On the status of algebraic multigrid preconditioners

4th Workshop of the INRIA-UIUC Joint Laboratory for Petascale Computing
Urbana-Champaign, November 22-24, 2010

MG@CS@Illinois
Bill Gropp
Hormozd Gahvari

Luke Olson
James Lai
Steven Dalton

Nathan Bell (Nvidia)
Jacob Schroder (UCB)

Potential INRIA collaborators
Stéphane Lanteri
Laura Grigori
Luc Giraud

wanted: more robust, efficient, and principled solver

algebraic **multilevel** **projections**

problem: AMG designed for basic problems

problem: AMG does not scale

problem: no sense of optimality in AMG

show the potential for AMG

highlight optimality

high-performance progress

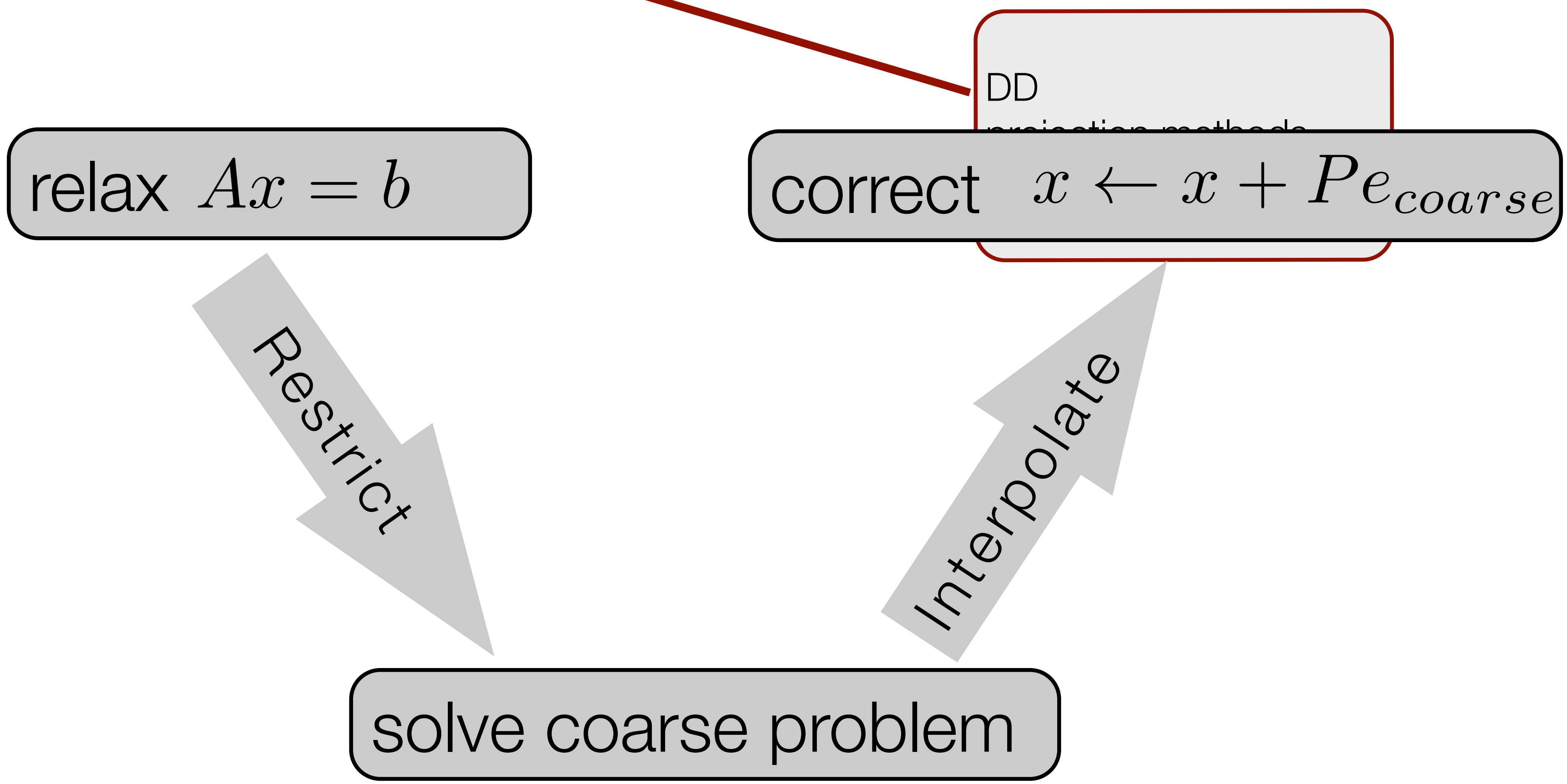
goal: strengthen and increase applicability through the Joint Lab

Multilevel view

Jacobi, Gauss-Seidel
Chebychev
ILU
Kaczkmarz

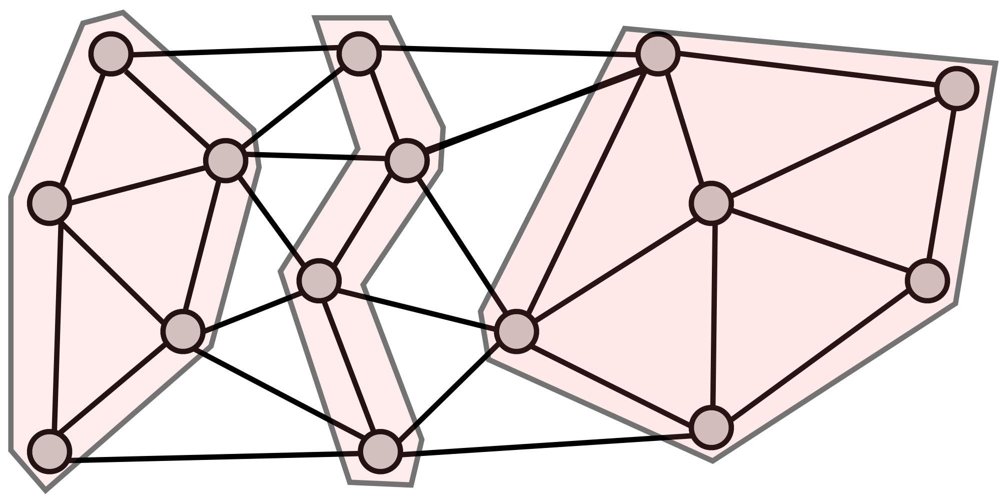
attenuate high energy quickly with with relaxation

attenuate low energy error through coarse-grid correction



AMG Framework

- aggregation: groups of fine nodes form coarse nodes



fine: 15
coarse: 3

- an initial interpolation pattern

- find an optimal interpolation operator P that contains low energy

- cycle:

$$x \leftarrow x + P(P^T A P)^{-1} P^T r$$

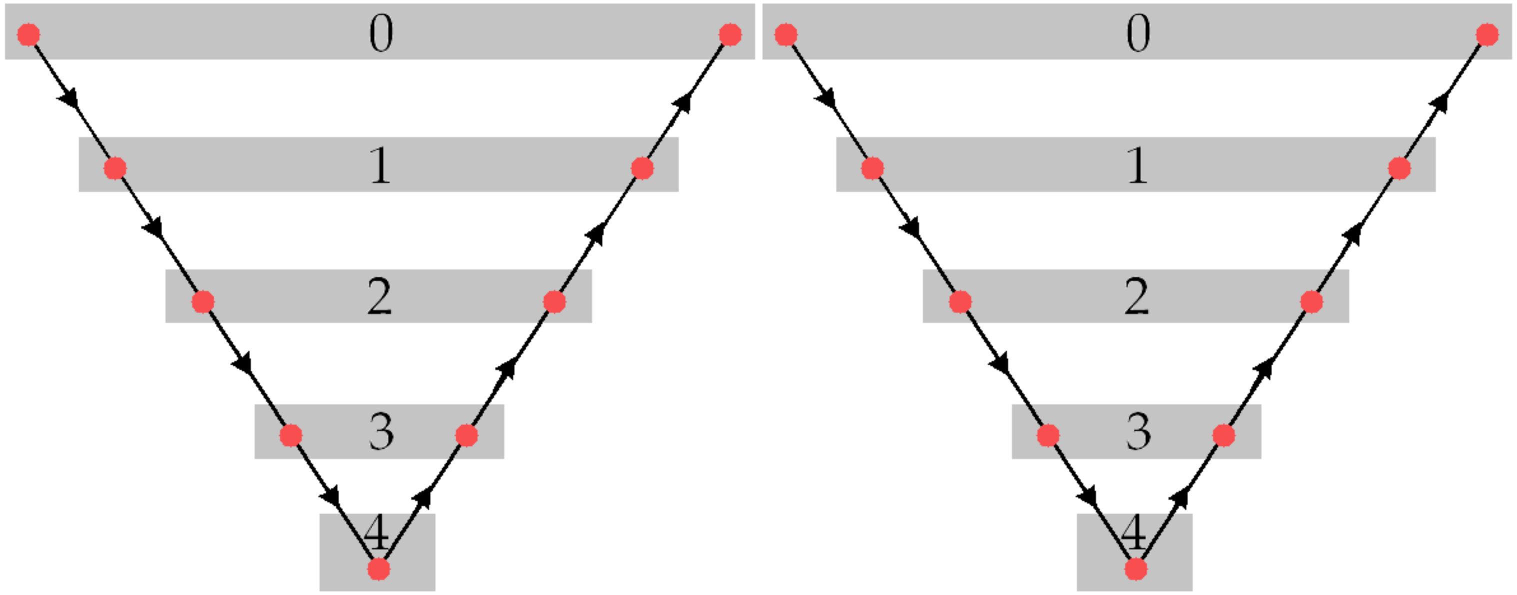
$$x \leftarrow x + P A_{coarse}^{-1} P^T r$$

