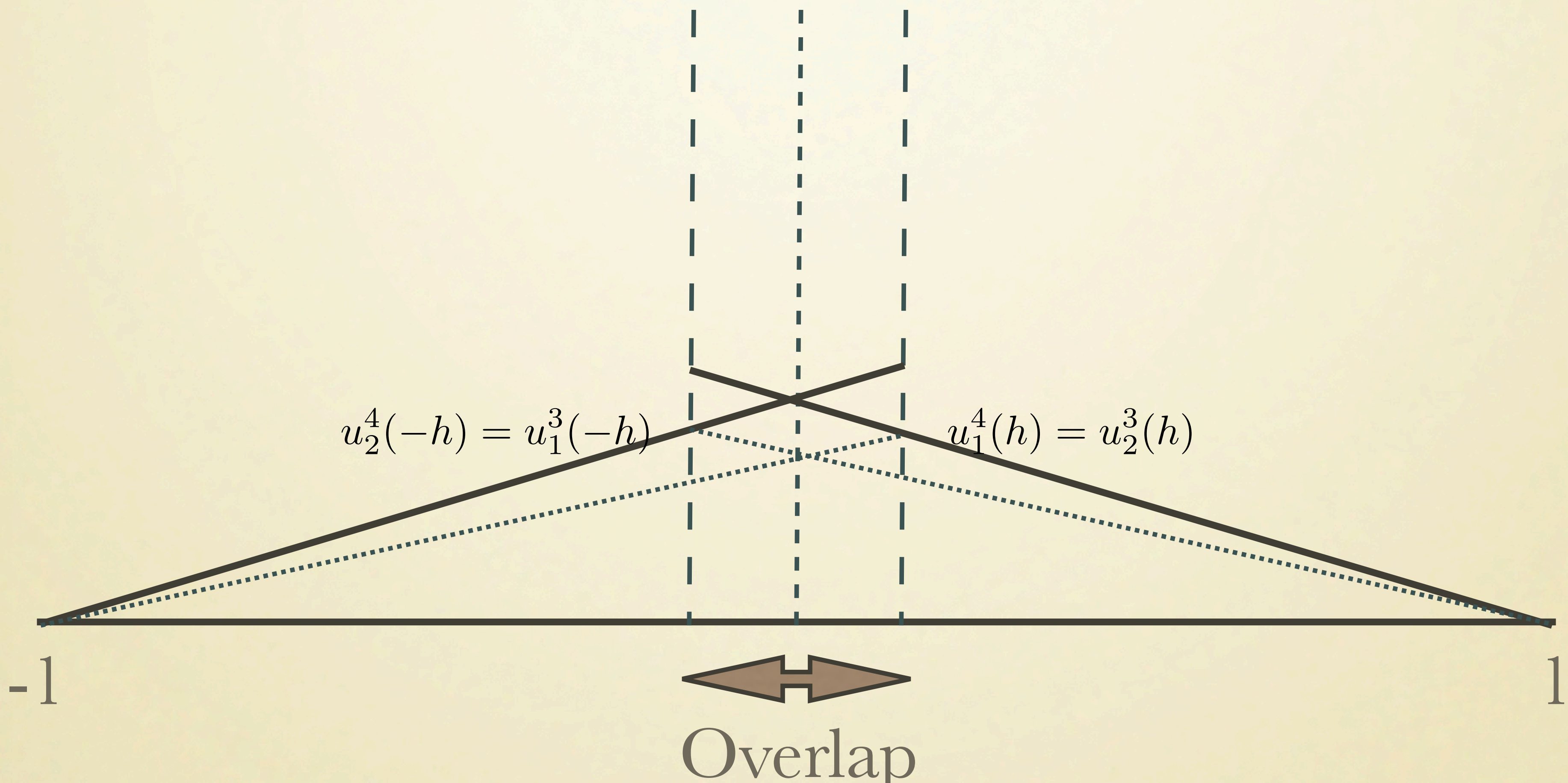

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



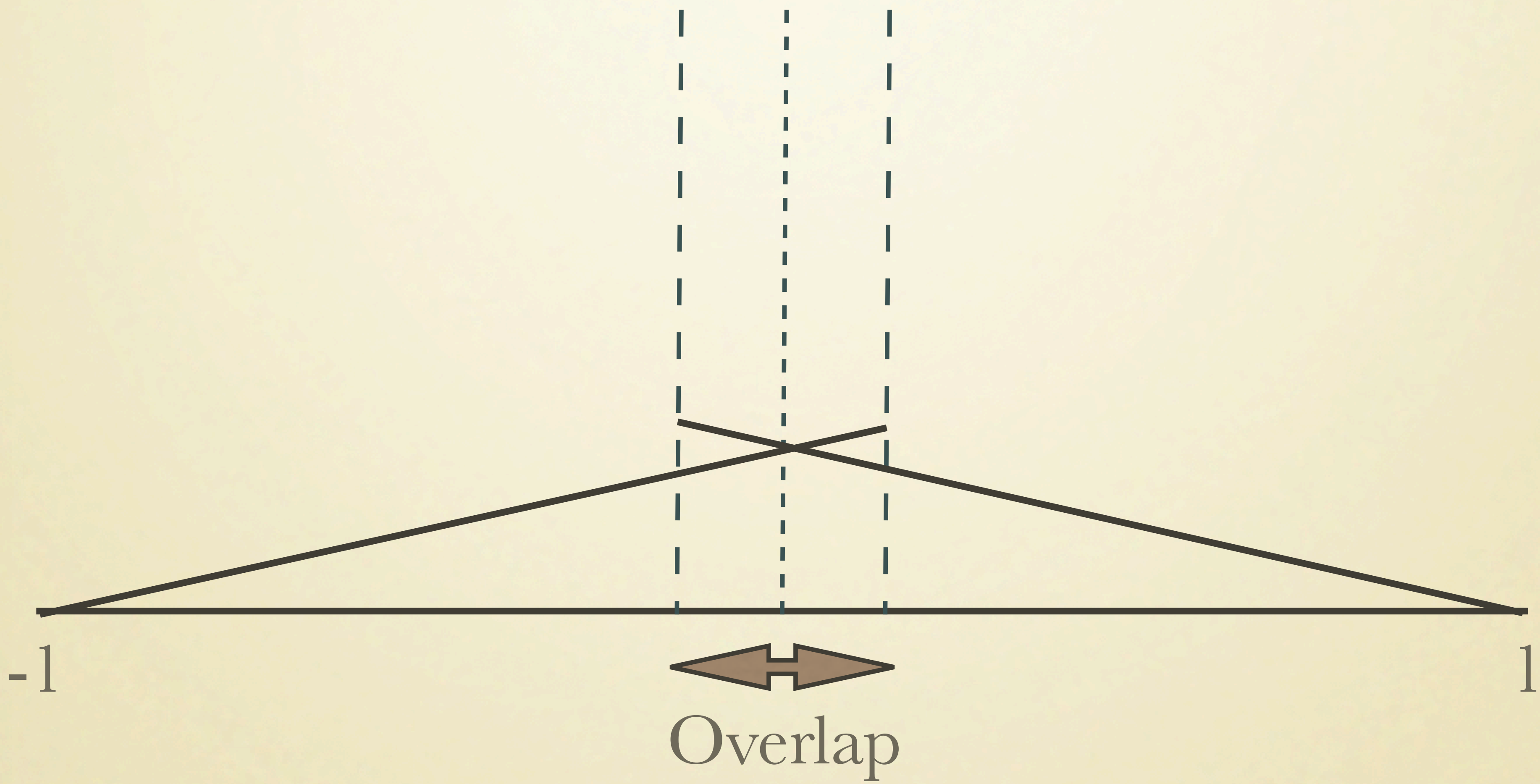

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



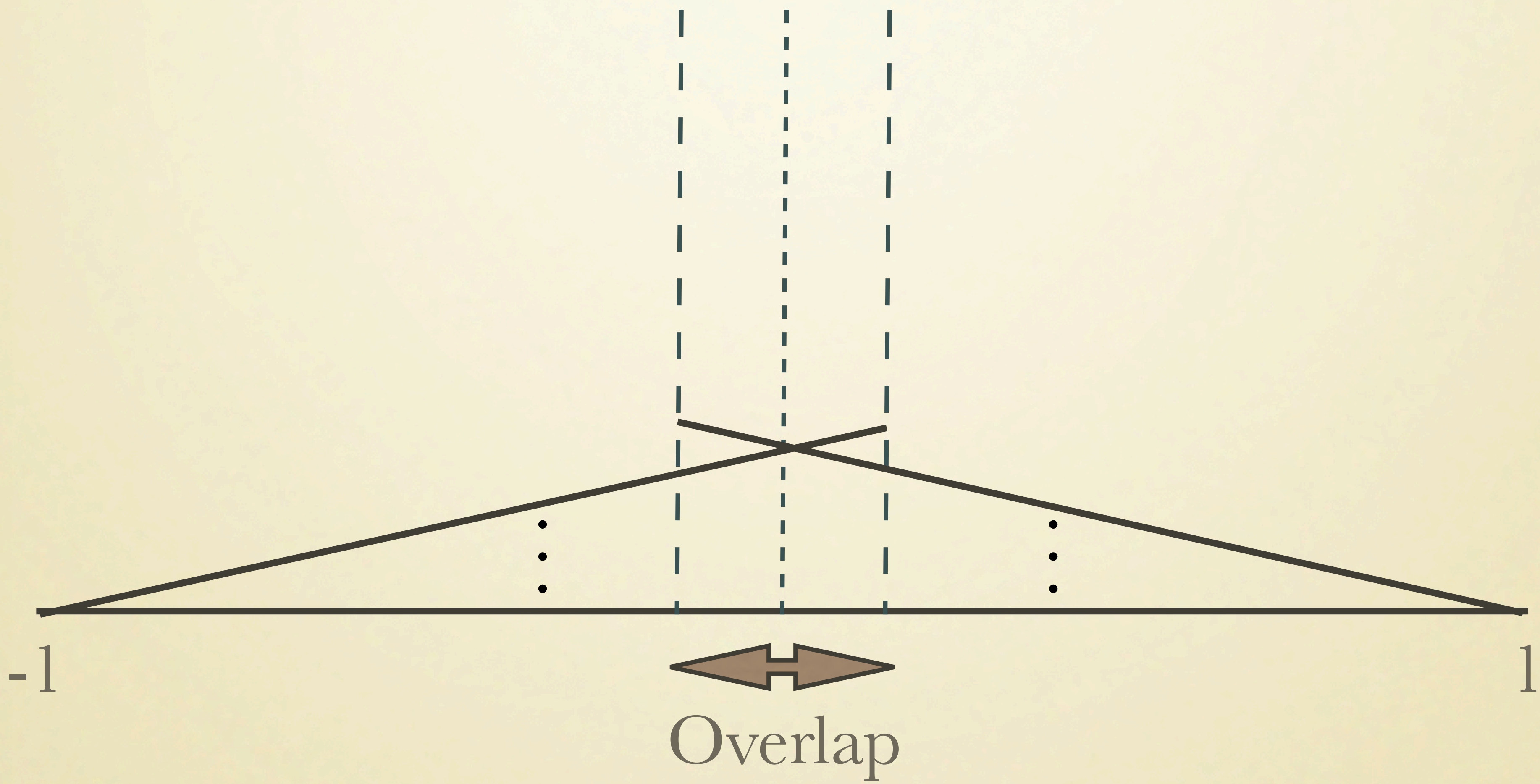

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap



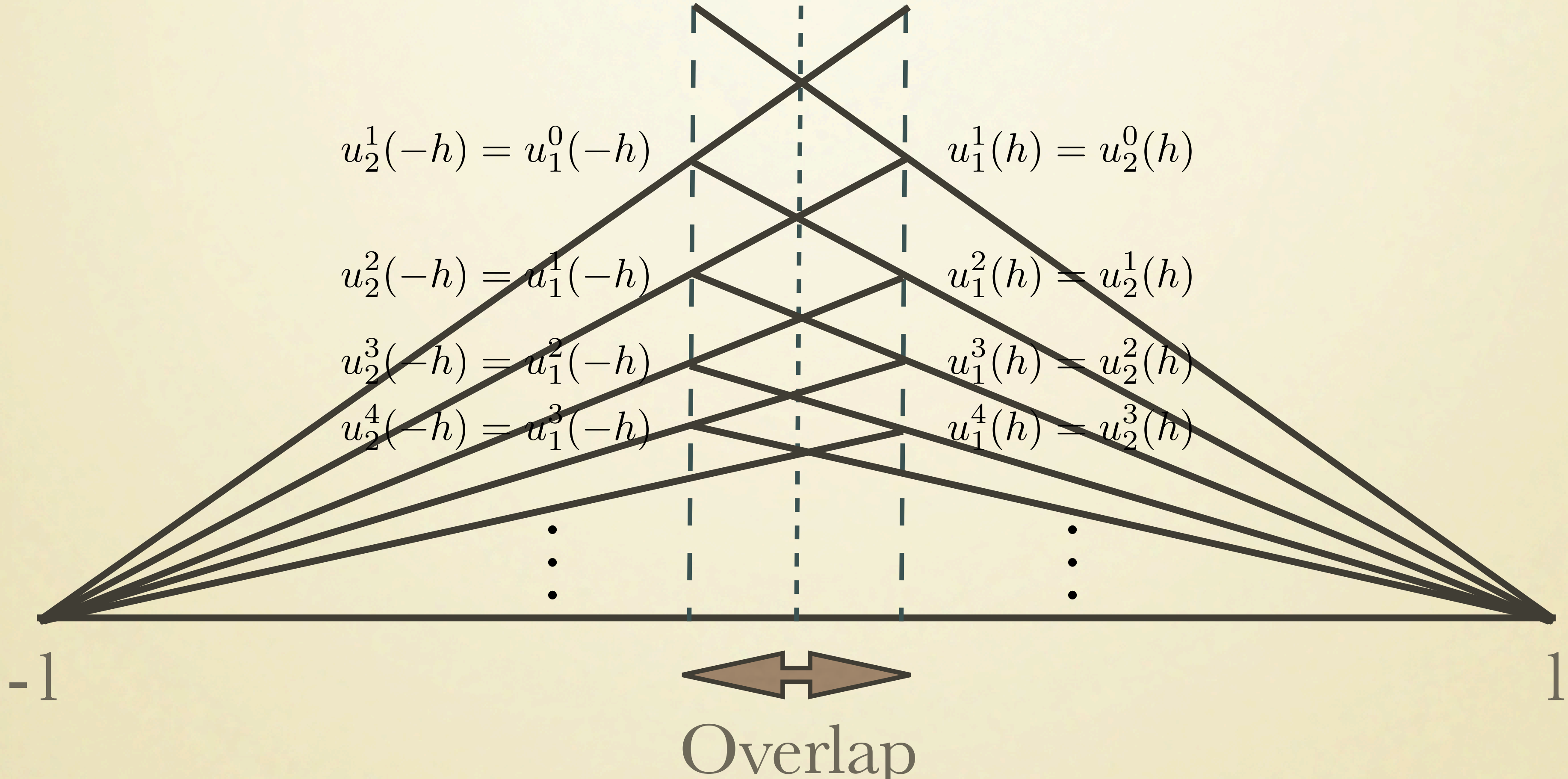

$$\Delta u = 0, \text{ on } [-1, 1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap




$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# Schwarz with large overlap




$$\Delta u = 0, \text{ on } [-1,1] \text{ with } u(-1) = u(1) = 0$$

# The Robin method

- Lions (1990)
- Used to accelerate convergence of Schwarz
- Free positive parameter: how to find its correct value?
- Convergence rate not demonstrated theoretically
- No need for overlap!

$$\mathcal{L}u_j^{k+1} = u_j^{k+1} - \Delta u_j^{k+1} = f_j$$

$$pu_j^{k+1} + \frac{\partial u_j^{k+1}}{\partial \mathbf{n}_{jl}} = pu_l^k + \frac{\partial u_l^k}{\partial \mathbf{n}_{jl}} \text{ on } \partial\Omega_j \cap \partial\Omega_l \text{ for } l \in \mathcal{N}(\Omega_j)$$

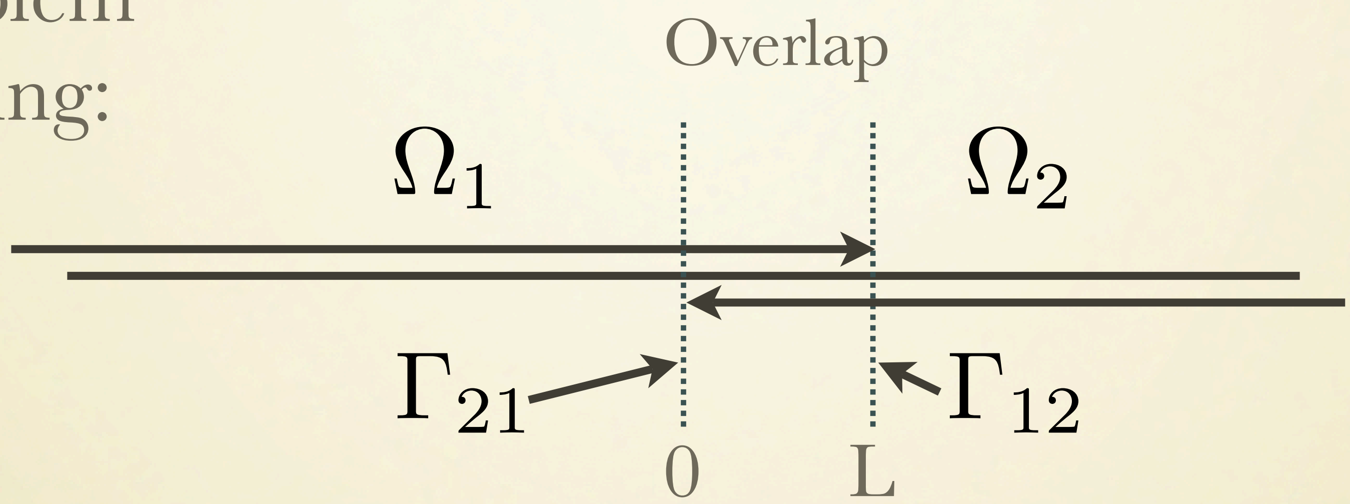
$$u_j^{k+1} = u_0 \text{ on } \partial\Omega_j \cap \partial\Omega$$

# Fourier analysis

- Study simple 2D problem
- Only 2 subdomains
- Fourier transform in the tangent direction to the separating interface between domains
- Solve the remaining ODE
- Obtain convergence rate of the algorithm

# Fourier analysis

Problem setting:




$$(\eta - \Delta)u(x, y) = 0, \text{ on } \Omega$$

Boundary conditions: solution decays at infinity

Subdomains:

$$\Omega_1 = [-\infty, L] \times \mathbb{R} \text{ and } \Omega_2 = [0, \infty] \times \mathbb{R}$$

# Fourier analysis

Two subproblems:

$$\begin{array}{rcll} (\eta-\Delta)u_1^{n+1} & = & 0 & \text{in } \Omega_1, \\ u_1^{n+1}(L,y) & = & u_2^n(L,y) & \text{on } \Gamma_{12}, \end{array} \qquad \begin{array}{rcll} (\eta-\Delta)u_2^{n+1} & = & 0 & \text{in } \Omega_2, \\ u_2^{n+1}(0,y) & = & u_1^n(0,y) & \text{on } \Gamma_{21}. \end{array}$$

Fourier transforming in the y direction:

$$\begin{array}{rcll} (\eta+k^2-\partial_{xx})\hat{u}_1^{n+1} & = & 0 & \text{in } \Omega_1, \\ \hat{u}_1^{n+1}(L,k) & = & \hat{u}_2^n(L,k) & \text{on } \Gamma_{12}, \end{array} \qquad \begin{array}{rcll} (\eta+k^2-\partial_{xx})\hat{u}_2^{n+1} & = & 0 & \text{in } \Omega_2, \\ \hat{u}_2^{n+1}(0,k) & = & \hat{u}_1^n(0,k) & \text{on } \Gamma_{21}. \end{array}$$

Solving in the x direction:

$$\hat{u}_1^n(x,k) = \hat{u}_2^{n-1}(L,k)e^{-\sqrt{k^2+\eta}(x-L)}, \qquad \hat{u}_2^n(x,k) = \hat{u}_1^{n-1}(0,k)e^{-\sqrt{k^2+\eta}x}$$

Convergence rate of classical Schwarz (Gander 2006 SINUM):

$$\rho_{cla} = \rho_{cla}(k,\eta,L) = e^{-\sqrt{k^2+\eta}L}$$

# Optimized approach

- Inspired by the Robin problem:

$$\begin{aligned}(\eta-\Delta)u_1^{n+1} &= 0 && \text{in } \Omega_1, & (\eta-\Delta)u_2^{n+1} &= 0 && \text{in } \Omega_2,\\ (\partial_x+S_1)u_1^{n+1} &= (\partial_x+S_1)u_2^n && \text{on } \Gamma_{12}, & (\partial_x+S_2)u_2^{n+1} &= (\partial_x+S_2)u_1^n && \text{on } \Gamma_{21}.\end{aligned}$$

We are looking for the best possible forms of in Fourier space

Proceeding as before leads to the solutions: $\left(\sigma_r(k)=\mathcal{F}(S_r)\right)$

$$\hat{u}_1^n(x,k) = \frac{\sigma_1(k)-\sqrt{k^2+\eta}}{\sigma_1(k)+\sqrt{k^2+\eta}}e^{-\sqrt{k^2+\eta}(x-L)}\hat{u}_2^{n-1}(L,k), \qquad \hat{u}_2^n(x,k) = \frac{\sigma_2(k)+\sqrt{k^2+\eta}}{\sigma_2(k)-\sqrt{k^2+\eta}}e^{-\sqrt{k^2+\eta}x}\hat{u}_1^{n-1}(0,k)$$

New convergence rate:

$$\rho_{opt}=\rho_{opt}(k,\eta,L)=\frac{\sigma_1(k)-\sqrt{k^2+\eta}}{\sigma_1(k)+\sqrt{k^2+\eta}}\frac{\sigma_2(k)+\sqrt{k^2+\eta}}{\sigma_2(k)-\sqrt{k^2+\eta}}e^{-2\sqrt{k^2+\eta}L}$$

# Optimized approach

The choice

$$\sigma_1(k) = \sqrt{k^2+\eta},\ \sigma_2(k) = -\sqrt{k^2+\eta}$$

leads to the convergence of the algorithm in 2 iterations

$$\rho_{opt} = 0$$

The operators are <u>not local</u> operators in physical space!

An approximation is sought such that all frequencies have an optimal decay rate:

$$\boxed{\sigma_1^{app}(k) = p_1 + q_1 k^2,\ \sigma_2^{app}(k) = -p_2 - q_2 k^2}$$

# Various choices (one sided)

Taylor zeroth order: $\sigma_1^{app}(k) = \sqrt{\eta}$

Taylor second order: $\sigma_1^{app}(k) = \sqrt{\eta} + \frac{1}{2\sqrt{\eta}}k^2$

Zeroth order optimized: $k(L,\eta,p) = \frac{\sqrt{L(2p + L(p^2-\eta))}}{L}$

$$\rho_{OO0}(k_{\min}, L, \eta, p^*) = \rho_{OO0}(k(p^*), L, \eta, p^*)$$

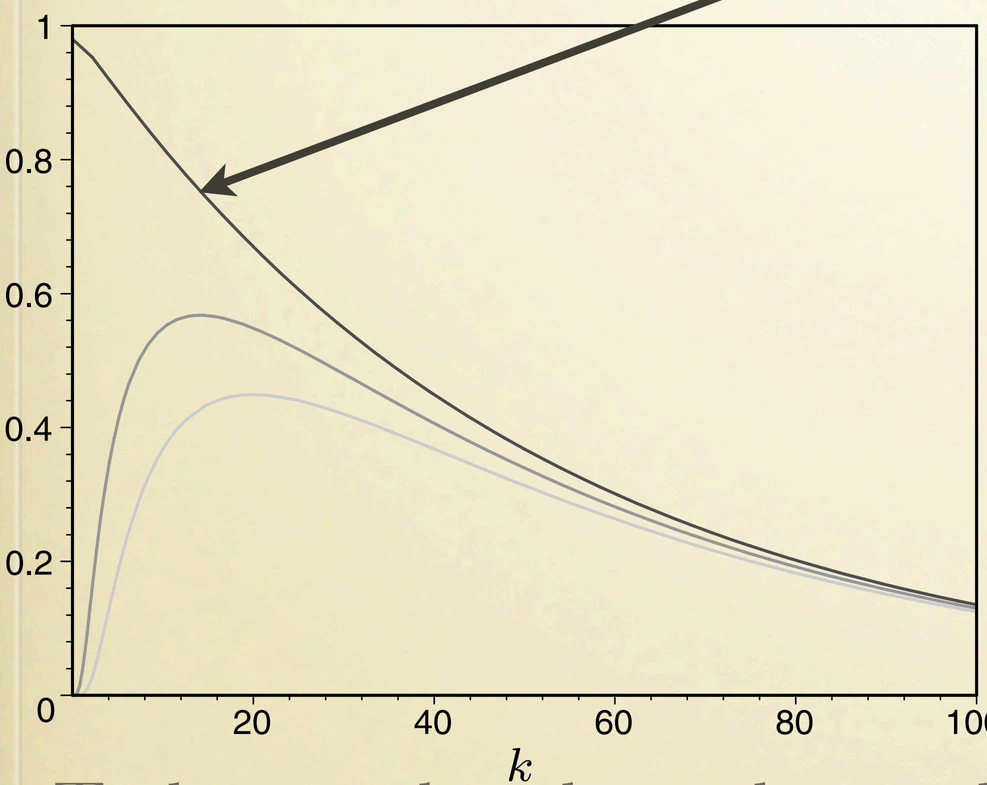
Zeroth order optimized (no overlap): $p^* = ((k_{\min}^2 + \eta)(k_{\max}^2 + \eta))^{\frac{1}{4}}$

Second order optimized: very long and complex formulas for p and q ...

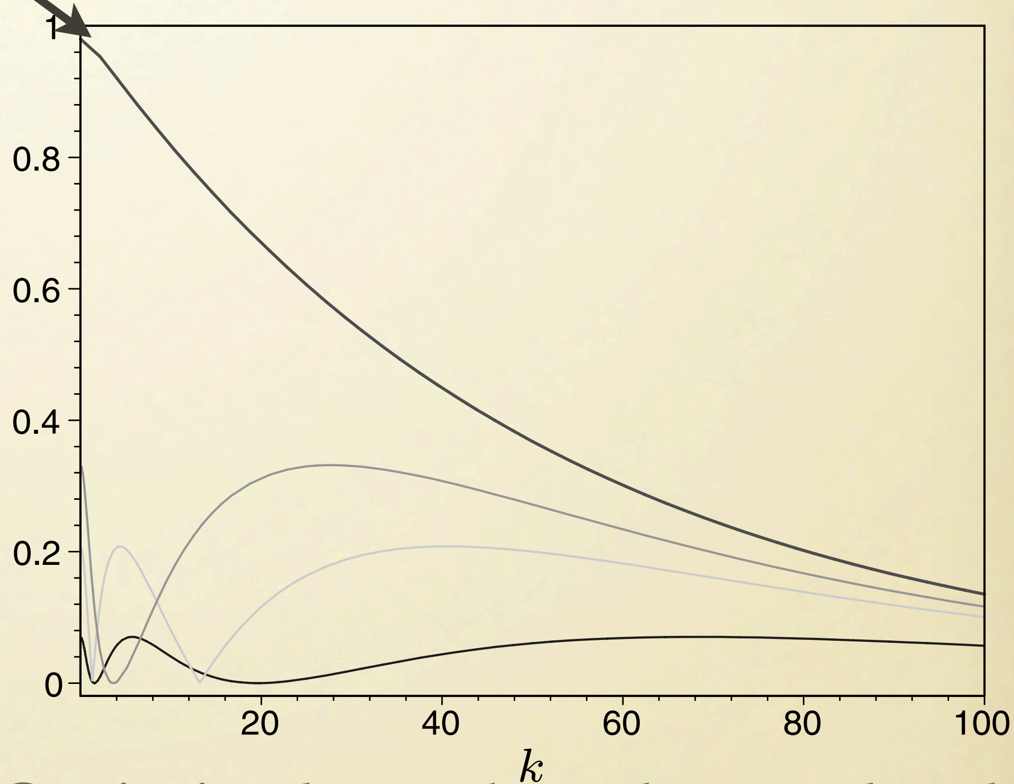
Details see Gander (SINUM 2006)

# Convergence rates

Classical Schwarz



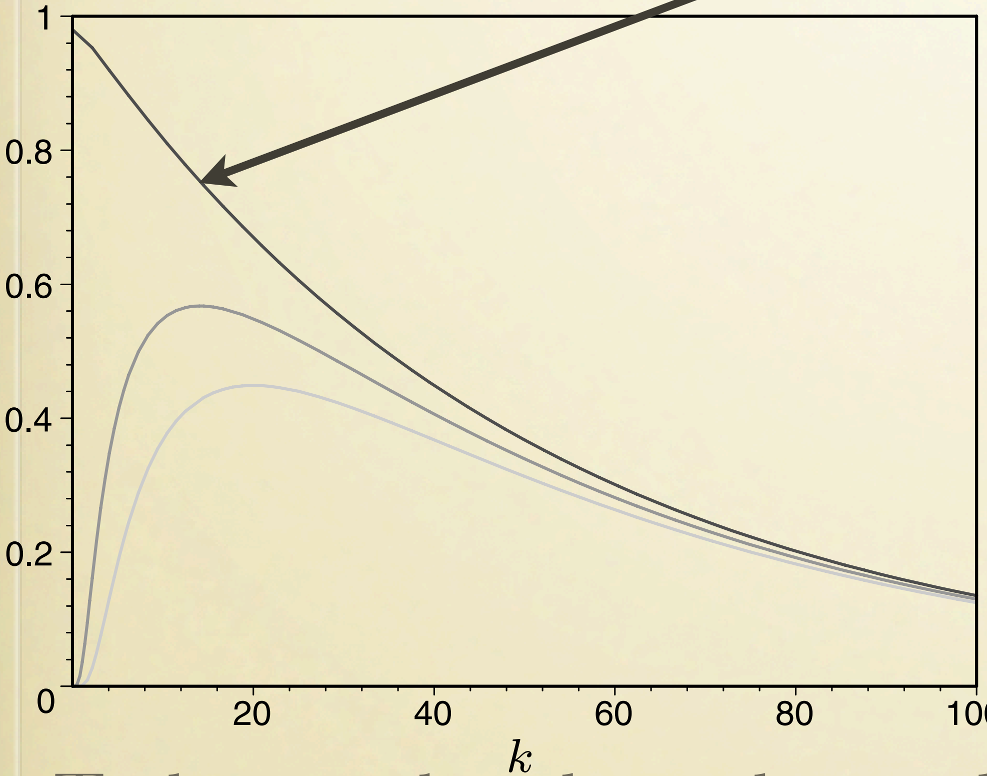

Taylor zeroth order and second order

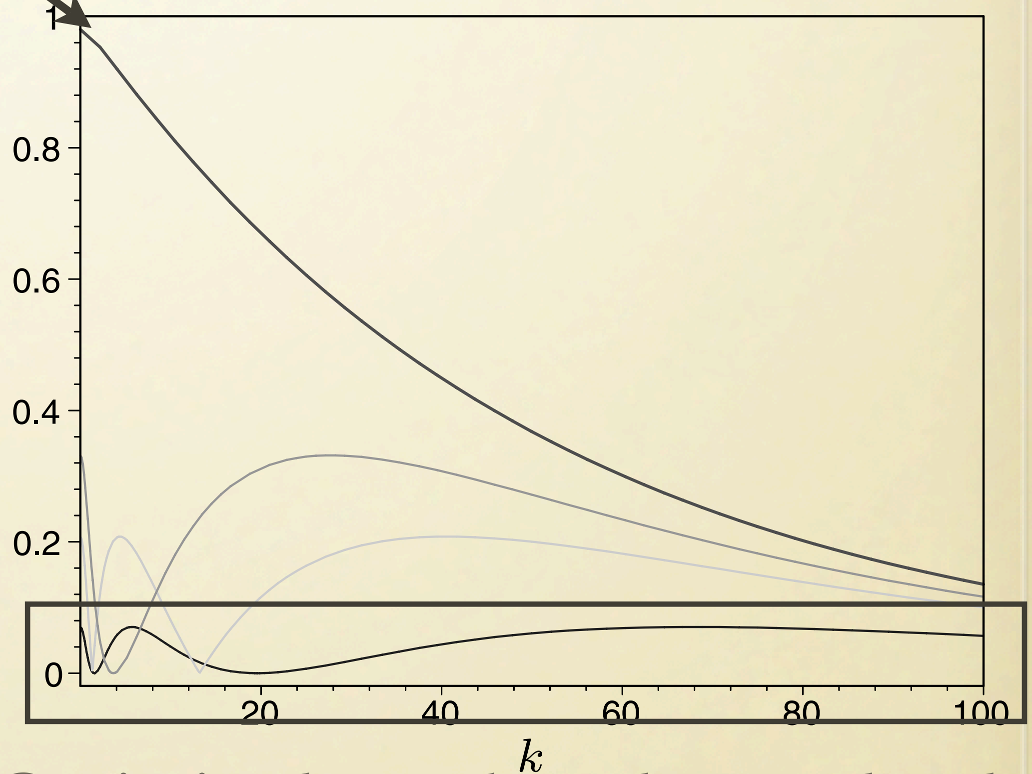



Optimized zeroth and second order with two-sided zeroth order

# Convergence rates

Classical Schwarz




Taylor zeroth order and second order

Optimized zeroth and second order with two-sided zeroth order

# Optimized Schwarz: algebraic results

- SGT 2007: show how to modify existing Schwarz algorithm to yield optimized versions
- The augmented or "enhanced" system is rediscovered
- Spectral elements are natural candidates:
  - Overlapping grids are cumbersome to construct
  - Block preconditioning costly: FDM when possible
  - Optimal preconditioner is known (SD Kim 2006)
  - Q1-GLL based problem costly to invert does not scale: use MG or other solver (opt Schwarz?)to invert