AMG Components: what we need

- an idea of the low energy: physics, training, intuition $AB \approx 0$
- a strength measure to determine strong node couplings S_{ij}
- a parallel aggregation method Agg
- low complexity, optimal interpolation P

Setup

- better cycling
  - richer coarse grids
  - “parallel” cycling

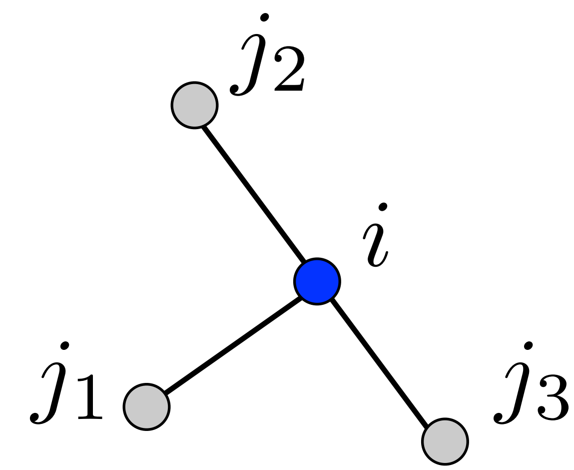


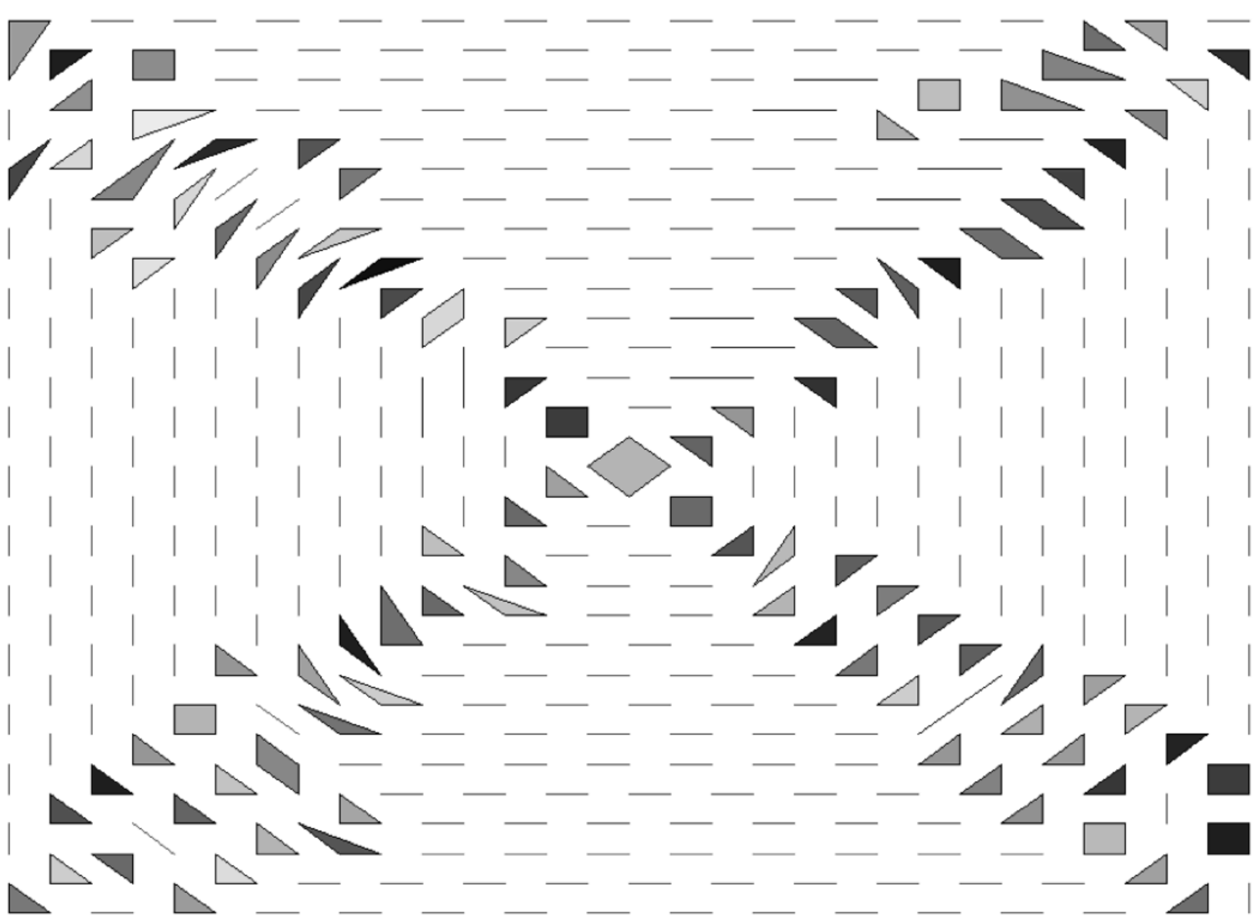
Solve

*** Olson, Schroder, Tuminaro, *A new perspective on strength measures in algebraic multigrid*, 2009

Evolution Measure

1. drop point source at a node
2. evolve/diffusion point source with A
3. evaluate diffusivity at neighbors in comparison to known low energy





- efficient
- parameter insensitive
- Euler flow
- wave problems
- high-order
- discontinuous elements

*** Bell, Olson, *Lloyd Aggregation*, 2008
*** Bell, Dalton, Olson, *AMG on the GPU, 2010*

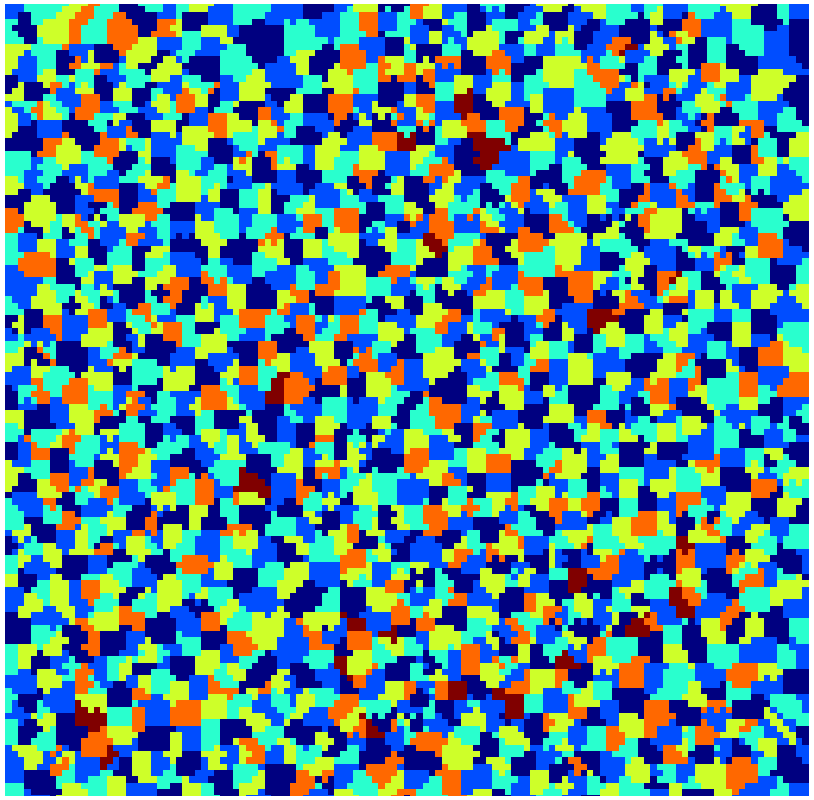
Flexible aggregation

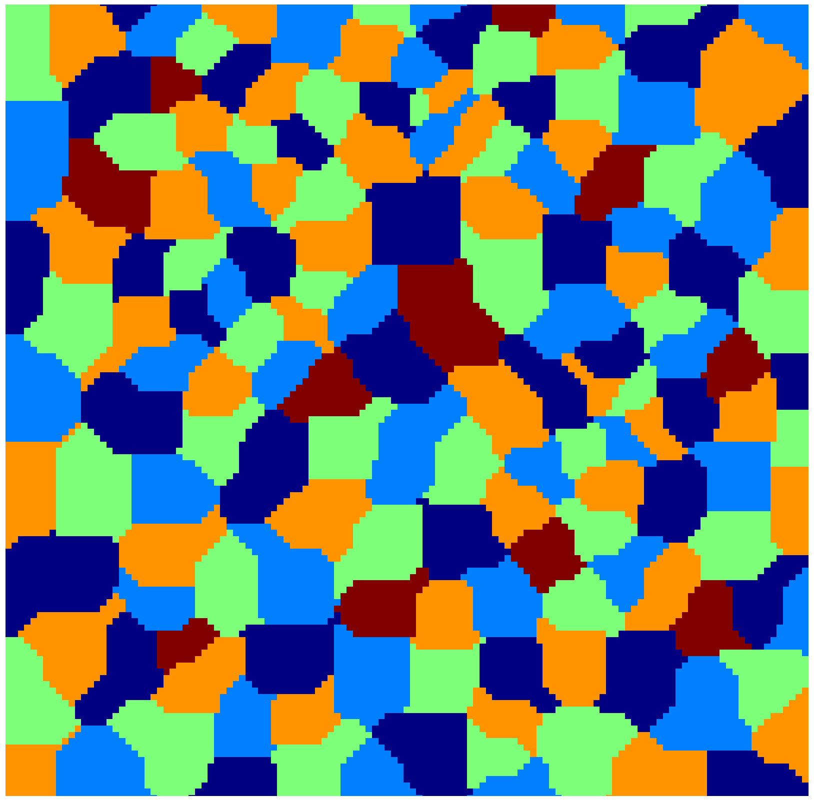
- we use two approaches:

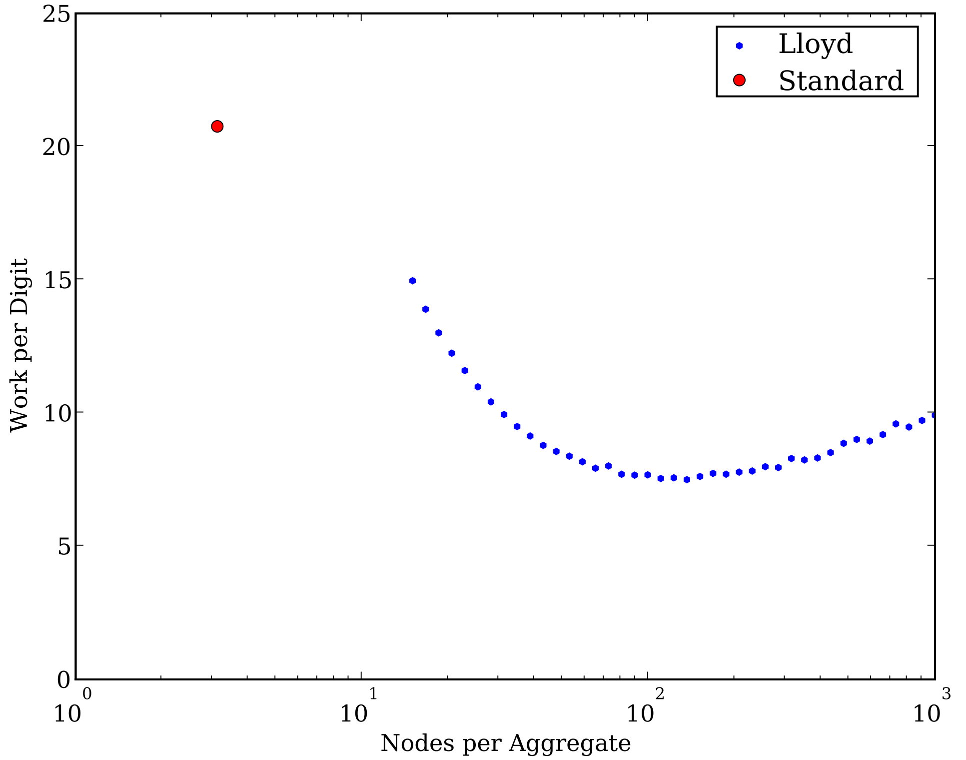
  1. MIS(2): parallel coarse grids == serial coarse grids

  2. shortest-path: ability to tune bandwidth

- **need to combine**
- **need to develop autotuning for heterogeneous archs**





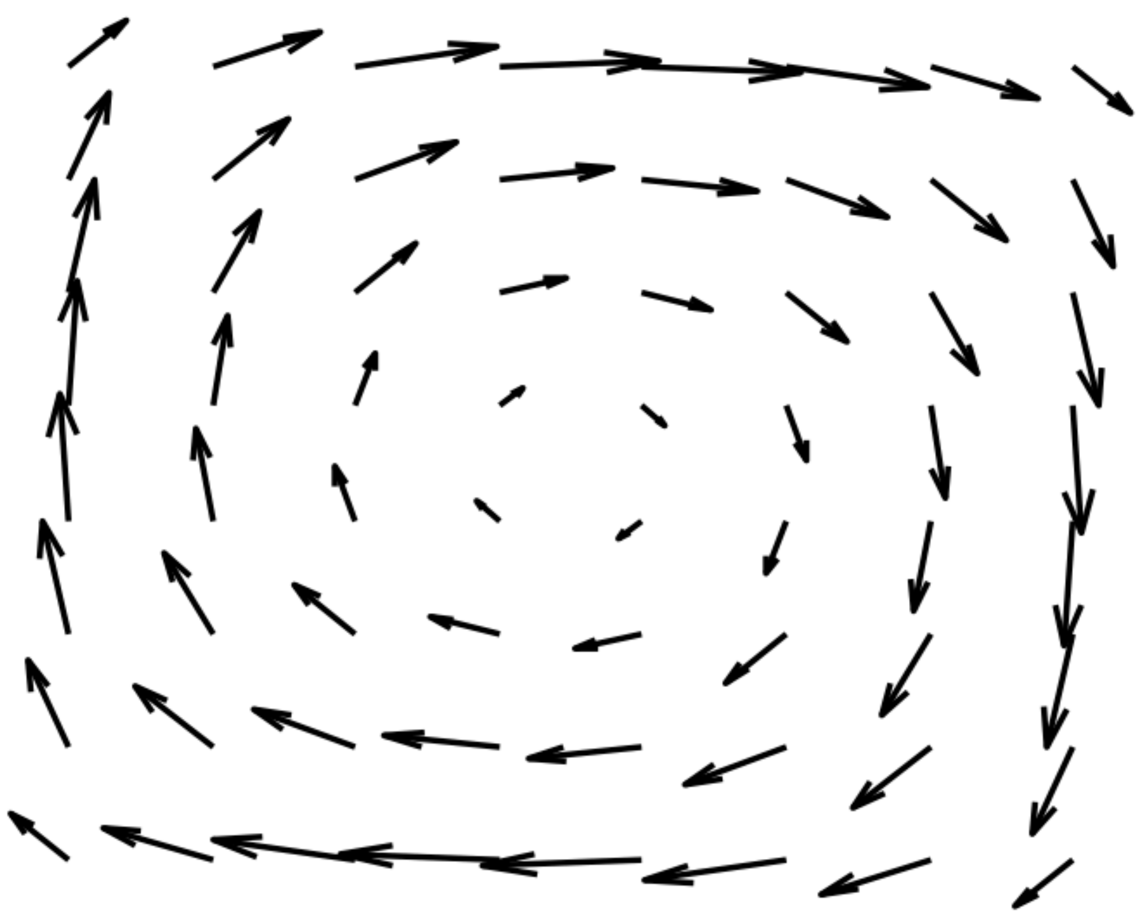


*** Olson, Schroder, Tuminaro, A *general interpolation strategy for algebraic multigrid using energy-minimization,* 2010.

optimizing energy

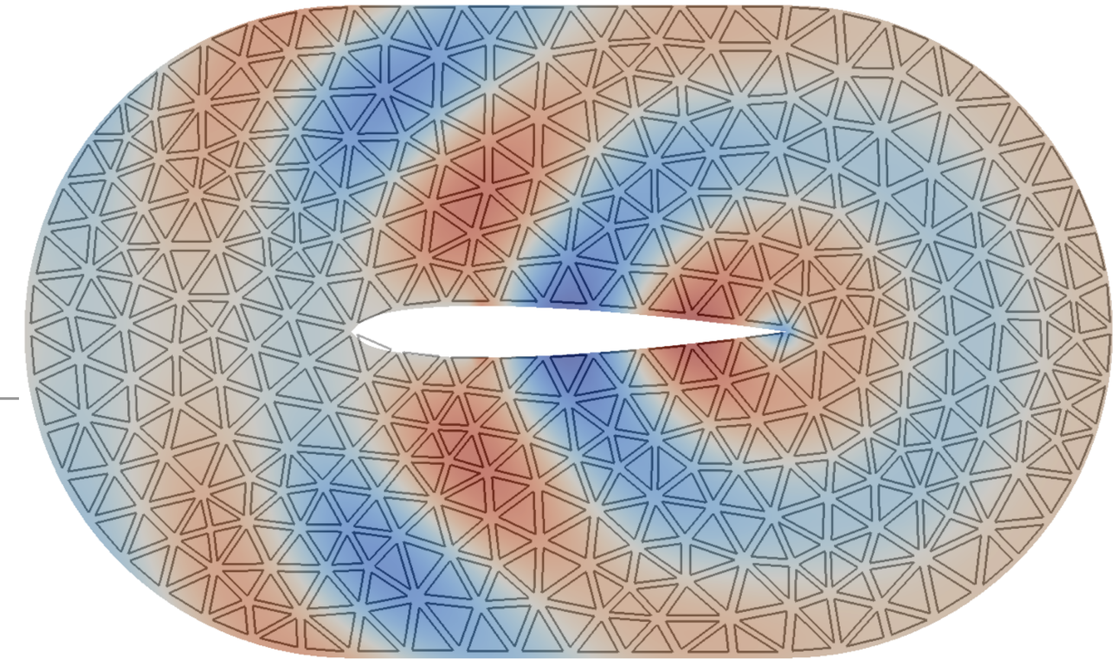
- P should have low energy (low A-norm or A^*A-norm)
  1. determine sparsity pattern
  2. minimize energy column-wise (parallel)