## Optimized Schwarz: algebraic results

Inverting:
$$\mathbf{u}_j^{n+1} = \tilde{A}_j^{-1}\mathbf{f}_j + \tilde{A}_j^{-1}\sum_{k=1}^{J}\tilde{B}_{jk}\mathbf{u}_k^n$$

At convergence:
$$(I - \tilde{A}_j^{-1}\sum_{k=1}^{J}\tilde{B}_{jk})\mathbf{u}_k = \tilde{A}_j^{-1}\mathbf{f}_j$$

Apply restriction extension operators:
$$\{I - \sum_{j,k=1}^{J}\tilde{R}_j^T\tilde{A}_j^{-1}\tilde{B}_{jk}R_k\}\mathbf{u} = \sum_{j=1}^{J}\tilde{R}_j^T\tilde{A}_j^{-1}R_j\mathbf{f}$$

MxV operation:
$$M^{-1}\bar{A}\mathbf{u} = M^{-1}\bar{\mathbf{f}}$$

# Optimized Schwarz: algebraic results

Inverting: $$\mathbf{u}_j^{n+1} = \tilde{A}_j^{-1}\mathbf{f}_j + \tilde{A}_j^{-1}\sum_{k=1}^{J}\tilde{B}_{jk}\mathbf{u}_k^n$$

At convergence: $$(I - \tilde{A}_j^{-1}\sum_{k=1}^{J}\tilde{B}_{jk})\mathbf{u}_k = \tilde{A}_j^{-1}\mathbf{f}_j$$

Apply restriction extension operators:

$$\{I - \sum_{j,k=1}^{J}\tilde{R}_j^T\tilde{A}_j^{-1}\tilde{B}_{jk}R_k\}\mathbf{u} = \sum_{j=1}^{J}\tilde{R}_j^T\tilde{A}_j^{-1}R_j\mathbf{f}$$

MxV operation: $$M^{-1}\bar{A}\mathbf{u} = M^{-1}\bar{\mathbf{f}}$$

## Optimized Schwarz: algebraic results

Inverting: $$\mathbf{u}_j^{n+1} = \tilde{A}_j^{-1}\mathbf{f}_j + \tilde{A}_j^{-1}\sum_{k=1}^{J}\tilde{B}_{jk}\mathbf{u}_k^n$$

At convergence: $$(I - \tilde{A}_j^{-1}\sum_{k=1}^{J}\tilde{B}_{jk})\mathbf{u}_k = \tilde{A}_j^{-1}\mathbf{f}_j$$

Apply restriction extension operators:

$$\{I - \sum_{j,k=1}^{J}\tilde{R}_j^T\tilde{A}_j^{-1}\tilde{B}_{jk}R_k\}\mathbf{u} = \sum_{j=1}^{J}\tilde{R}_j^T\tilde{A}_j^{-1}R_j\mathbf{f}$$

MxV operation: $$M^{-1}\bar{A}\mathbf{u} = M^{-1}\bar{\mathbf{f}}$$

# Optimizing optimized Schwarz!!

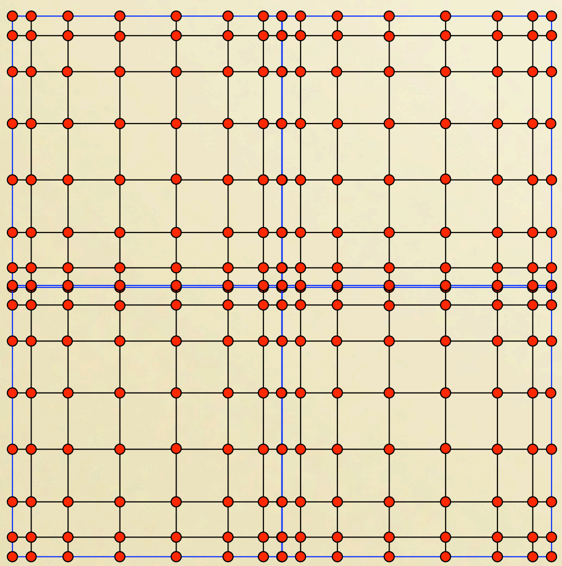
- Schwarz for SEM: efficient implementation (Fischer 97,+Miller and Tufo 98, + Tufo 99) (3D)
- Constraints imposed by new architectures
- Loosing symmetry: 2 MxV instead of 1
- OS no overlapping region to construct
- FDM lost?

# Optimizing for cache

$$\mathbf{z}_j = \sum_{k=1}^{J} \tilde{B}_{jk} \mathbf{u}_k^n$$

Suppose a non-overlapping domain in 2D: $N^2$ unknowns

$\mathbf{z}_j$

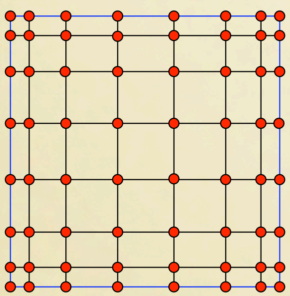


# Optimizing for cache

$$\mathbf{z}_j = \sum_{k=1}^{J} \tilde{B}_{jk}\mathbf{u}_k^n$$

Suppose a non-overlapping domain in 2D: $N^2$ unknowns

$\mathbf{z}_j$



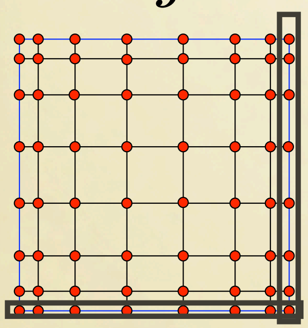
# Optimizing for cache

$$\mathbf{z}_j = \sum_{k=1}^{J} \tilde{B}_{jk}\mathbf{u}_k^n$$

Suppose a non-overlapping domain in 2D: $N^2$ unknowns

$\mathbf{z}_j$

$2N$ Boundary nodes

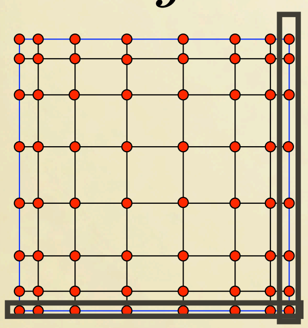


# Optimizing for cache

$$\mathbf{z}_j = \sum_{k=1}^{J} \tilde{B}_{jk} \mathbf{u}_k^n$$

Suppose a non-overlapping domain in 2D: $N^2$ unknowns

$\mathbf{z}_j$



$2N$ Boundary nodes

$$\tilde{A}_j^{-1} \mathbf{z}_j = \tilde{A}_j^{-1} \sum_{k=1}^{J} \tilde{B}_{jk} \mathbf{u}_k^n$$

| | Normal | Optimized |
|---|---|---|
| Size: | $N^2 \times N^2$ | $N^2 \times 2N$ |
| MxV cost: | $O(N^4)$ | $O(N^3)$ |

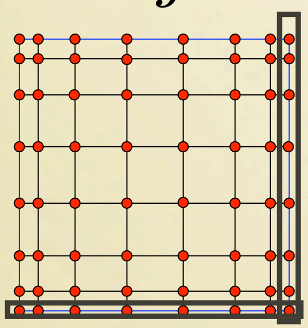
Rectangular

# Optimizing for cache

$$\mathbf{z}_j = \sum_{k=1}^{J} \tilde{B}_{jk} \mathbf{u}_k^n$$

Suppose a non-overlapping domain in 2D: $N^2$ unknowns

$\mathbf{z}_j$



$2N$ Boundary nodes

$$\tilde{A}_j^{-1} \mathbf{z}_j = \tilde{A}_j^{-1} \sum_{k=1}^{J} \tilde{B}_{jk} \mathbf{u}_k^n$$

| | Normal | Optimized |
|---|---|---|
| Size: | $N^2 \times N^2$ | $N^2 \times 2N$ |
| MxV cost: | $O(N^4)$ | $O(N^3)$ |

Rectangular

Cost identical to 2D FDM or "interface" system approach

# Creating the augmented system from a weak form

- The normal derivative can be written in terms of the original bilinear operator (Toselli, Widlund 2005)
- Avoids the difficult duality pairing for functions on the edges of the subdomains

$$
\begin{aligned}
T_j(w_j^{k+1}, \phi_j) &= \int_{\Omega_j} \nabla \phi_j \cdot \nabla w_j^{k+1} + \int_{\Omega_j} \phi_j \Delta w_j^{k+1} \\
&= \int_{\Omega_j} \nabla \phi_j \cdot \nabla w_j^{k+1} + \int_{\Omega_j} \phi_j w_j^{k+1} - f_j(\phi_j) \\
&= a_j(w_j^{k+1}, \phi_j) - f_j(\phi_j)
\end{aligned}
$$

Where we pick $\phi_j \in H^1(\partial \Omega_j)$

# Creating the augmented system from a weak form

Boundary condition is:

$$
\begin{aligned}
T_j(w_j^{k+1}, \phi_j) &= \sum_{l \in \mathcal{N}(\Omega_j)} T_j(w_j^{k+1}, \phi_j|_{\Gamma_{jl}}) \\
&= \sum_{l \in \mathcal{N}(\Omega_j)} \{\int_{\Gamma_{jl}} p\phi_j(w_l^k - w_l^{k+1}) - T_l(w_l^k, \phi_j|_{\Gamma_{jl}})\} \\
&= -\int_{\Omega_j} p\phi_j w_j^{k+1} + \sum_{l \in \mathcal{N}(\Omega_j)} \{\int_{\Gamma_{jl}} p\phi_j w_l^k - T_l(w_l^k, \phi_j|_{\Gamma_{jl}})\}
\end{aligned}
$$

where a sum on neighbors appears.

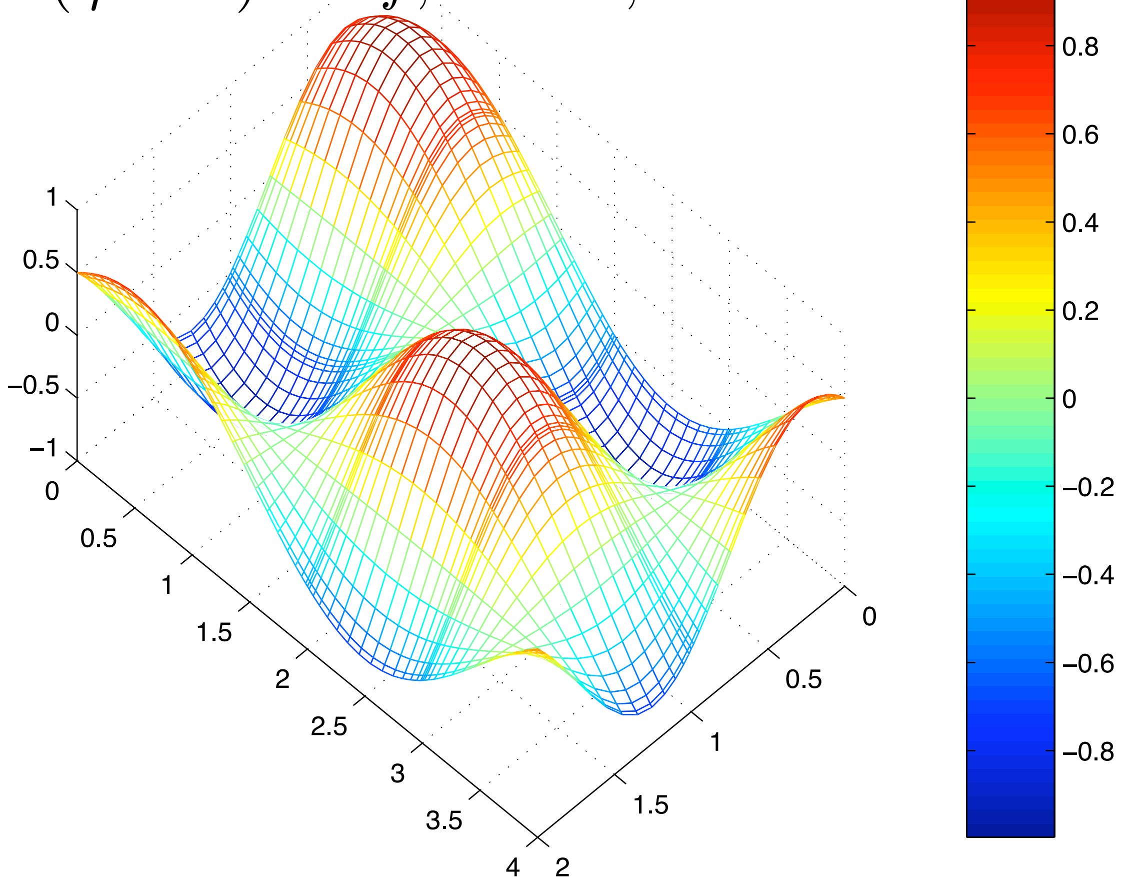
Leads to the relaxed form required by the algorithm

$$
\boxed{
\begin{aligned}
a_j^h(u_j^{n+1}, \phi_j) + \sum_{l \in \mathcal{N}(\Omega_j)} < \phi_j, T(u_j^{n+1}, p, q, \tau) > |_{\Gamma_{jl}} &= f_j^h(\phi_j) + \sum_{l \in \mathcal{N}(\Omega_j)} f_l^h(\phi_l|_{\Gamma_{jl}}) \\
&- \sum_{l \in \mathcal{N}(\Omega_j)} a_l^h(u_l^n, \phi_l|_{\Gamma_{jl}}) + \sum_{l \in \mathcal{N}(\Omega_j)} < \phi_l, T(u_l^n, p, q, \tau) > |_{\Gamma_{jl}} \qquad (
\end{aligned}
}
$$