- **need to push toward more demanding non-symmetric applications**
- **need a formal non-symmetric process**



h	std.	opt.
1/64	>150	24
1/128	>150	28
1/256	>150	33
1/512	>150	33

Wave problems (Helmholtz)



- AMG problem: standard “low energy” modes (constant) break Nyquist rate

- answer: introduce wave modes at all levels $e^{i\omega x}$

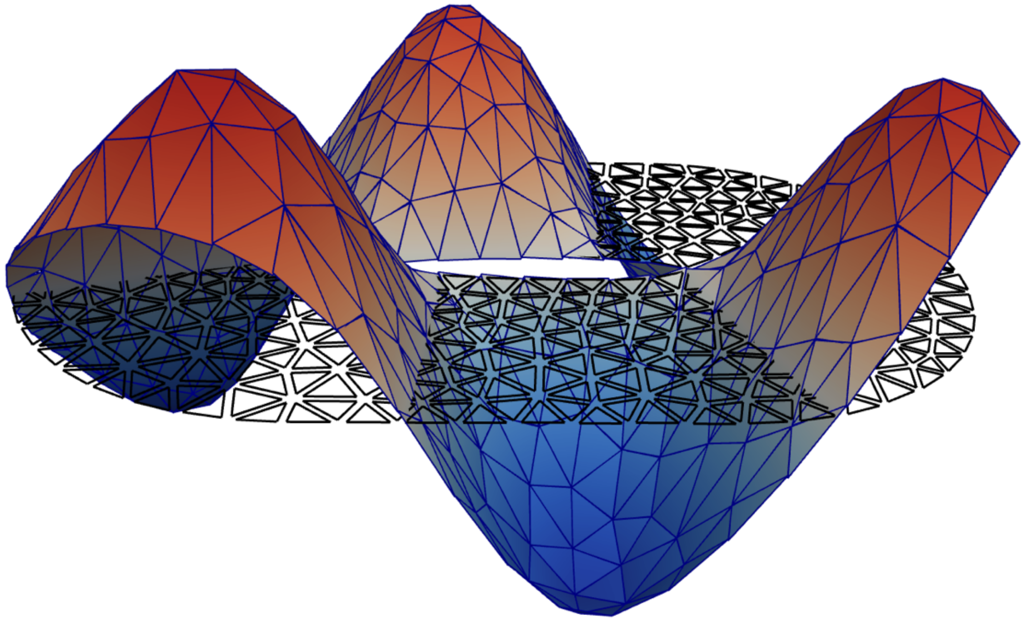
multiple wave modes: ⇒ $e^{i\omega \cos(\theta)x + \sin(\theta)y}$

infuse wave modes: ⇒ level 1: $\theta = 0, \frac{\pi}{2}$ level 2: $\theta = 0, \frac{\pi}{8}, \frac{3\pi}{8}, \frac{5\pi}{8}, \frac{7\pi}{8}$

adapt to grid through relaxation: ⇒ $Ae^{i\omega \mathbf{x}} = 0$

optimize interpolation: ⇒ $\|P\|_{A^*A}$

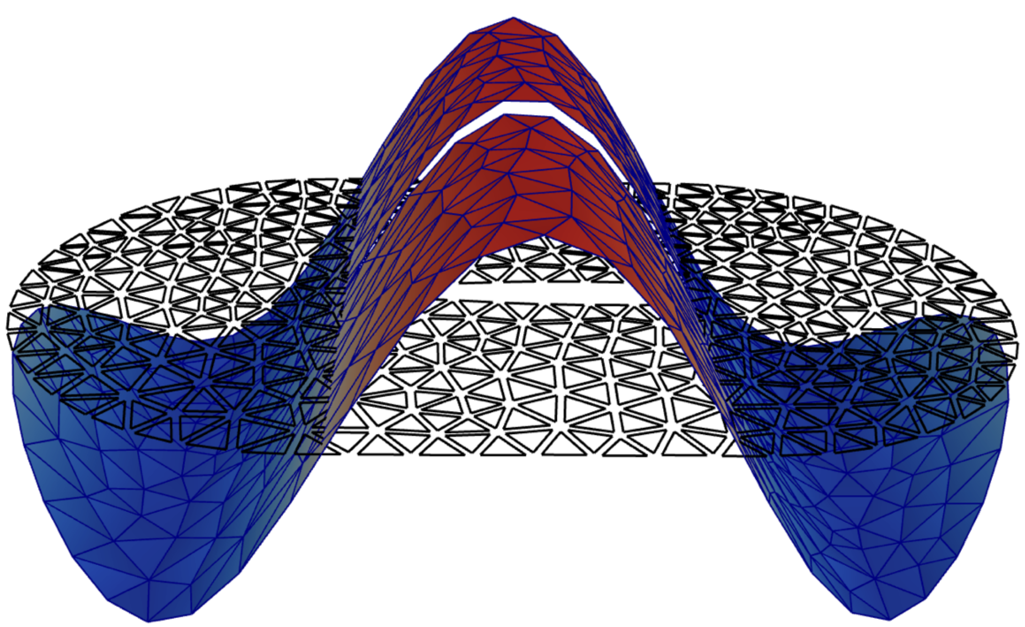
*** Olson, Schroder, *A smoothed aggregation multigrid method for Helmholtz problems*, 2010.



left singular vector

- **need to extend to Maxwell equations**
- **need to test robustness for lossy media**

injected mode



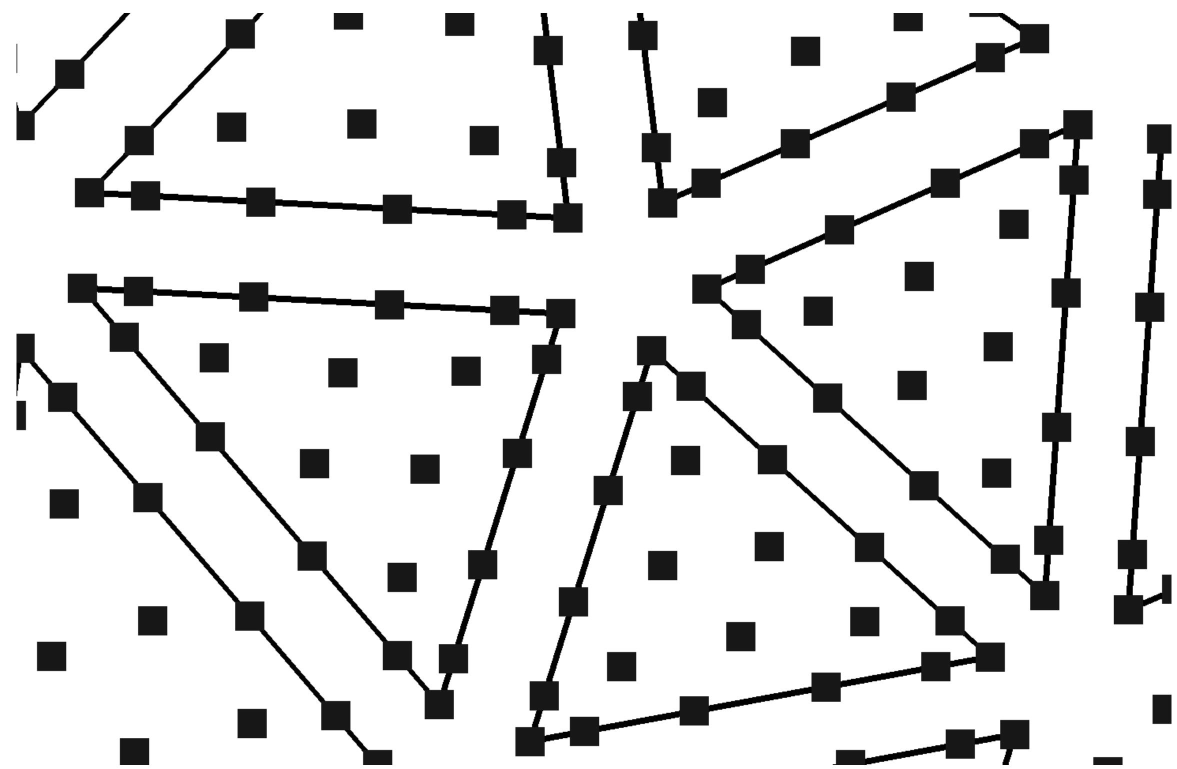
ppw	h	$h/2$	$h/4$	$h/8$	$h/12$
5.0	9	11	15	25	54
10.0	9	11	11	11	14
30.0	9	10	10	11	11
90.0	9	10	11	11	11
All	4	15	17	21	22

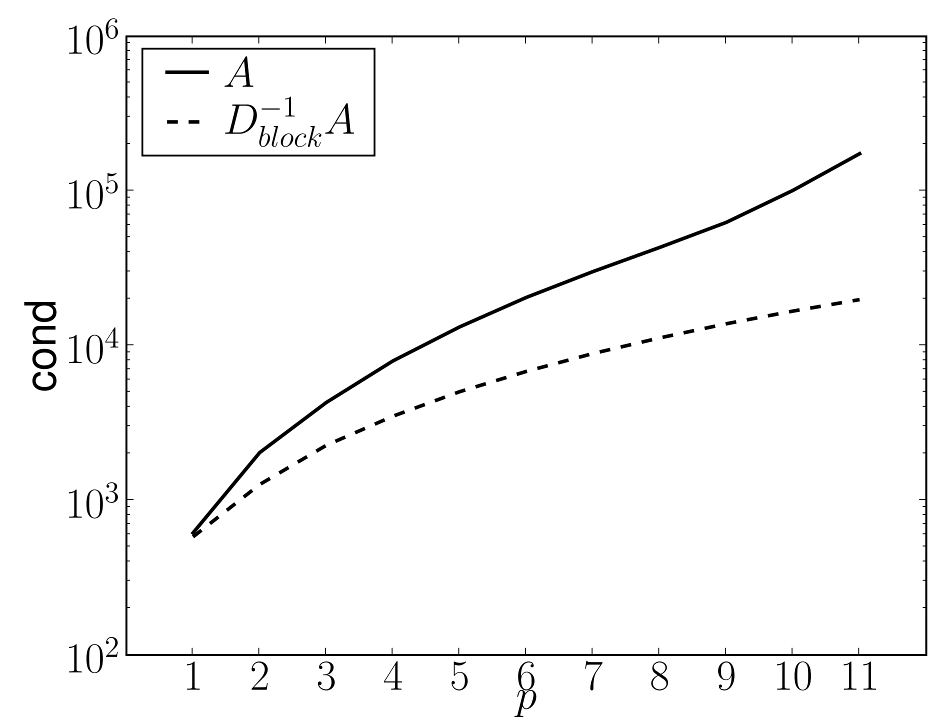
pgmres iterations, complexity

*** Olson, Schroder, *Smoothed aggregation multigrid solvers for high-order discontinuous galerkin methods*, 2010.

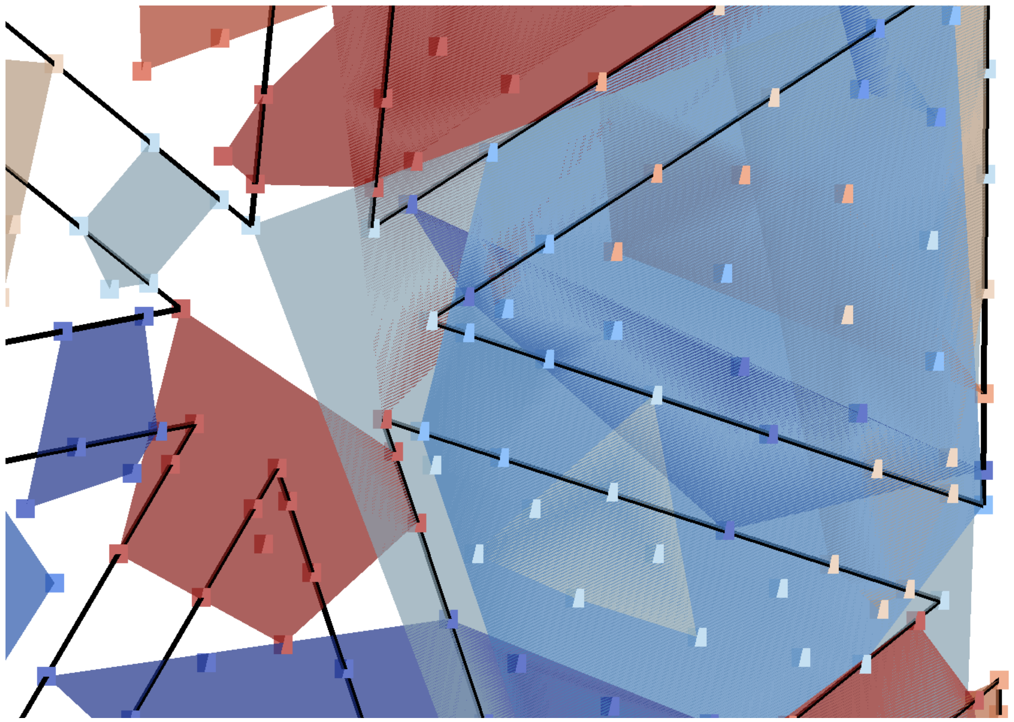
High-order discontinuous Galerkin

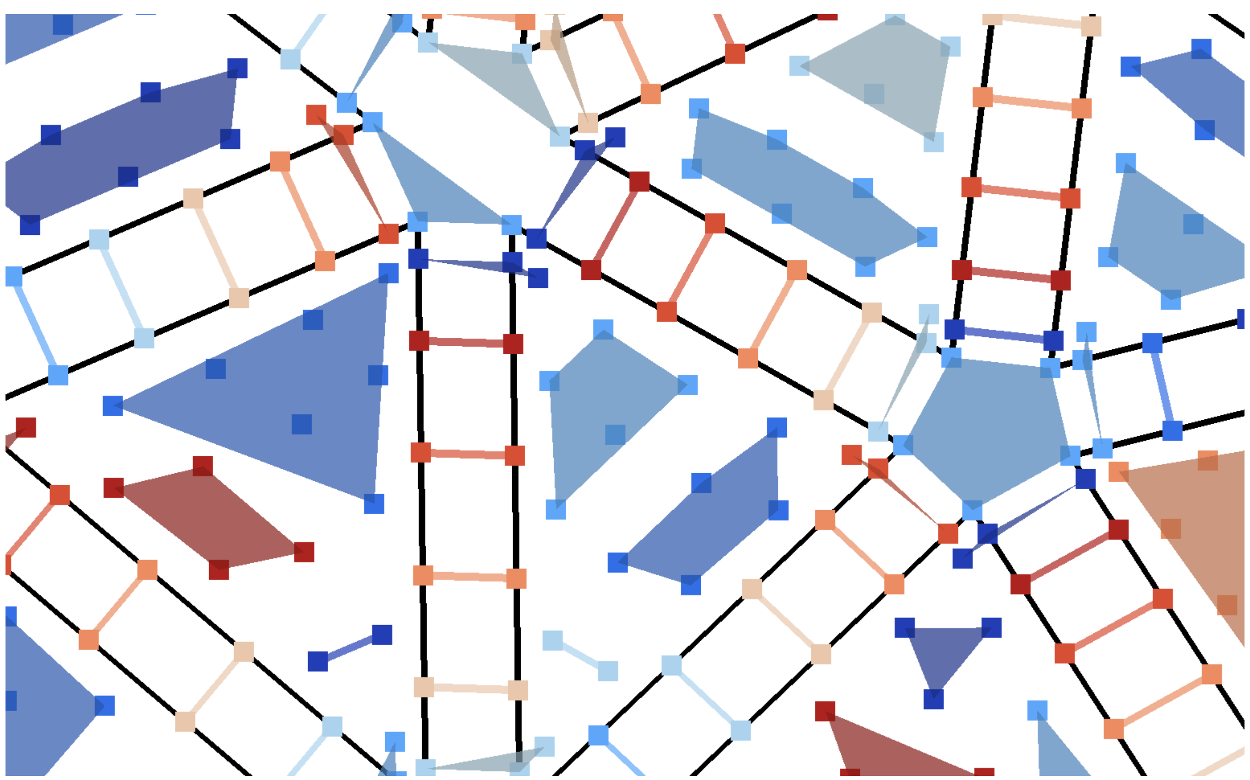
- different types of d.o.f.
- loss of locality
- increase in condition number
- most heuristics in AMG break down