# SEM simple problem

$$\mathcal{L}u = (\eta - \Delta)u = f, \quad \text{in } \Omega,$$

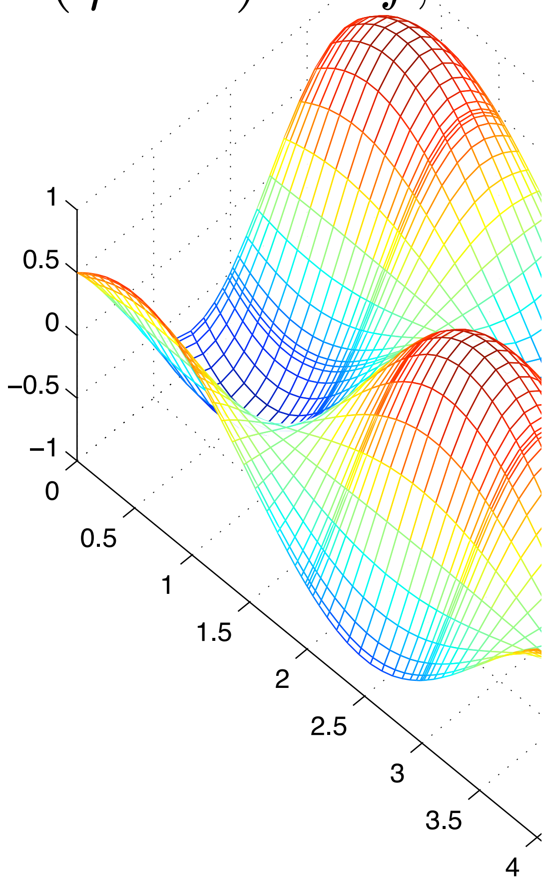
Gander 2006

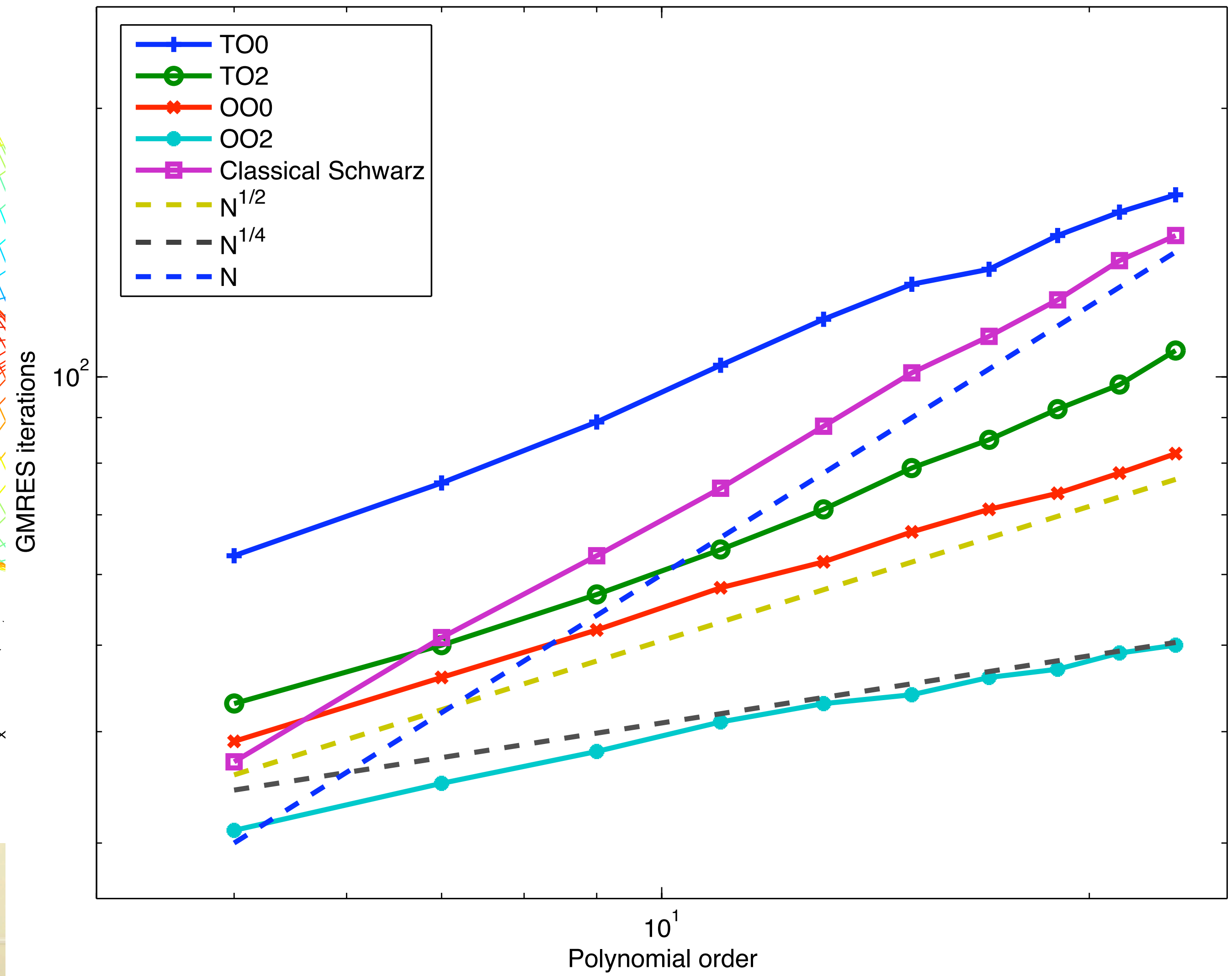


# SEM simple problem

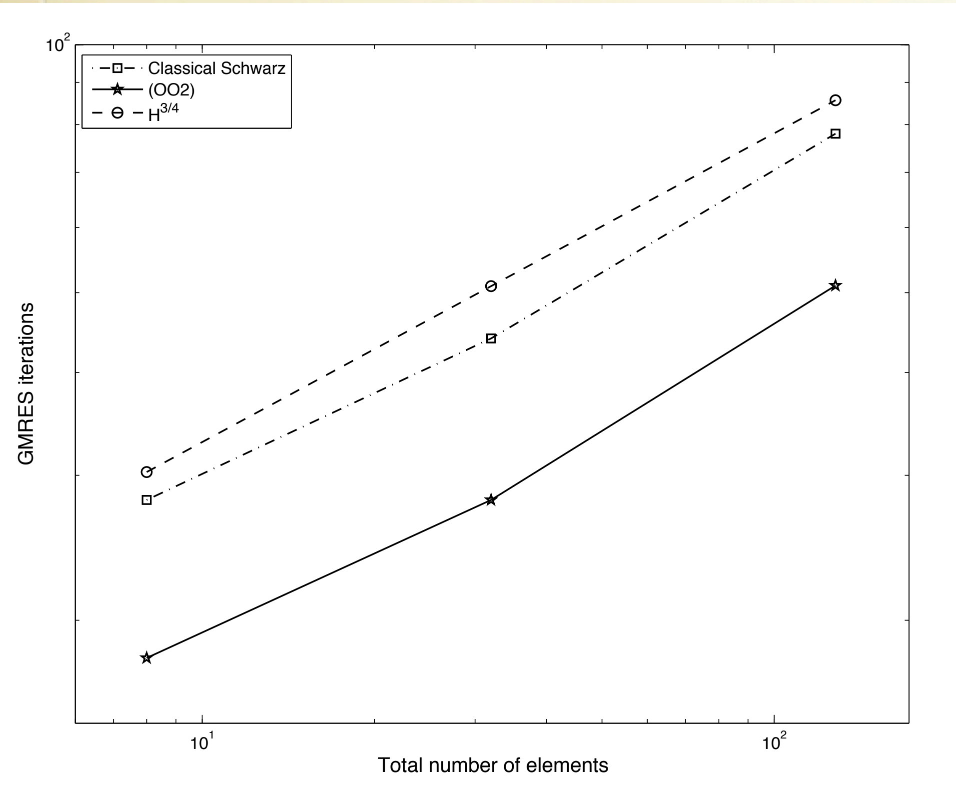
$$\mathcal{L}u = (\eta - \Delta)u = f, \quad \text{in } \Omega,$$

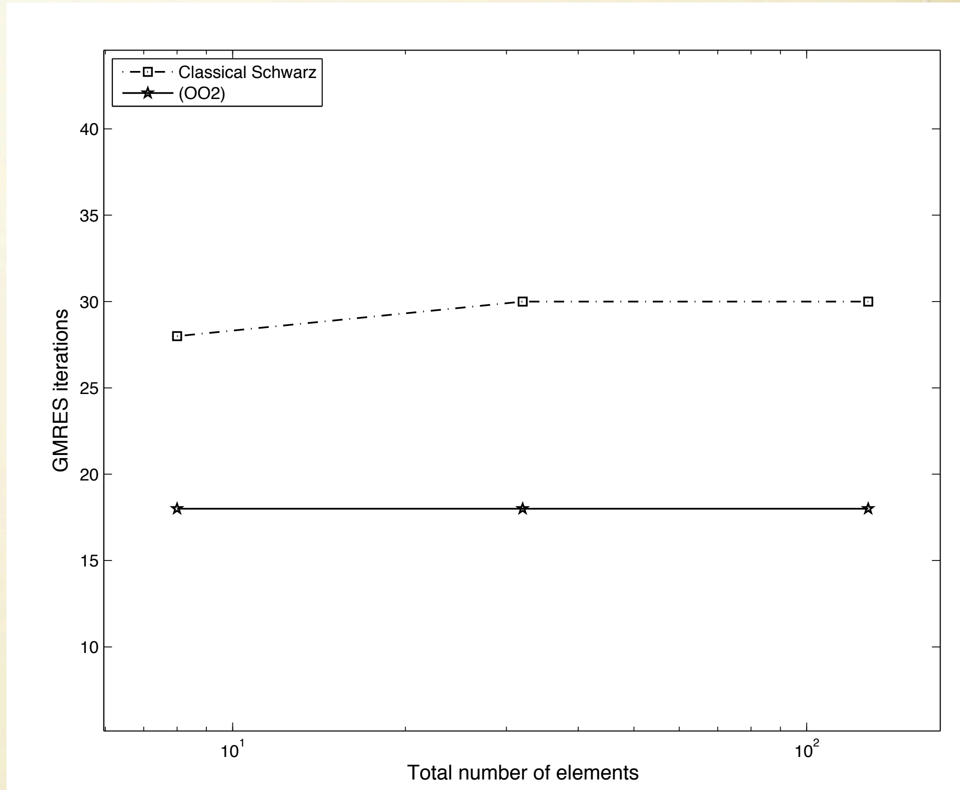
Gander 2006



# Coarse solver?




Qin and Xu 2006 SINUM

Scaling $\eta$ as $\frac{J}{4}(\frac{N}{4})^4$

# Primitive equations

Momentum: $$\frac{d\mathbf{v}}{dt} + f\,\mathbf{k} \times \mathbf{v} + \nabla\Phi + R\,T\,\nabla \ln p = 0$$

Thermodynamic: $$\frac{d\,T}{d\,t} - \frac{\kappa\,T\,\omega}{p} = 0$$

Continuity: $$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial \eta}\right) + \nabla \cdot \left(\mathbf{v}\,\frac{\partial p}{\partial \eta}\right) + \frac{\partial}{\partial \eta}\left(\dot{\eta}\,\frac{\partial p}{\partial \eta}\right) = 0$$

HOMME: high order multiscale modeling environment

# Primitive equations: SI

Hydrostatic assumption: $$\frac{\partial \Phi}{\partial \eta} = -\frac{R\,T}{p}\frac{\partial p}{\partial \eta}.$$

Linearization (barotropic state): $T^r = 300K, \;\; p_s^r = 1000hPa$

Semi-Implicit:

$$\frac{dX}{dt} = \mathcal{M}(X)$$

Add zero: $$\frac{dX}{dt} = \mathcal{M}(X) + \mathcal{L}X - \mathcal{L}X = \mathcal{N}(X) - \mathcal{L}X$$

$$\frac{X^{n+1} - X^{n-1}}{2\Delta t} = \mathcal{N}(X^n) - \frac{1}{2}\mathcal{L}(X^{n+1} + X^{n-1}) = \mathcal{M}(X^n) + \mathcal{L}X^n - \frac{1}{2}\mathcal{L}(X^{n+1} + X^{n-1})$$

$$\frac{X^{n+1} - X^{n-1}}{2\Delta t} = \mathcal{M}(X^n) - \frac{1}{2}\Delta_{tt}\mathcal{L}X$$

"Time diffusion"

# PE: VERTICAL STRUCTURE MATRIX

Results of hydrostatic assumption and vertical coordinate choice:

$$p(\eta, p_s) = A(\eta)p_0 + B(\eta)p_s$$

$$\mathbf{A} = R\mathbf{H}^r\mathbf{T} + RT^r P,$$

$$G^r - \Delta t^2 \mathbf{A}\nabla^2 G^r = B - \Delta t \mathbf{A}\nabla \cdot \mathcal{V}$$

Solve for each k:

$$\left(\nabla^2 - \frac{1}{\Delta t^2 \lambda_k}\right)\Gamma_k^r = C_k$$

Time dependance

Series of 2D Helmholtz

Barotropic eigenmodes of atmosphere

Backsub:

$$\begin{aligned} D &= \Delta t^{-1}\,\mathbf{A}^{-1}\,(B - G^r) \\ \ln p_s &= \mathcal{P} - \Delta t\, P \cdot D \\ T &= \mathcal{T} - \Delta t\,\mathbf{T}D \\ \mathbf{v} &= \mathcal{V} - \Delta t\,\nabla G^r \end{aligned}$$

(Thomas and Loft 2005)

# PE: VERTICAL STRUCTURE MATRIX

Results of hydrostatic assumption and vertical coordinate choice:

$$p(\eta, p_s) = A(\eta) p_0 + B(\eta) p_s$$

$$\mathbf{A} = R\mathbf{H}^r\mathbf{T} + RT^r P,$$ Diagonalize

$$G^r - \Delta t^2 \mathbf{A} \nabla^2 G^r = B - \Delta t \mathbf{A} \nabla \cdot \mathcal{V}$$

Solve for each k:

$$\left( \nabla^2 - \frac{1}{\Delta t^2 \lambda_k} \right) \Gamma_k^r = C_k$$

Time dependance

Series of 2D Helmholtz

Barotropic eigenmodes of atmosphere

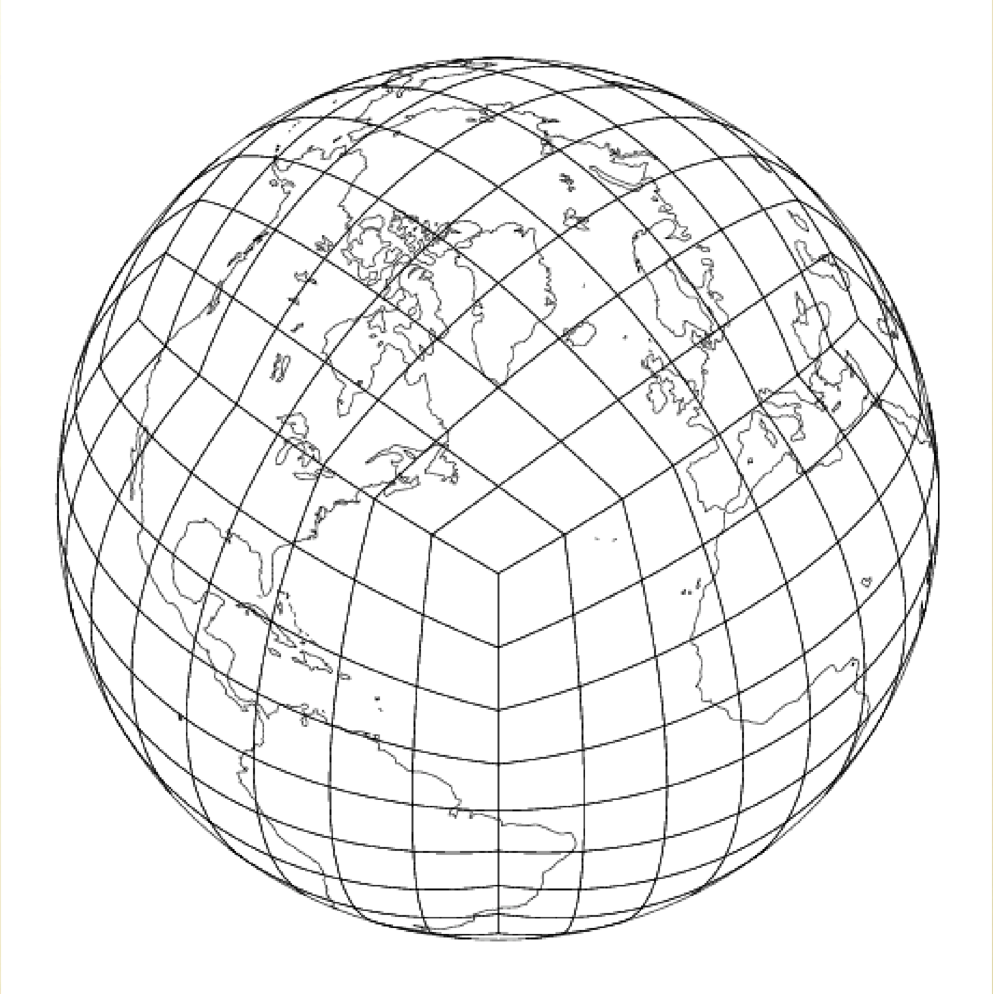
Backsub:

$$\begin{aligned} D &= \Delta t^{-1}\,\mathbf{A}^{-1}(B - G^r) \\ \ln p_s &= \mathcal{P} - \Delta t\, P \cdot D \\ T &= \mathcal{T} - \Delta t\,\mathbf{T} D \\ \mathbf{v} &= \mathcal{V} - \Delta t\, \nabla G^r \end{aligned}$$