High-order discontinuous Galerkin: Poisson



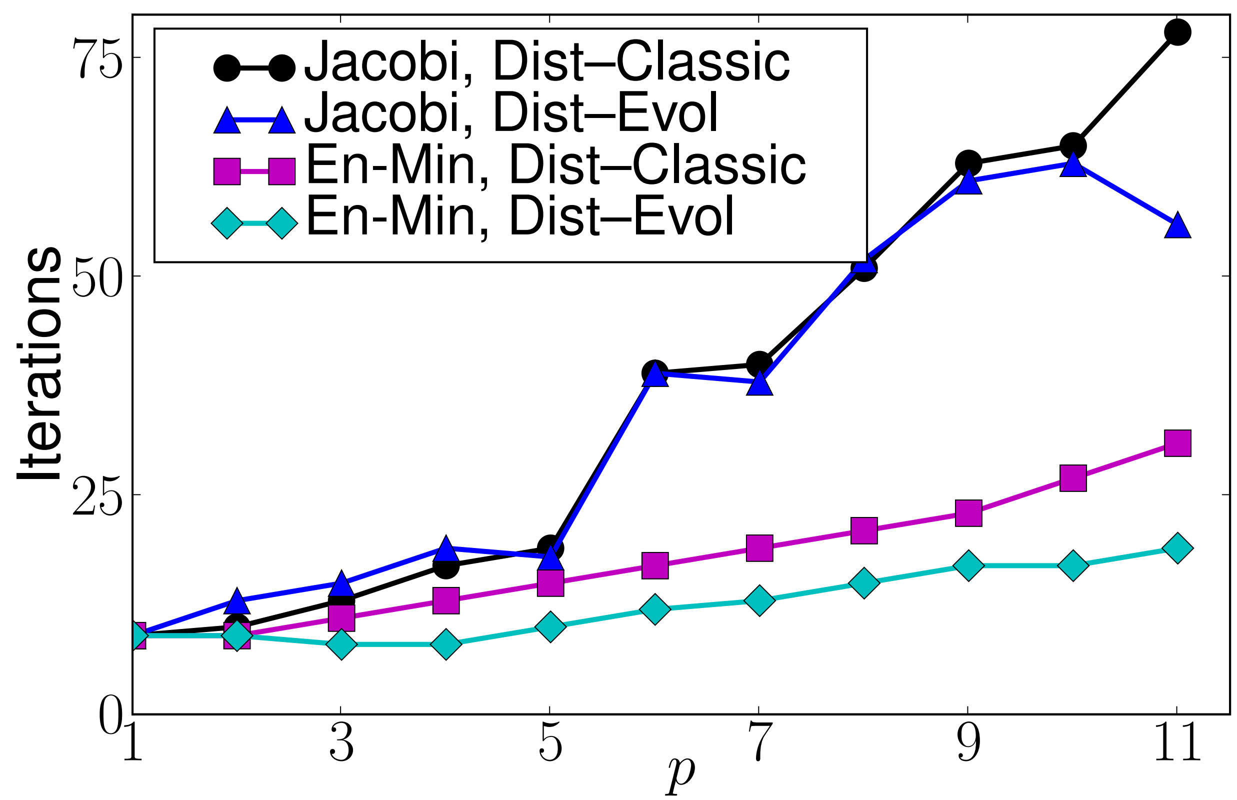


key ingredients:

- conforming aggregations step
- adapt the near null space
- optimal interpolation

needs:

- extend to Maxwell equations



AMG on the GPU

need for acceleration

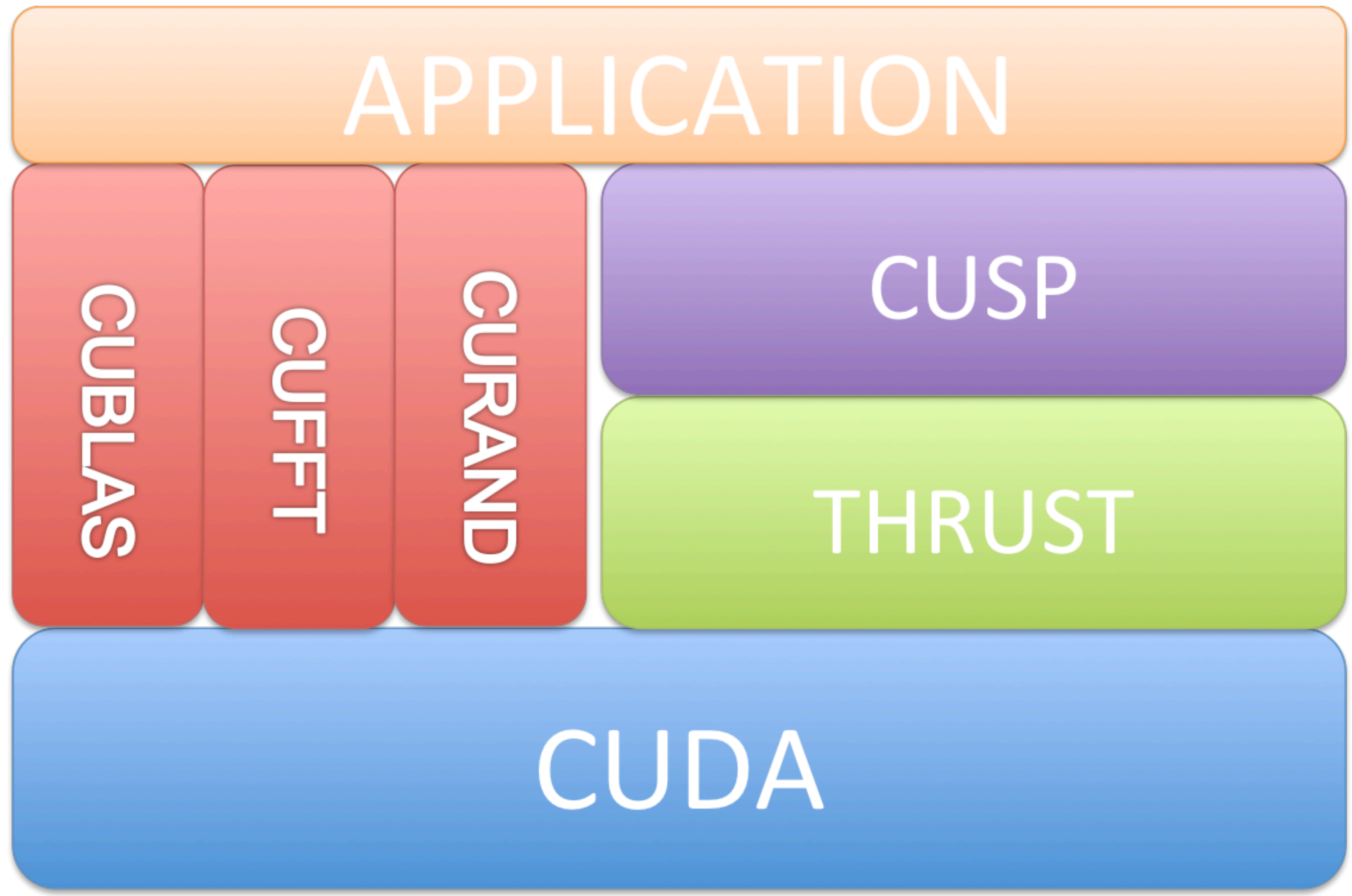
multicore difficult to optimize for AMG

efficient GPU kernels: SpMV

*** Bell, Dalton, Olson, *AMG on the GPU*

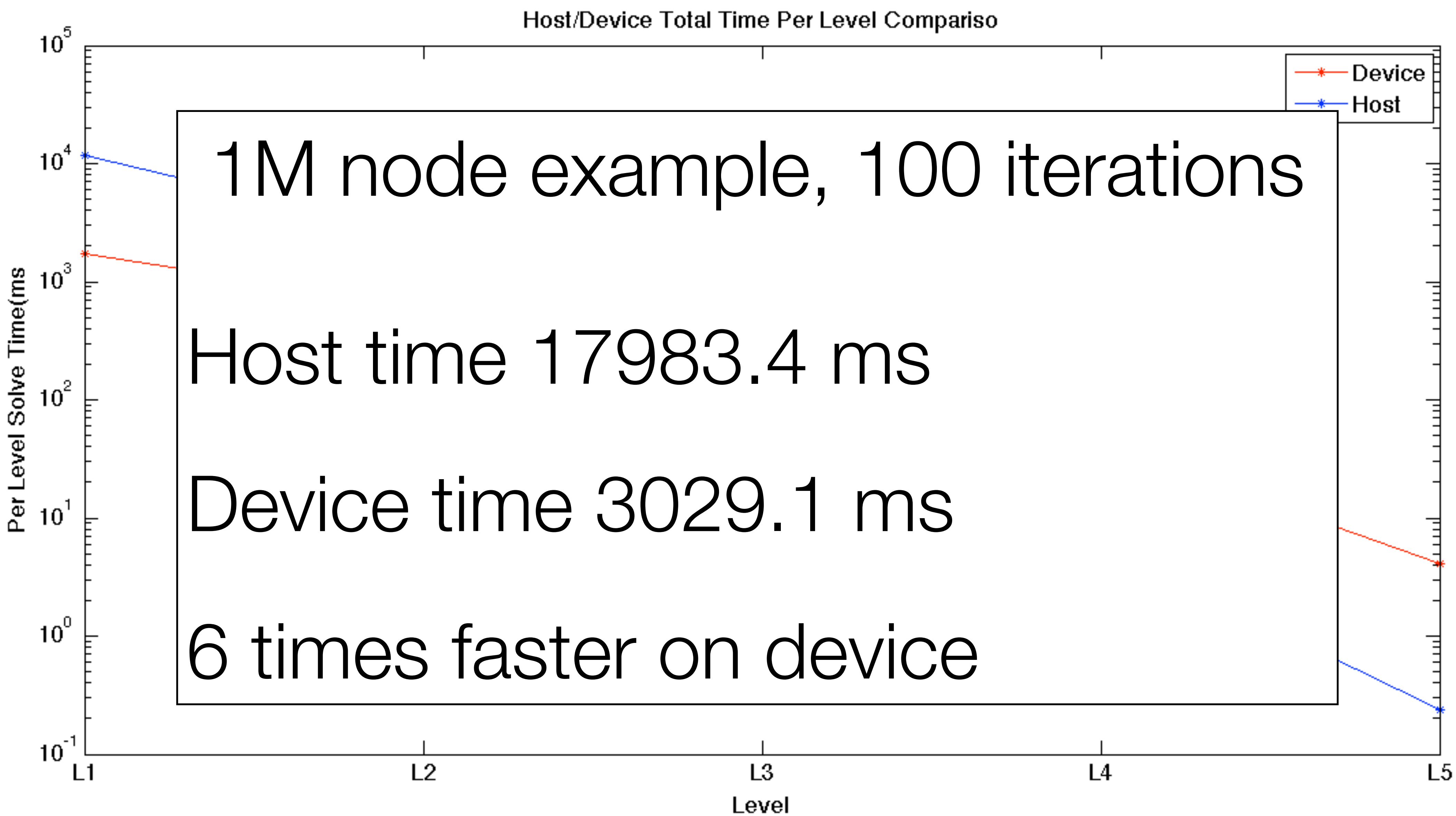
AMG on the GPU

- Cusp (Cuda) framework
- have developed hooks for use with Jumpshot
- target: setup+solve < 1sec 1M dof problem
- host: many seconds