(Thomas and Loft 2005)

# Cubed sphere

- Equiangular projection
- Sadourny (72), Rancic (96), Ronchi (96)
- Most models moving towards this approach
- SFC: Dennis 2003



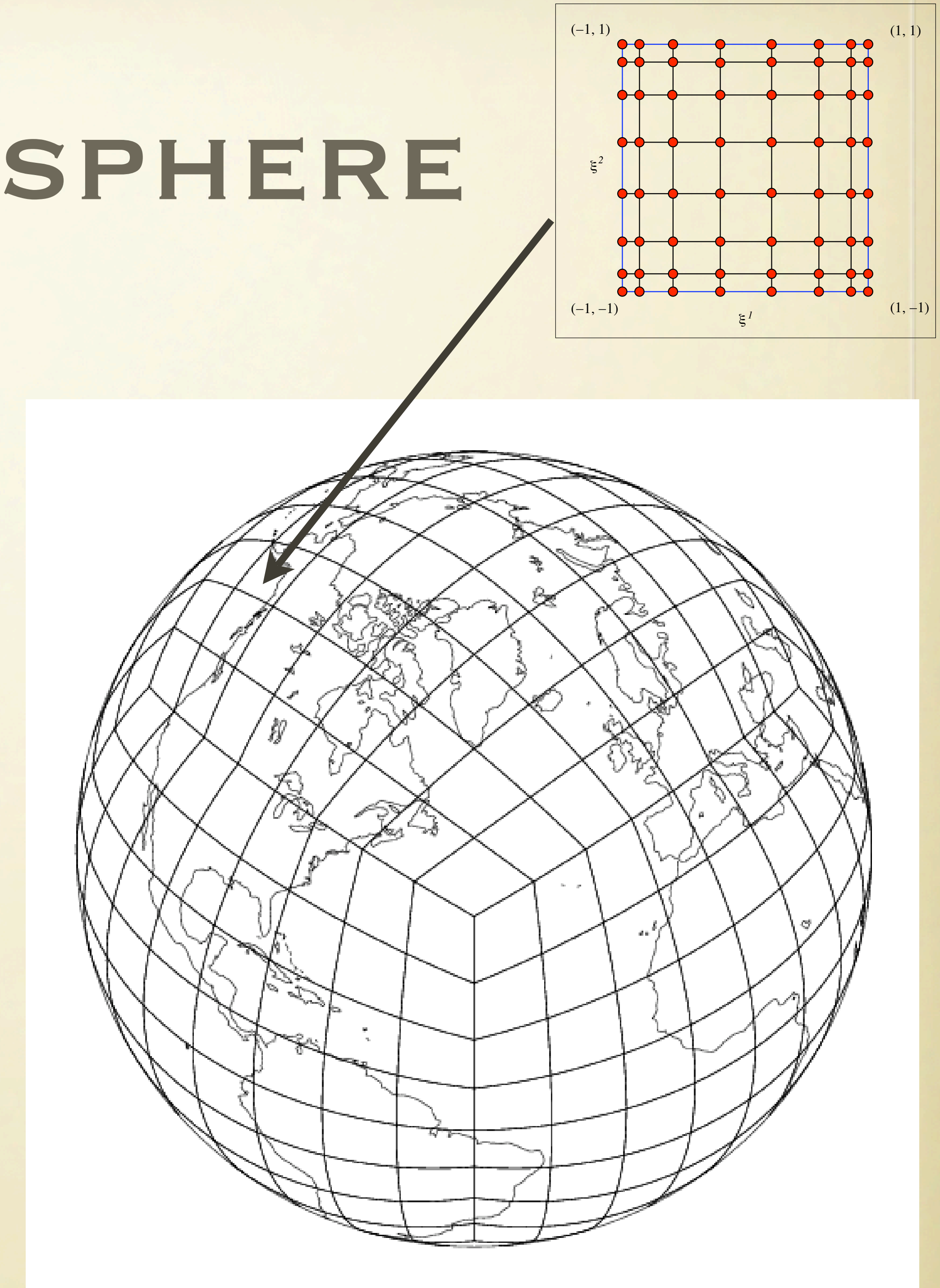
Metric tensor

$$g_{ij} = \frac{1}{r^4 \cos^2 x_1 \cos^2 x_2} \begin{bmatrix} 1 + \tan^2 x_1 & -\tan x_1 \tan x_2 \\ -\tan x_1 \tan x_2 & 1 + \tan^2 x_2 \end{bmatrix}.$$

Rewrite div and vorticity

$$g\, \nabla \cdot \mathbf{v} = \frac{\partial}{\partial x^j} \left( g\, u^j \right), \quad g\, \zeta = \epsilon_{ij} \frac{\partial u_j}{\partial x^i}.$$

# Cubed sphere




- Equiangular projection
- Sadourny (72), Rancic (96), Ronchi (96)
- Most models moving towards this approach
- SFC: Dennis 2003

Metric tensor

$$g_{ij} = \frac{1}{r^4 \cos^2 x_1 \cos^2 x_2} \begin{bmatrix} 1+\tan^2 x_1 & -\tan x_1 \tan x_2 \\ -\tan x_1 \tan x_2 & 1+\tan^2 x_2 \end{bmatrix}.$$

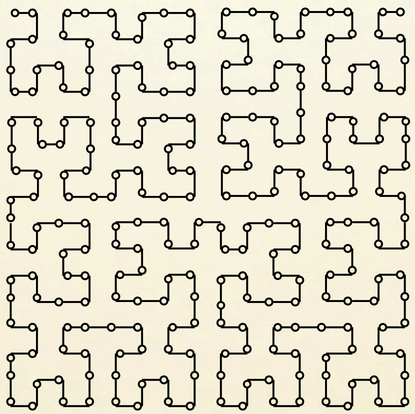
Rewrite div and vorticity

$$g\,\nabla\cdot\mathbf{v} = \frac{\partial}{\partial x^j}\,(\,g\,u^j\,), \quad g\,\zeta = \epsilon_{ij}\,\frac{\partial u_j}{\partial x^i}.$$

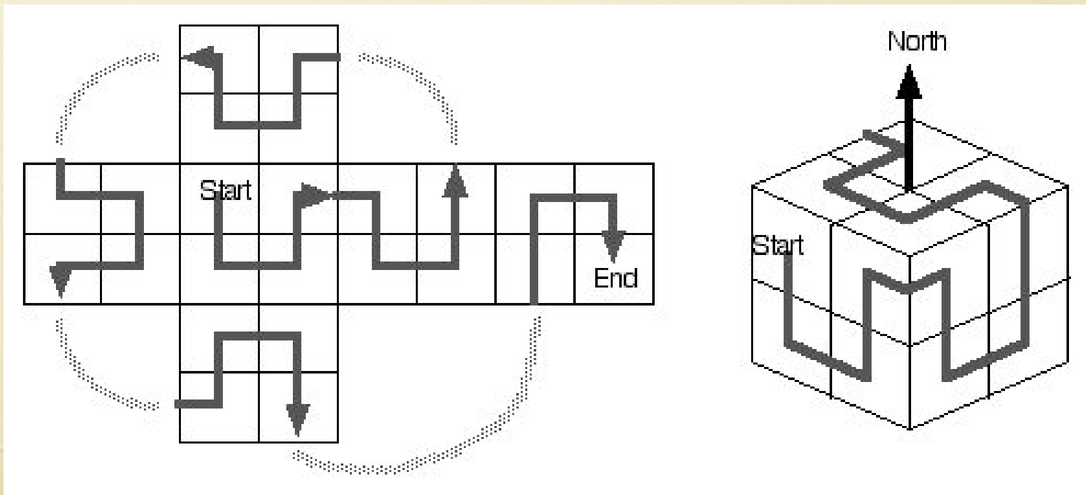
# Mesh partitioning

Space filling curves (Dennis 2003):

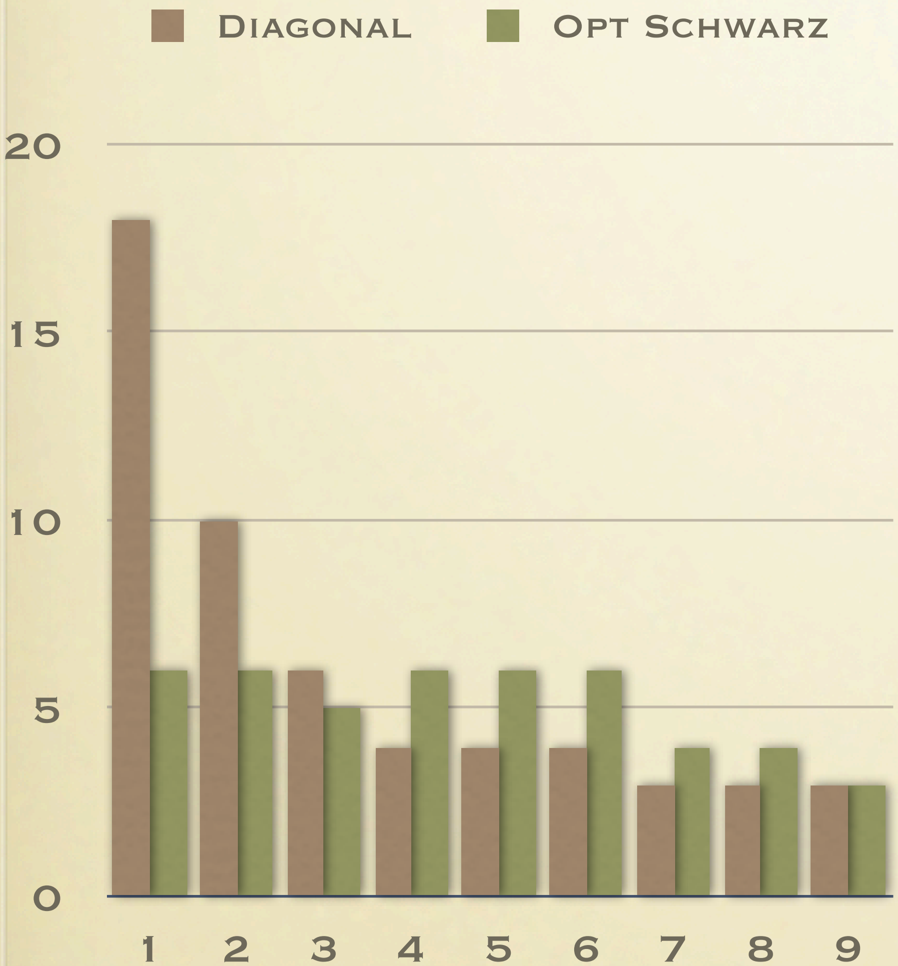
Hilbert SFC:



Cubed sphere:

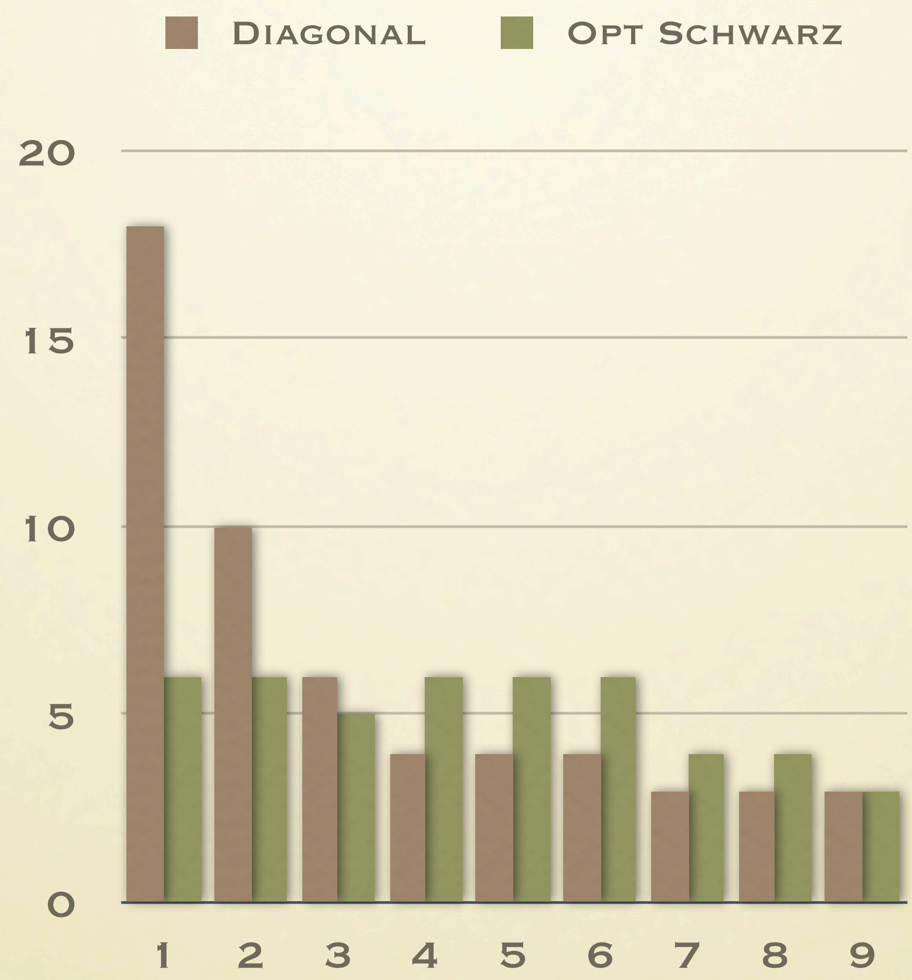


# Convergence per mode

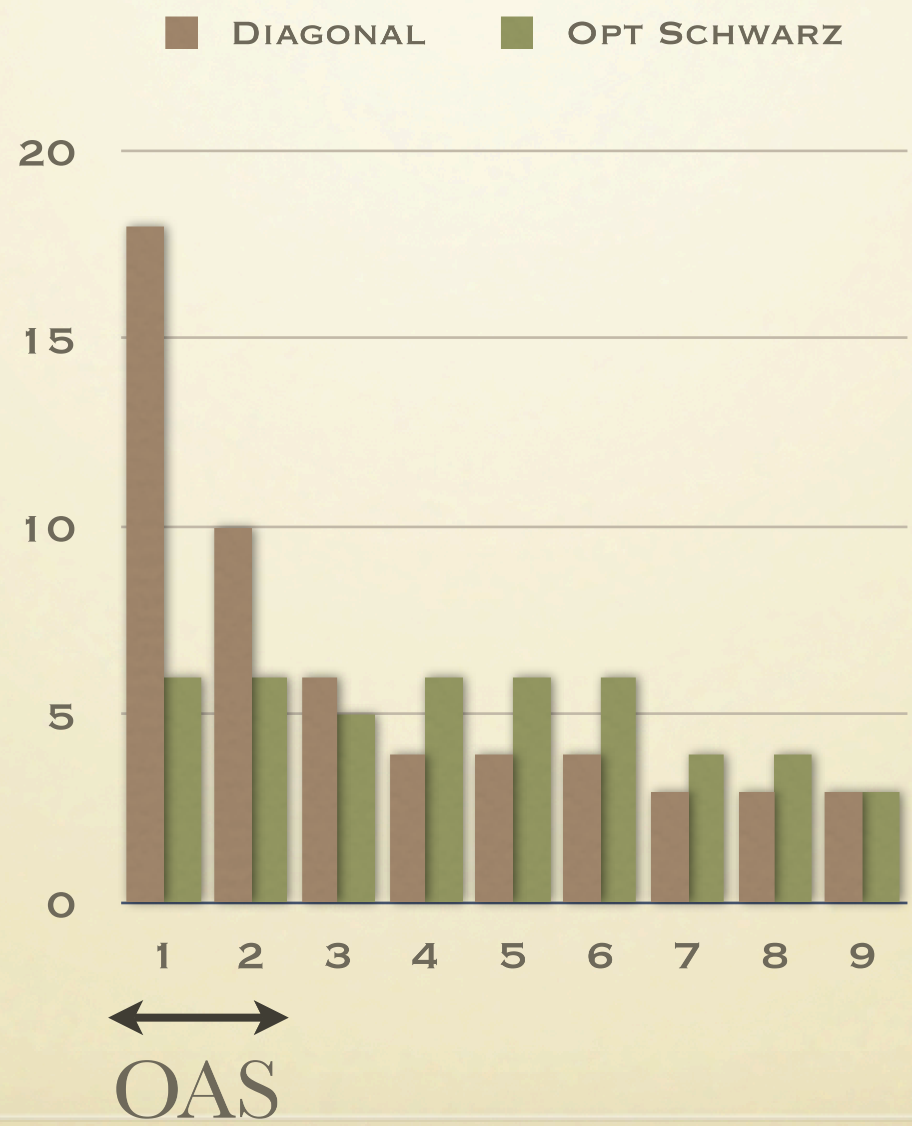



- Communication cost identical
- Twice the cost of CG per iteration
- Diagonal O(N) while OS is $O(N^3)$
- Best strategy: use OS on first few barotropic modes and diagonal elsewhere
- No coarse solver needed: because of time dependance