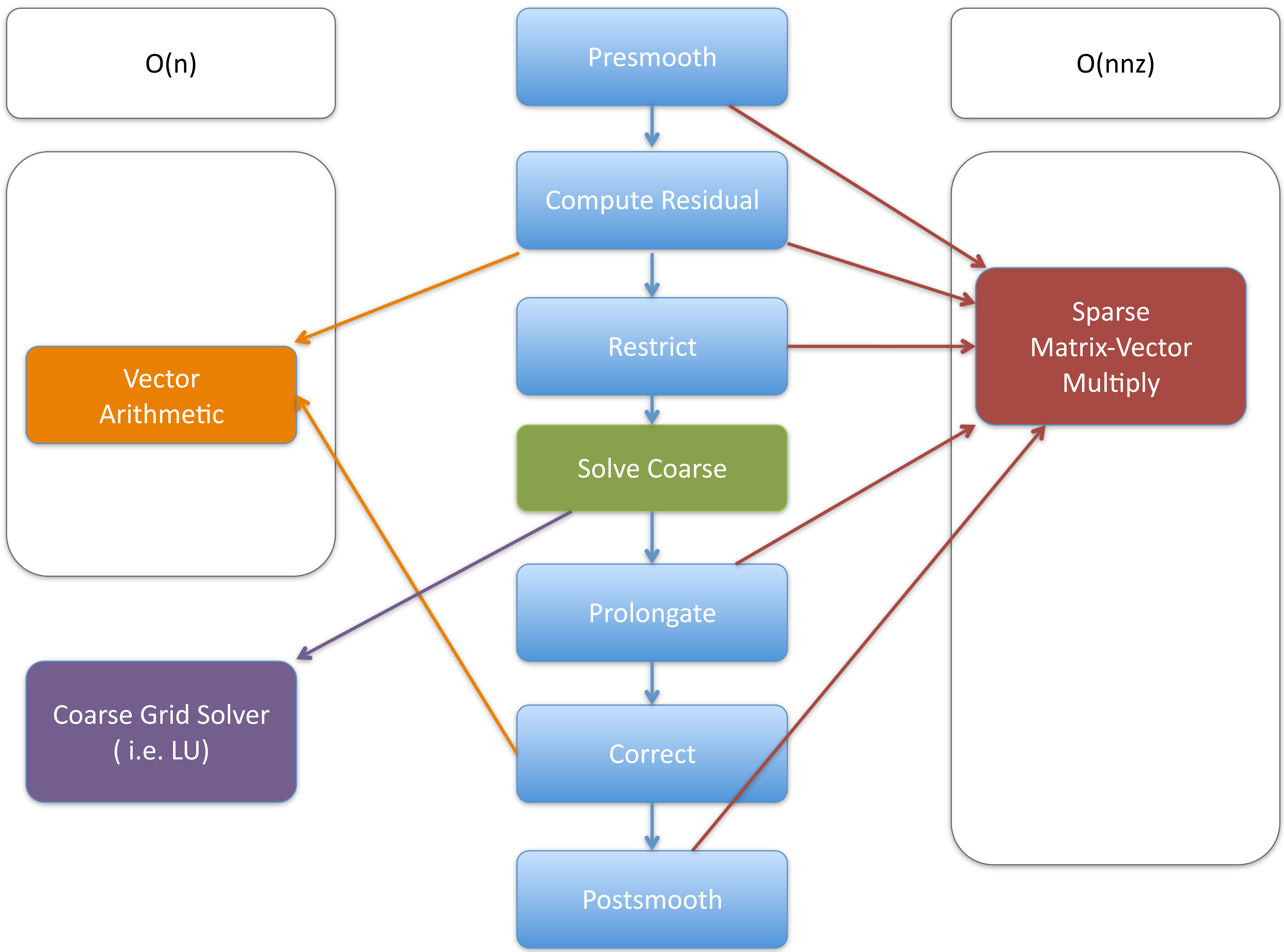


Solve phase

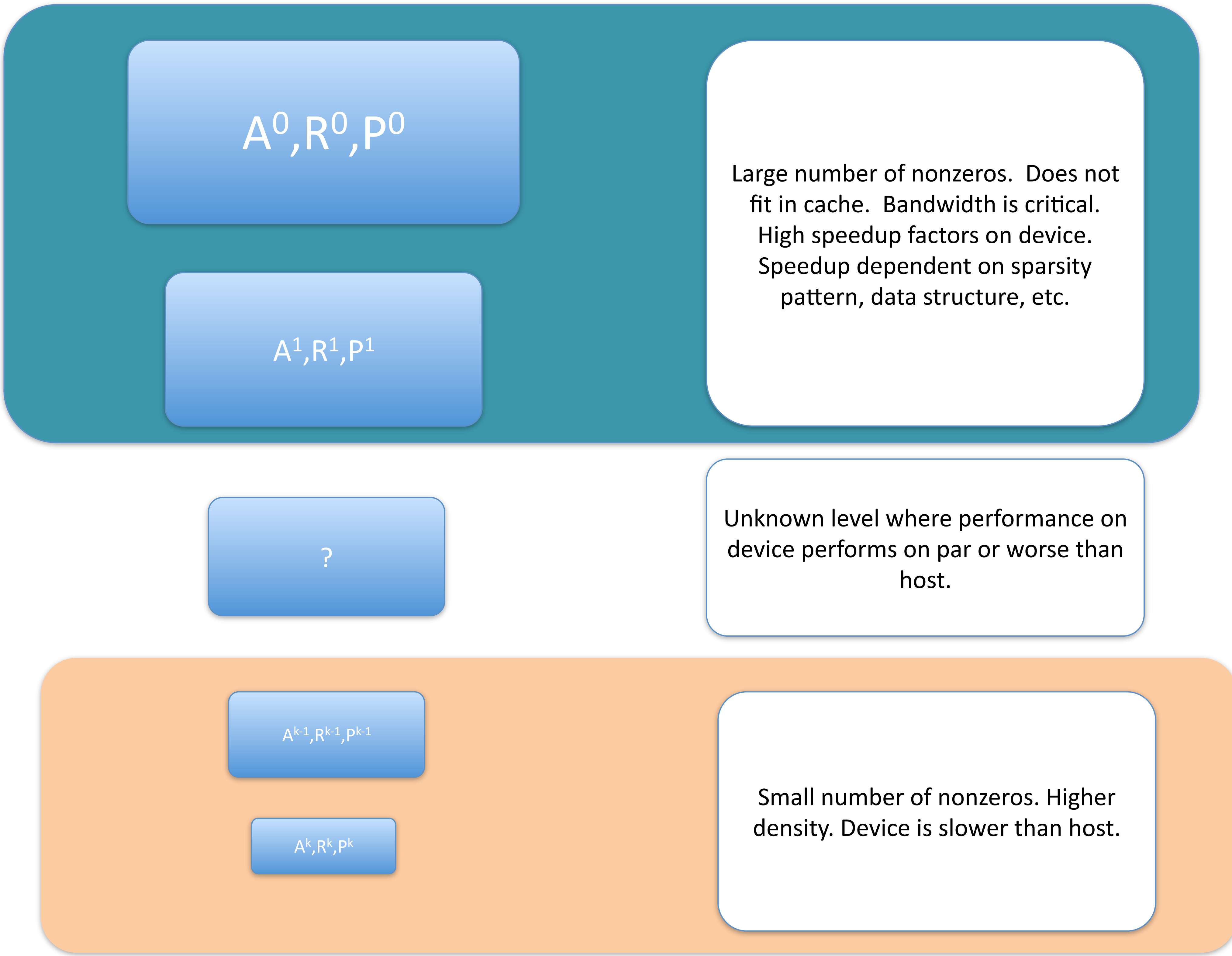
- relies on fast SpMV
- performance hit: GPU kernel launches
- performance hit: transfer
- Device: likes sparsity
- host: likes density



AMG Solve phase: parallel operations



Sparse Mat-Vec



Setup phase

- strength of connection
- aggregation
- tentative interpolation
- smoothed interpolation
- transpose
- Galerkin triple product
- relaxation

1M node example

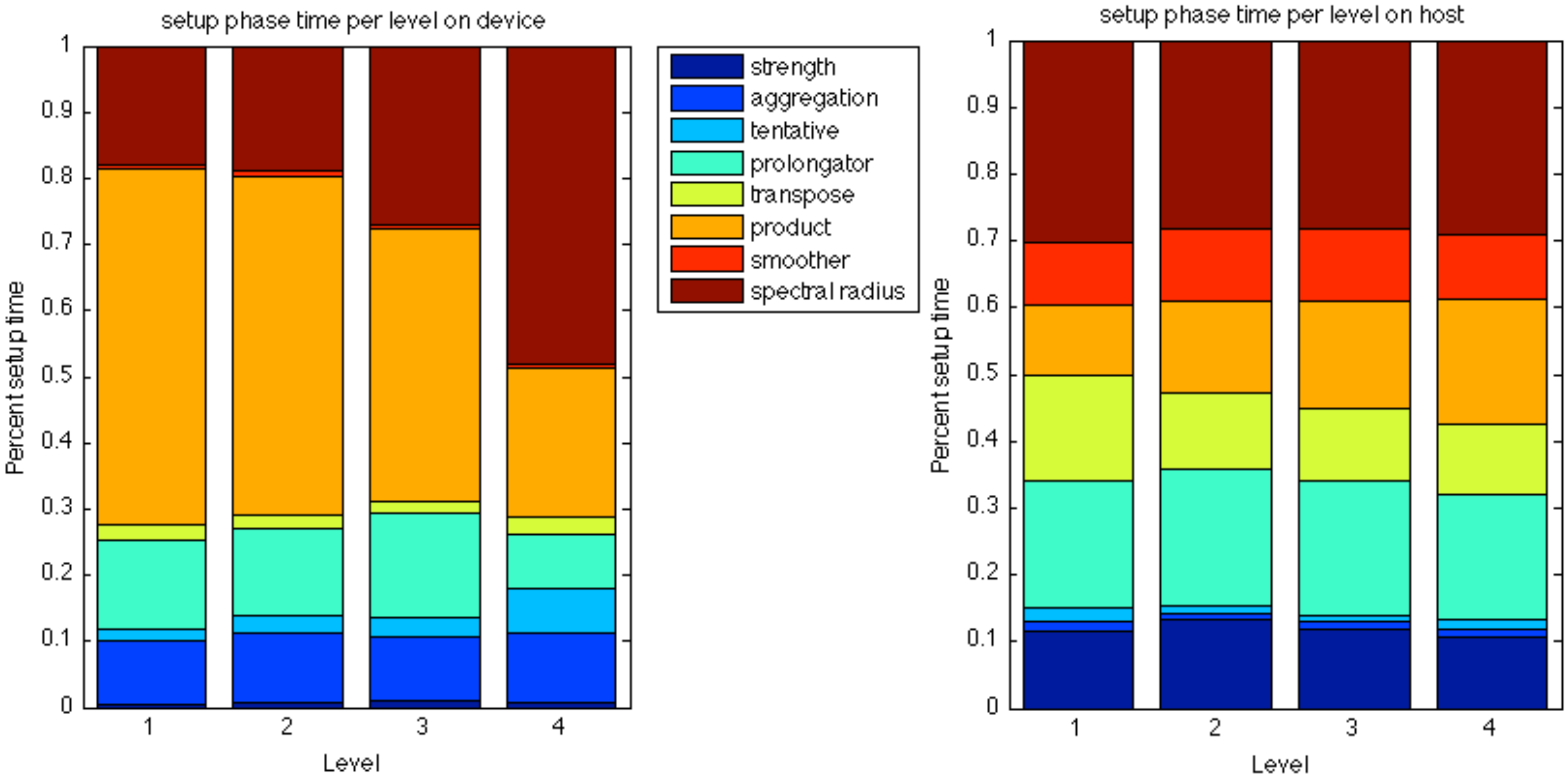
Host time 4558.97 ms

Device time 716.86 ms

6 times faster on device

Number of levels	5	MIS(2)
Level	Unknowns	Nonzeros
1	1048576	5238784
2	146193	1586645
3	8680	128908
4	363	5291
5	18	172

Setup Phase



setup phase time per level on device
Percent setup time
Level
strength
aggregation
tentative
prolongator
transpose
product
smoother
spectral radius
setup phase time per level on host
Percent setup time
Level

collaboration: E&M

Scientific objectives

- Design, analysis and validation of numerical methods and high performance resolution algorithms for the computer simulation of wave propagation problems in complex domains and heterogeneous media

Research directions

NACHOS

- Systems of linear PDEs with variable coefficients
  - Discretization
    - Discontinuous finite element (DG) methods on unstructured meshes
    - High order polynomial interpolation
    - *p*-, *h*- and *hp*-adaptivity
    - Numerical treatment of complex propagation media models
  - Resolution
    - Accurate and efficient time integration strategies
    - Domain decomposition (DD) methods
  - High peformance computing
    - Algorithmic aspects (parallel resolution methods)
    - Implementation issues (mixed SIMD/MIMD programming model)

collaboration: E&M

Computational electromagnetics

- System of Maxwell equations
- Dispersive propagation media
- Applications involve the interaction of electromagnetic waves with,
  1. biological tissues (biocem),
  2. geological media (georadar).

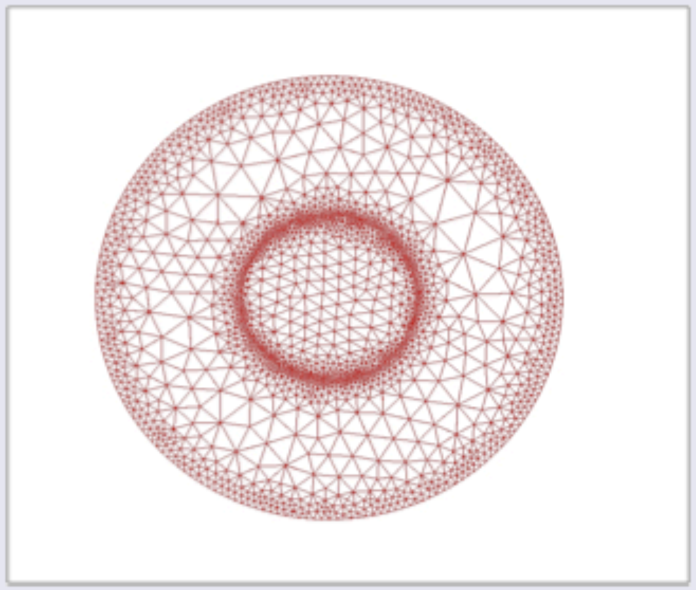
James Clerk Maxwell (1831-1879)

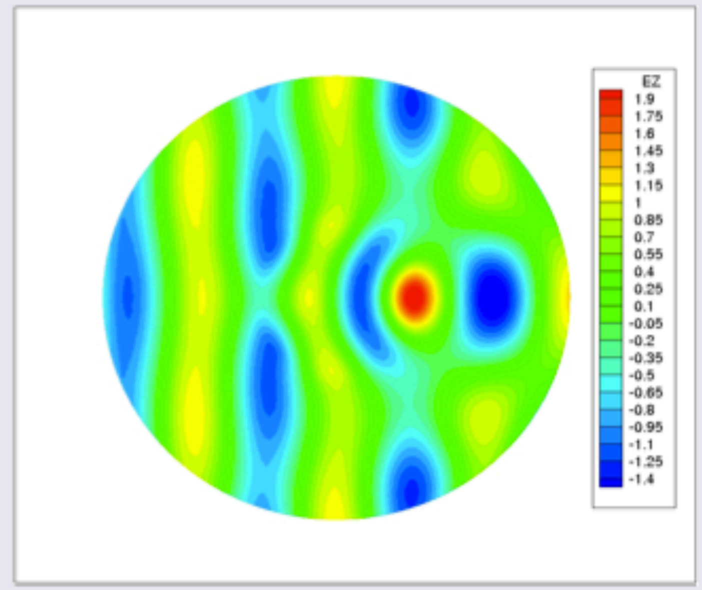
Computational geoseismics

- System of elastodynamic equations
- Viscoelastic propagation media
- Applications deal with the propagation of seismic waves,
  1. generated by an explosive source (earthquake dynamics),
  2. in the deep subsurface (resource prospection).

Nachos team, Domain Decomposition

Classical (non-optimized) Schwarz algorithm





Method	L2 error on E_z	N_s	# iter BiCGStab ($\varepsilon = 10^{-6}$)
DGTH-$\mathbb{P}_1$	0.16400	4	317
DGTH-$\mathbb{P}_2$	0.05701	4	650
DGTH-$\mathbb{P}_3$	0.05519	4	1067
DGTH-$\mathbb{P}_4$	0.05428	4	1619
DGTH-$\mathbb{P}_i$	0.05487	4	352

Nachos team, Domain Decomposition

Optimized Schwarz algorithm (case 1)

Method	L2 error on E_z	N_s	# iter BiCGStab ($\varepsilon = 10^{-6}$)
DGTH-$\mathbb{P}_1$	0.16457	4	52 (6.1) [a]
-	0.16467	16	83 (4.7)
DGTH-$\mathbb{P}_2$	0.05705	4	61 (10.7)
-	0.05706	16	109 (6.7)
DGTH-$\mathbb{P}_3$	0.05519	4	71 (15.0)
-	0.05519	16	139 (8.2)
DGTH-$\mathbb{P}_4$	0.05427	4	83 (19.5)
-	0.05527	16	170 (10.3)
DGTH-$\mathbb{P}_i$	0.05486	4	49 (7.2)
-	0.05491	16	81 (5.1)

[a] # iter classical/# iter optimized

directions for collaboration

Current implementation in 2D and 3D

- At the discrete level, Schur complement type system
  - Krylov solver for the interface system
  - Sparse direct solver (MUMPS) at the subdomain level
- Message passing programming using MPI

Potential collaboration topics in the framework of the joint laboratory

- On the methdological side

  AMG method for DG discretization of the frequency domain Maxwell equations

- On the computational side

  AMG used as a subdomain solver in the 3D case
  Hybrid programming for exploiting multiple parallelism levels
  (e.g. multi-threaded AMG)