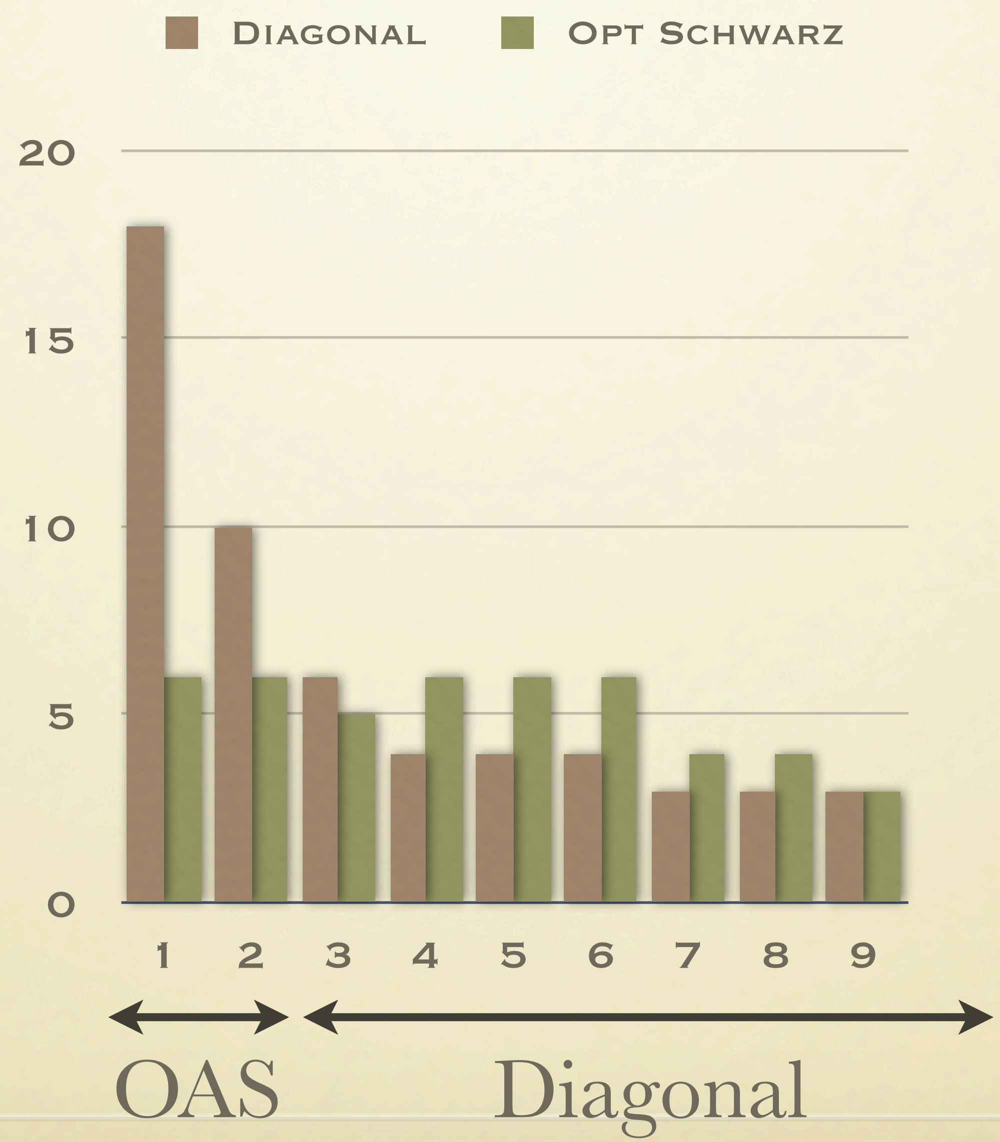
# New approach



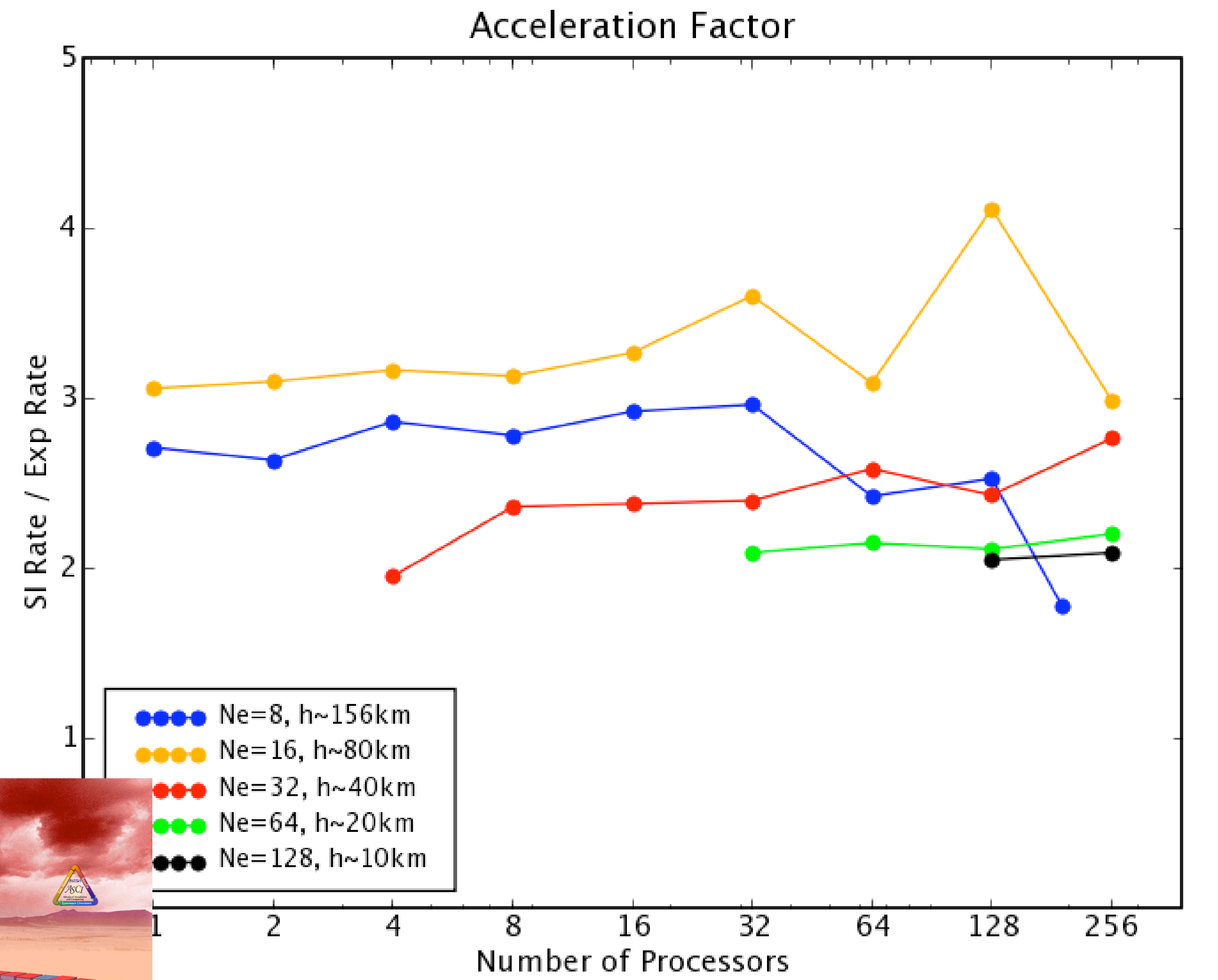
# New approach

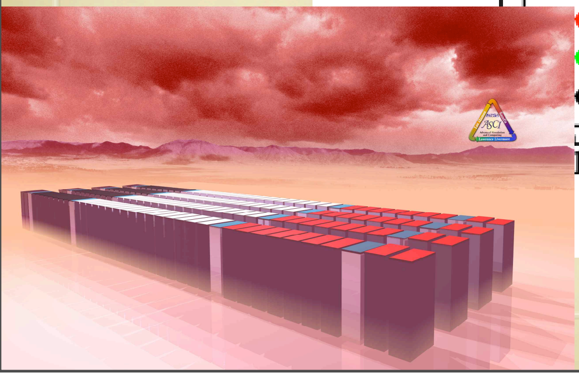


# New approach



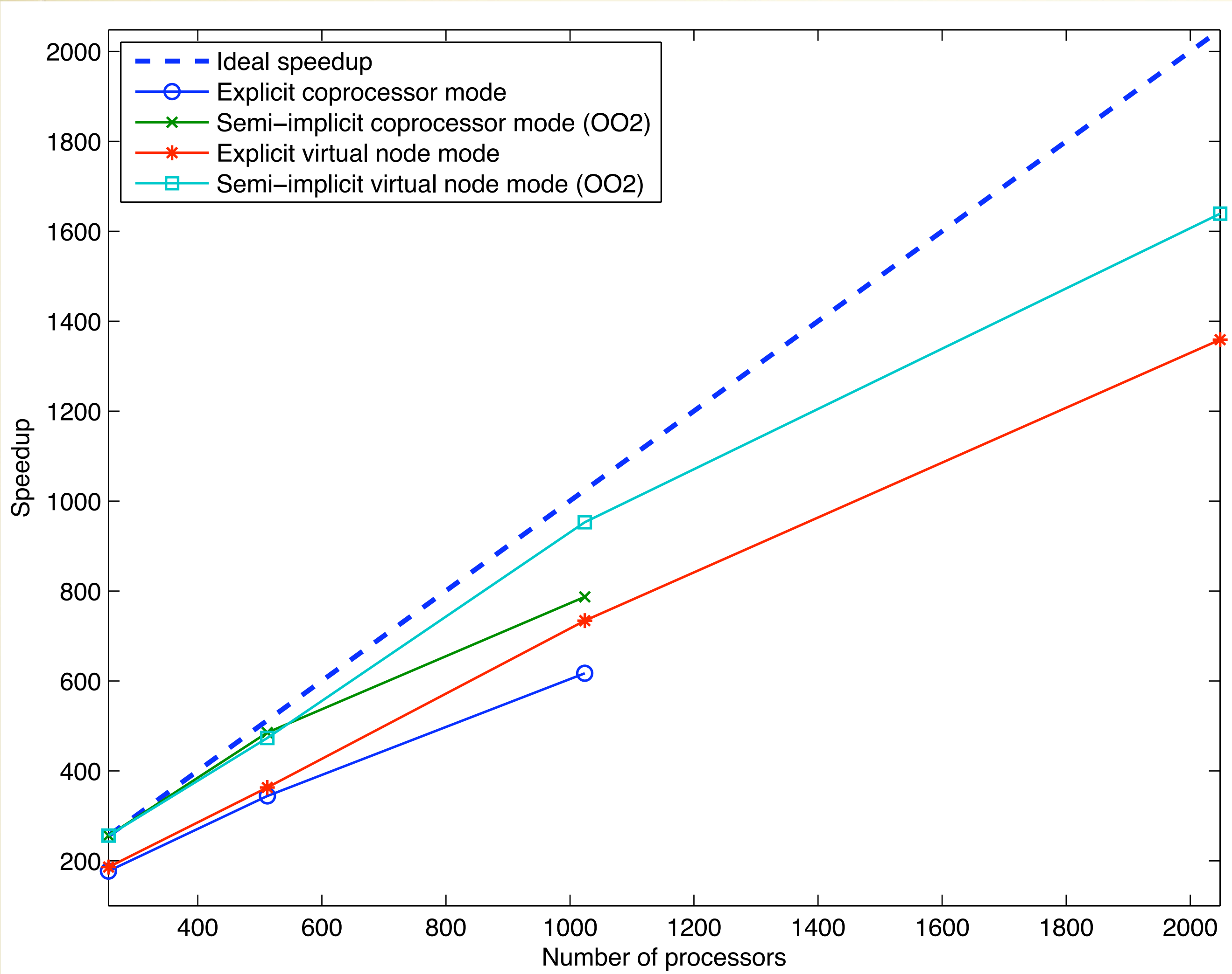
# Si vs Exp: Red Storm



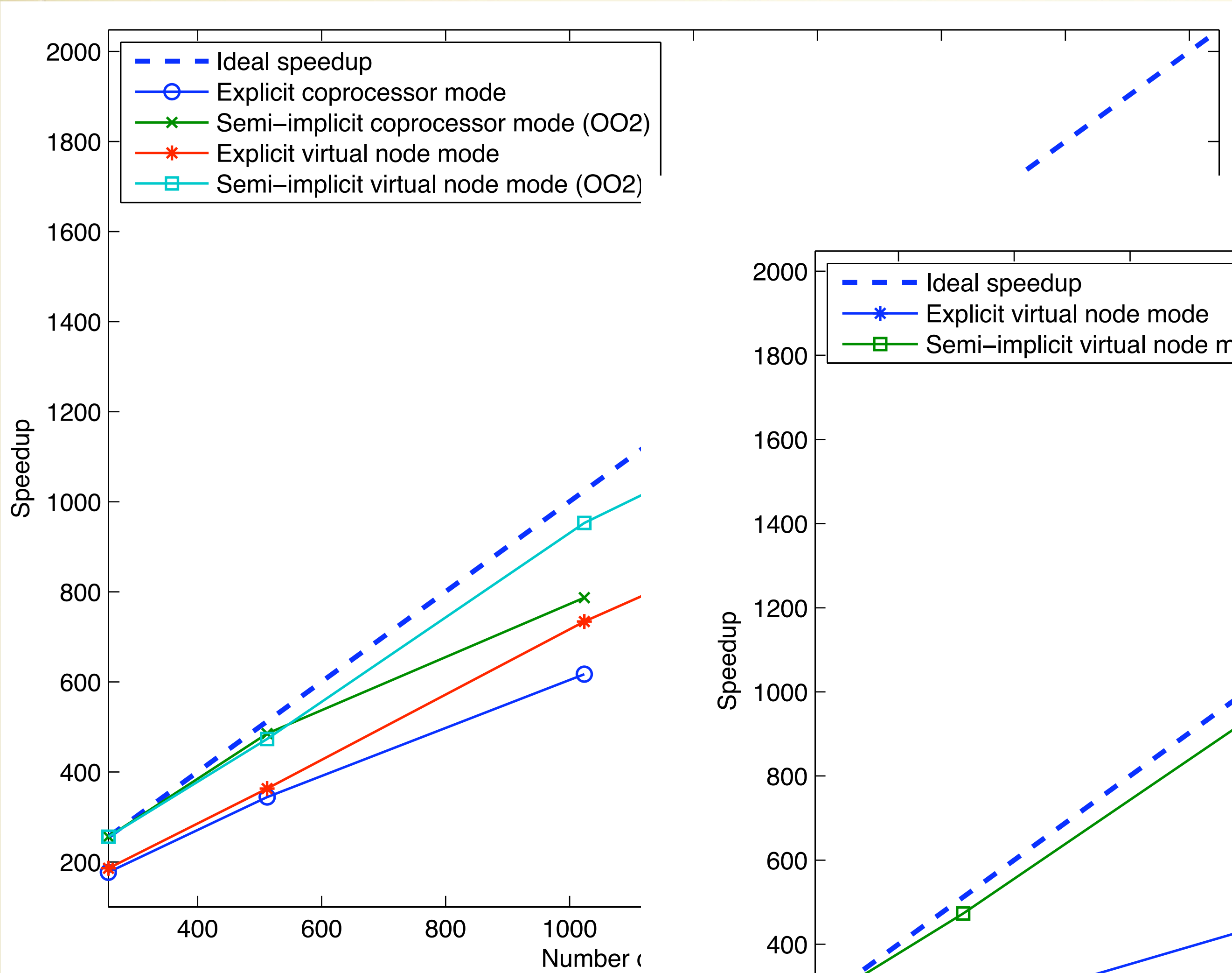
W. Spotz : Sandia National Labs

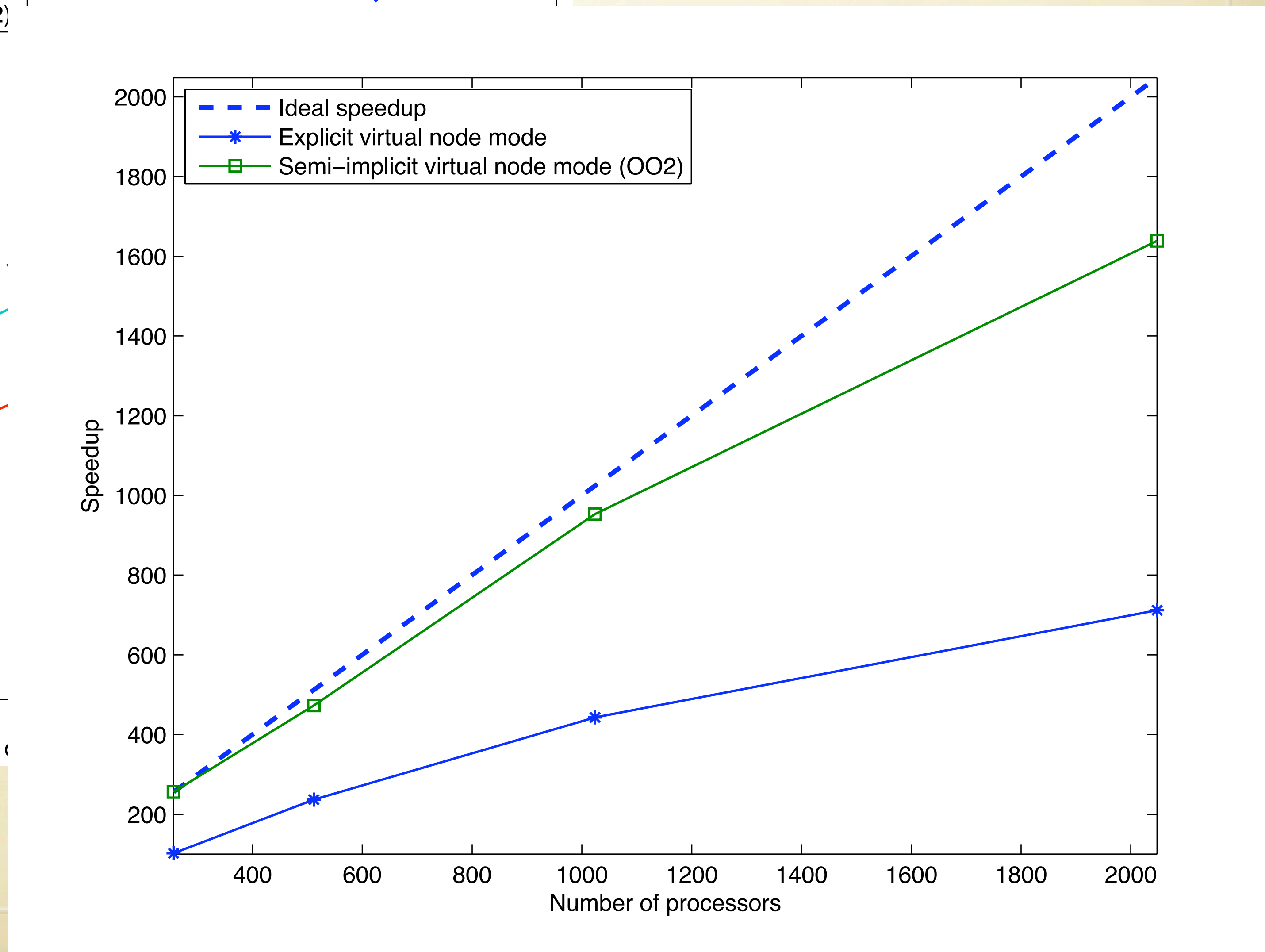
# Si vs Exp: Blue gene




ne=32, 40km

# Si vs Exp: Blue gene




ne=32,
40km

# Bite the Bullet

# Stratified Compressible Euler

Conservation law: $\underline{\mathrm{U}}_t + \nabla \cdot \mathbf{F}(\underline{\mathrm{U}}) = S(\underline{\mathrm{U}})$

With:

$$
\begin{aligned}
\underline{\mathrm{U}} &\equiv (\rho, \rho u, \rho w, \Theta)^T = (\rho, U, W, \Theta)^T \\
\mathbf{F}(\underline{\mathrm{U}}) &\equiv (F, G) \\
F &= (U, \frac{UU}{\rho} + p, \frac{WU}{\rho}, \frac{U\Theta}{\rho})^T \\
G &= (W, \frac{UW}{\rho}, \frac{WW}{\rho} + p, \frac{W\Theta}{\rho})^T \\
S(\underline{\mathrm{U}}) &= (0, 0, -g\rho, 0)^T \\
p &= p_0(\frac{R\Theta}{p_0})^\gamma
\end{aligned}
$$

Remove hydrostatic state:

$$
\begin{aligned}
p =& \bar{p}(z) + p' \\
\rho =& \bar{\rho} + \rho' \\
\Theta =& \bar{\rho}(z)\bar{\theta}(z) + \Theta'
\end{aligned}
$$

# DG formulation:

Integrate over control volume $\Omega_k$ using:

$$\int_{\Omega_k} \varphi_h \frac{d\underline{\mathrm{U}}_h}{dt}\, d\Omega = \int_{\Omega_k} \varphi_h S(\underline{\mathrm{U}}_h)\, d\Omega + \int_{\Omega_k} \mathbf{F}(\underline{\mathrm{U}}_h) \cdot \nabla \varphi_h \, d\Omega - \int_{\partial\Omega_k} \varphi_h \mathbf{F} \cdot \hat{n}\, ds$$

Lax-Friedrich numerical flux:

$$\widehat{\mathbf{F}}(\underline{\mathrm{U}}_h^+, \underline{\mathrm{U}}_h^-) \cdot \hat{n} = \frac{1}{2}\left[(\mathbf{F}(\underline{\mathrm{U}}_h^+) + \mathbf{F}(\underline{\mathrm{U}}_h^-)) \cdot \hat{n} - \alpha(\underline{\mathrm{U}}_h^+ - \underline{\mathrm{U}}_h^-)\right]$$

Galerkin based on GLL points + <u>exact integration</u>

$$u_h^k = \sum_{i=0}^{N} \sum_{j=0}^{N} u_{ij} h_i(x) h_j(y)$$

Leads to semi-discrete problem solved using ROW:

$$\frac{d\underline{\mathrm{U}}_h}{dt} = L_h(\underline{\mathrm{U}}_h)$$

Filtering is applied: Boyd-Vandeven

# Cheap Implicitness: Rosenbrock methods

- The Jacobian matrix is included in the DIRK order conditions
- Each stage requires solution to <u>linear problem only</u>
- Viewed as one Newton iteration per RK stage
- Used for stiff chemical reaction problems in the geosciences with success
- Used traditionally for parabolic PDEs

# Rosenbrock

Start with DIRK:

$$k_i = \Delta t f(u^n + \sum_{j=1}^{i-1} a_{ij} k_j + a_{ii} k_i)$$

$$u^{n+1} = u^n + \sum_{i=1}^{s} b_i k_i$$

LDIRK:

$$k_i = \Delta t f(g_i) + \Delta t \left.\frac{\partial f}{\partial u}\right|_{g_i} a_{ii} k_i$$

Linearize around: $u^n + \sum_{j=1}^{i-1} a_{ij} k_j$

$$g_i = u^n + \sum_{j=1}^{i-1} a_{ij} k_j$$

$$u^{n+1} = u^n + \sum_{i=1}^{s} b_i k_i$$

Replace Jacobian at $g_i$ with $J \equiv \left.\frac{\partial f}{\partial u}\right|_{u^n}$

$$k_i = \Delta t f(u^n + \sum_{j=1}^{i-1} \alpha_{ij} k_j) + \Delta t J \sum_{j=1}^{i} \gamma_{ij} k_j$$

Rosenbrock:

$$u^{n+1} = u^n + \sum_{i=1}^{s} b_i k_i$$

MxV products

# Avoiding multiplications

Suppose $h_i = \sum_{j=1}^{i} \gamma_{ij} k_j$ then $k_i = \frac{1}{\gamma_{ii}} h_i - \sum_{j=1}^{i-1} c_{ij} h_j$

The modified Rosenbrock is

$$(\frac{1}{\Delta t \gamma_{ii}} - J) h_i = f(u^n + \sum_{j=1}^{i-1} a_{ij} h_j) + \sum_{j=1}^{i-1} (\frac{c_{ij}}{\Delta t}) h_j$$

$$u^{n+1} = u^n + \sum_{j=1}^{s} m_j h_j \qquad \hat{u}^{n+1} = \hat{u}^n + \sum_{j=1}^{s} \hat{m}_j h_j$$

Free error estimation!

Where: $\Gamma = (\gamma_{ij})$ and

$$c = \mathrm{diag}(\gamma_{11}^{-1}, ..., \gamma_{ss}^{-1}) - \Gamma^{-1},$$
$$(a_{ij}) = (\alpha_{ij})\Gamma^{-1},$$
$$m^T = b^T \Gamma^{-1}.$$

Solving the linear system might not be the cheapest thing...

# Implicit vs Explicit

- Suppose matrix-vector and the RHS evaluation of the ODE have unit cost
- Suppose a s-stages explicit RK and s-stages SDIRK or Rosenbrock
- Find the total number of Krylov iterations one can afford to see some acceleration

# Implicit vs Explicit

- Suppose matrix-vector and the RHS evaluation of the ODE have unit cost
- Suppose a s-stages explicit RK and s-stages SDIRK or Rosenbrock
- Find the total number of Krylov iterations one can afford to see some acceleration

Rosenbrock:

$$\overline{iter} \leq \frac{CFL - 1}{prod \times accel}$$

Newton-Krylov (SDIRK):

$$\overline{iter} \times \overline{newton} \leq \frac{CFL - 1}{prod \times accel}.$$

# Implicit vs Explicit

- Suppose matrix-vector and the RHS evaluation of the ODE have unit cost
- Suppose a s-stages explicit RK and s-stages SDIRK or Rosenbrock
- Find the total number of Krylov iterations one can afford to see some acceleration

Rosenbrock:

$$\overline{iter} \leq \frac{CFL - 1}{prod \times accel}$$

Newton-Krylov (SDIRK):

$$\overline{iter} \times \overline{newton} \leq \frac{CFL - 1}{prod \times accel}.$$

(BiCGStab) prod=2

| $CFL$ / $accel$ | 1 | 2 | 4 |
|---|---|---|---|
| 10 | 4.5 | 2 | 1 |
| 20 | 9.5 | 4.75 | 2.4 |
| 30 | 14.5 | 7.25 | 3.6 |
| 100 | 49.5 | 24.75 | 12.4 |
| 500 | 249.5 | 124.75 | 62.4 |

# Rosenbrock-W: ROW

- Steihaug and Wolfbrandt 1979
- Rosenbrock-W: suppose Jacobian is not exact
- Same stability region as SDIRK if Jacobian is exact
- If not ... stability very hard to study
- Never used with high-order methods
- We need L-stability for PDEs...
- Dense output + error estimator is available

# L-stable Rosenbrock-W

- More order conditions for ROW methods
- p=s (p > 2) for ROW to be L-stable: impossible
- p<s we can get L-stable + W
- Stiffly accurate: no error reduction in RK stages
- Embedded method: error control

Combining ideas in
Hairer and Wanner (II)
we get an L-stable

| |
|---|
| $\gamma = 0.4358665$ |
| $a_{21} = 2.00000$ |
| $a_{31} = 1.41921$ |
| $a_{32} = -0.259$ |
| $a_{41} = 4.18476$ |
| $a_{42} = -0.285$ |
| $a_{43} = 2.29428$ |
| $m_1 = 0.24212$ |
| $m_2 = -1.223$ |
| $m_3 = 1.54526$ |
| $m_4 = 0.43586$ |

# L-stable Rosenbrock-W

- More order conditions for ROW methods
- p=s (p > 2) for ROW to be L-stable: impossible
- p<s we can get L-stable + W
- Stiffly accurate: no error reduction in RK stages
- Embedded method: error control

Combining ideas in Hairer and Wanner (II) we get an L-stable

<u>Error control</u>

| |
|---|
| $\gamma = 0.4358665$ |
| $a_{21} = 2.00000$ |
| $a_{31} = 1.41921$ |
| $a_{32} = -0.259$ |
| $a_{41} = 4.18476$ |
| $a_{42} = -0.285$ |
| $a_{43} = 2.29428$ |
| $m_1 = 0.24212$ |
| $m_2 = -1.223$ |
| $m_3 = 1.54526$ |
| $m_4 = 0.43586$ |

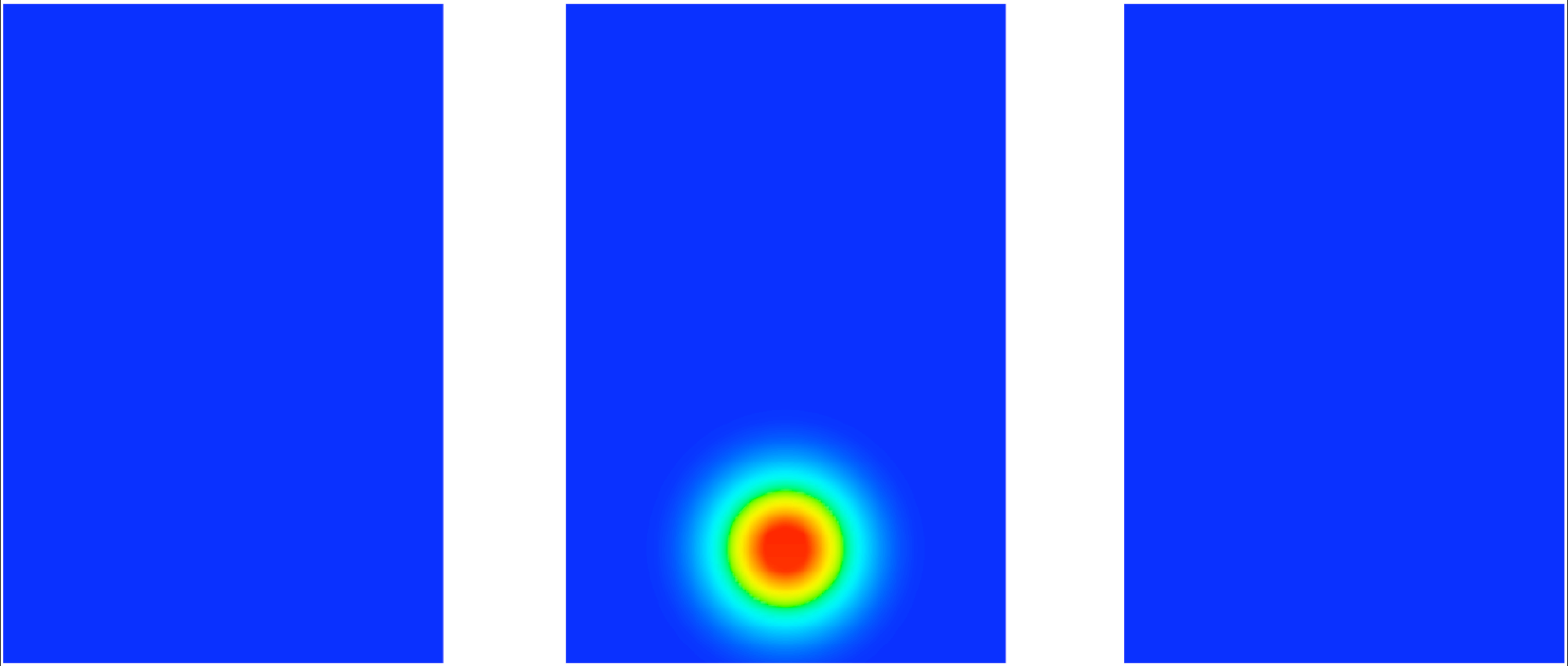
# Stratified Compressible Euler

- Close to incompressible flows: boundary effects
- At high resolutions the nonhydrostatic effects need to be considered: hydrostatic GCM can run at 10 km resolutions now (e.g. HOMME)
- Global next generation GCM will be nonhydrostatic
- High-order: ideally suited for wave propagation phenomena not well suited for shocks and steep gradients: limiting HOMs is research...

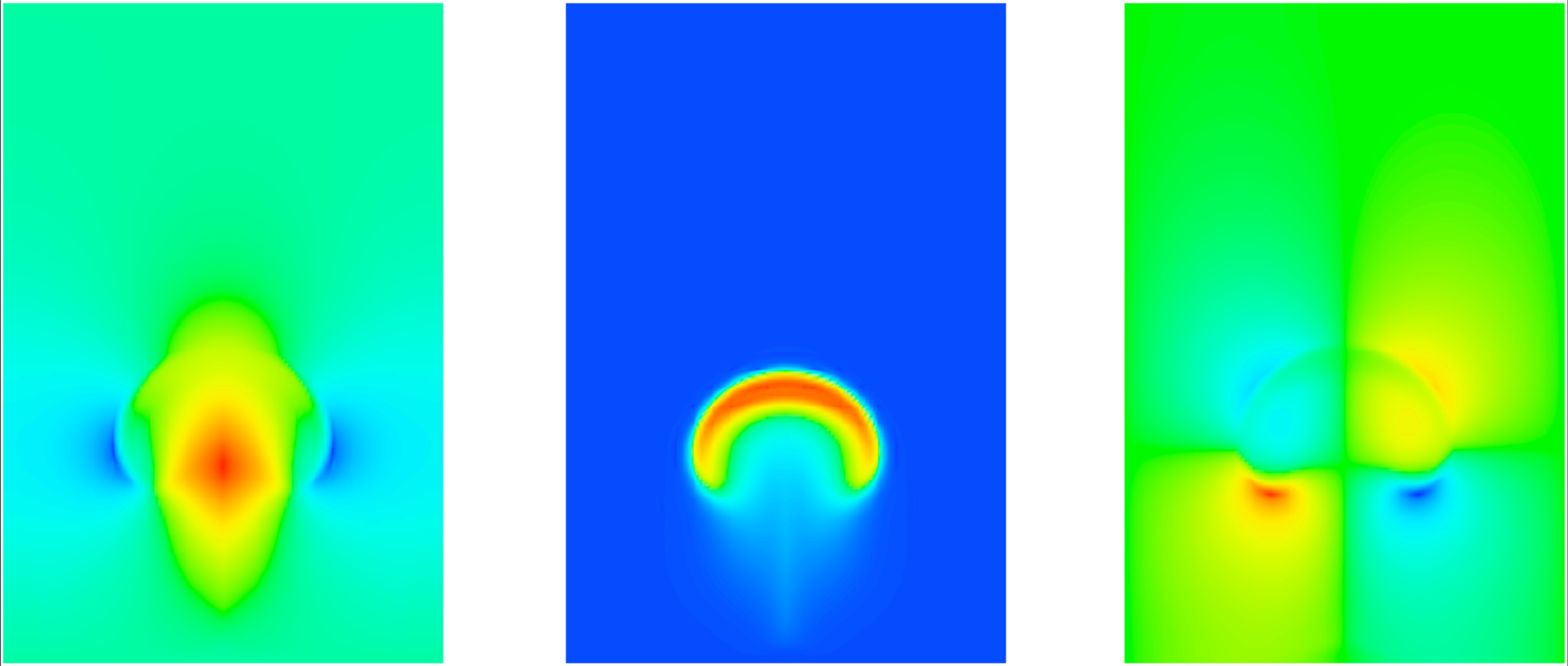
# Numerical experiment: Rising bubble

- Hydrostatically balanced flow
- Potential temperature perturbed
- Domain 1.0 km x 1.5 km resolution: 6m
- Robert (1993): slow moving large scale bubble with fast acoustic waves reflected
- Integrate for 1800 secs: bubble crashes onto top lid
- Block Jacobi preconditioning + Guillard and Viozat diffusive term (1998)
- dt fixed to 1 secs, acceleration observed ~6-8: compared to an SSP with CFL=2

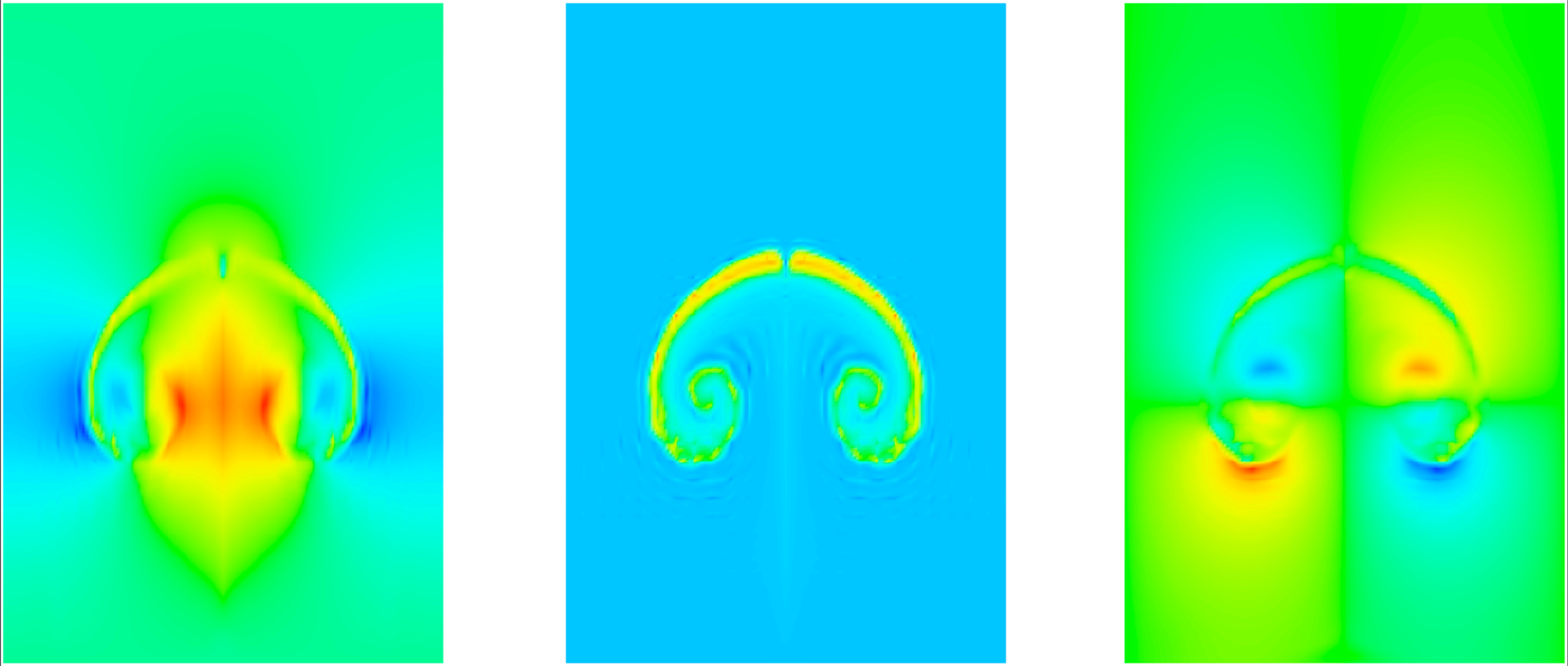
# Initial condition (t=0)



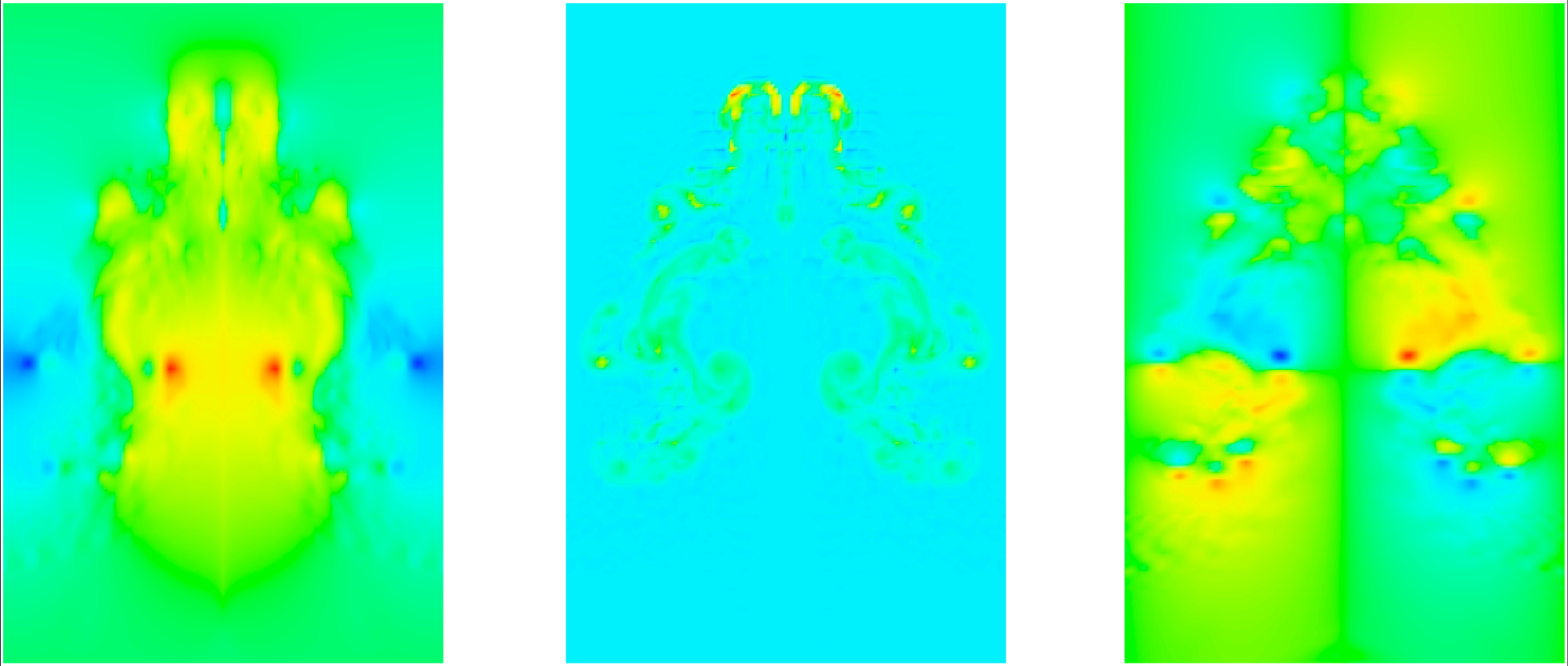
# t=360 seconds



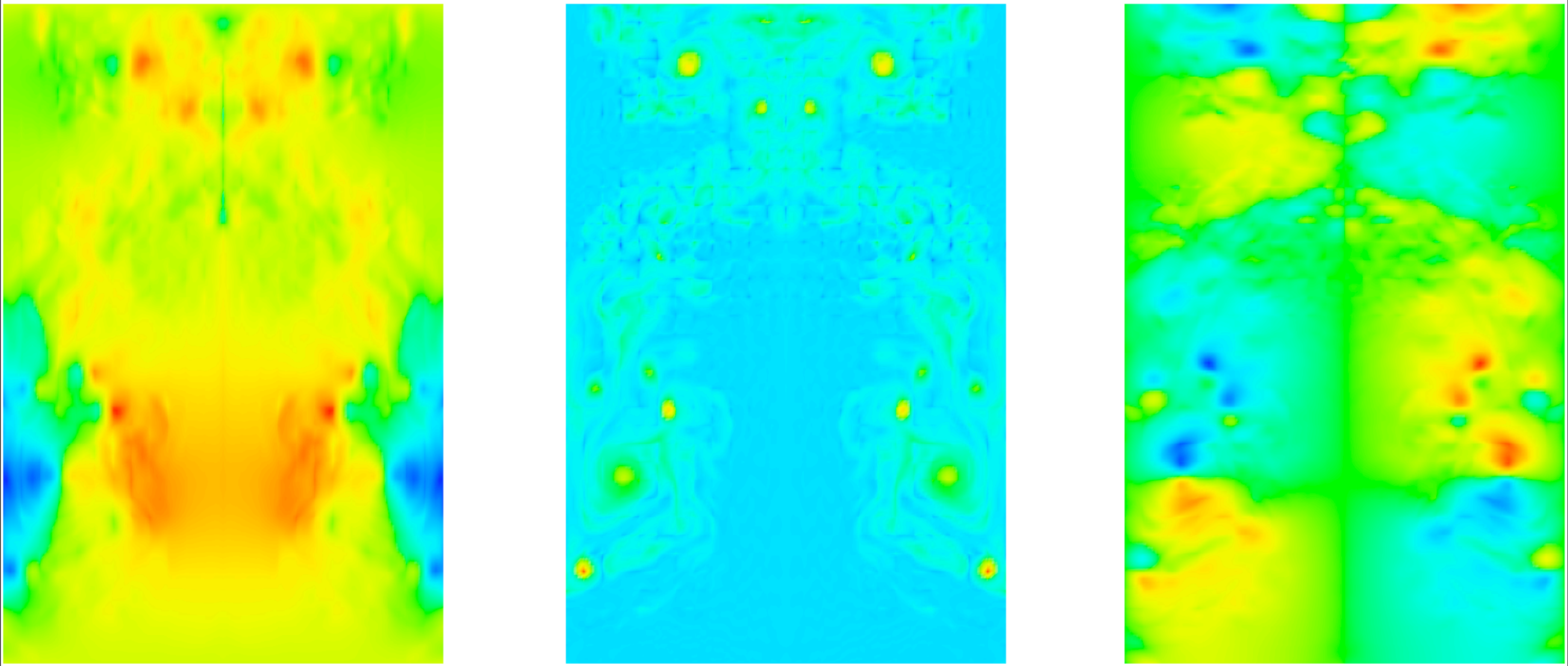
# t=720 seconds



# t=1080 seconds



# t=1800 seconds



# Conclusions and future work

- Petascale computing imposes constraints!
- The barotropic problem is solved optimally by OAS
- A very cache friendly OAS version was derived: compares to FDM

---

- A cheap way of achieving implicit time integration for PDEs was derived
- Possible to attain high-order in time: error control
- Stiffer problems will be considered: mountains and gravity waves in a channel
- Preconditioning ... + projections using lower polynomial degrees to construct better starting estimates
- ROW + SSP: L-stable Rosenbrock-W method with explicit part SSP (Joint with Prof. Sandu and E. Constantinescu: Virginia Tech)

# Acknowledgments

- The AMR work was funded under NSF grant CMG-0222282: An adaptive Mesh, Spectral Element Formulation of the Well-Posed Primitive Equations for Climate and Weather Modeling, NSF MRI Grant CNS-0421498, NSF MRI Grant CNS-0420873, NSF MRI Grant CNS-0420985
- Partly supported by the NSF-CMG grant: Adaptive High-Order Methods for Nonhydrostatic Numerical Weather Prediction 0530820.
- DOE Climate Change Prediction Program CCPP.

Email: amik@ucar.edu
Phone: 303-497-1287