

An assembled inexact Schur complement preconditioner



Joachim Schöberl

Computational Mathematics in Engineering
Institute for Analysis and Scientific Computing
Vienna University of Technology



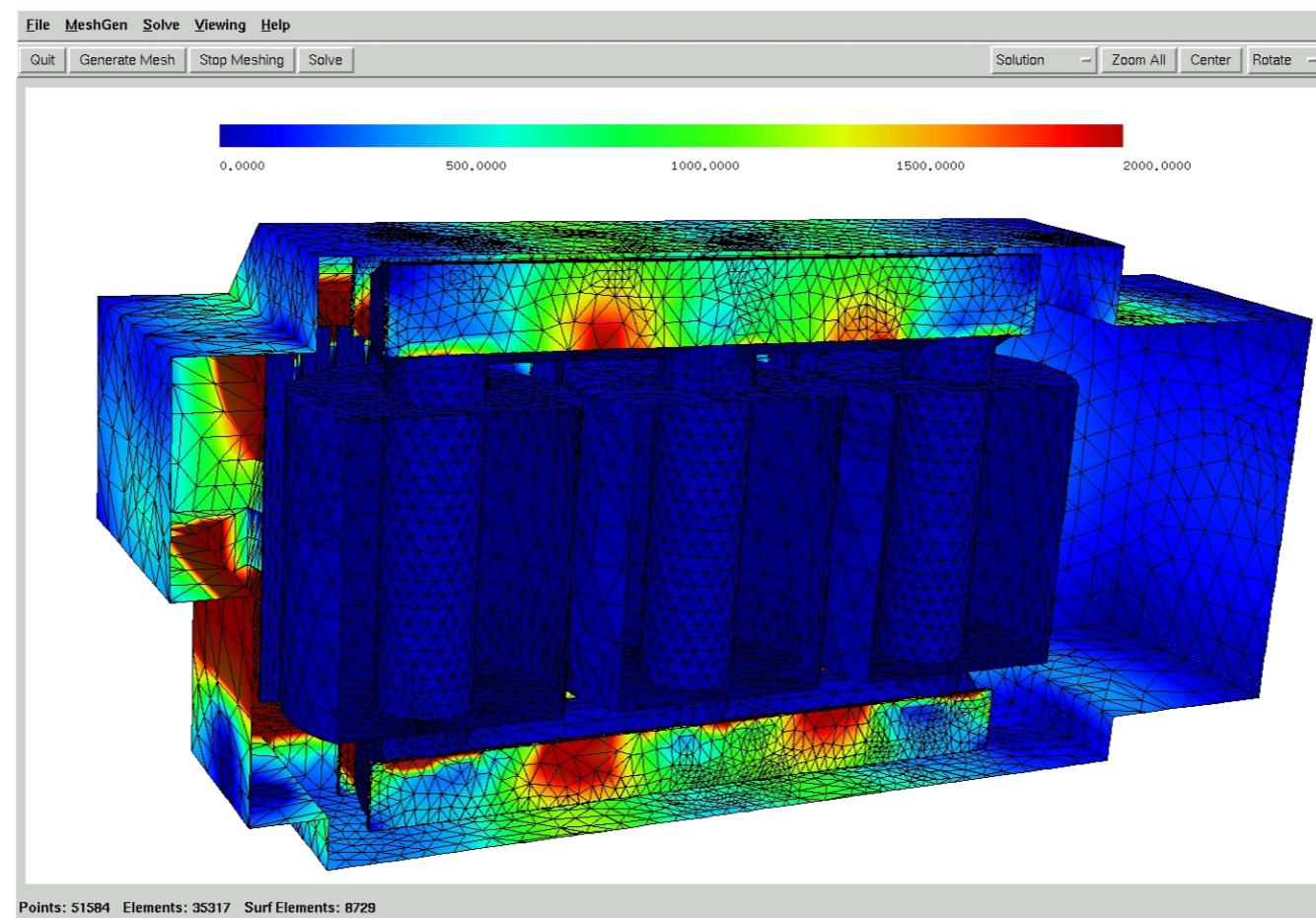
DD22, Lugano, September 2013

Netgen / NGsolve Software

- Mesh generator **Netgen**
 - automatic tetrahedral mesh generator
 - input geometry in CSG, STL, IGES / Step
 - special features for thin domains
 - arbitrary order curved elements
 - parallel mesh data structure
 - visualization
- Finite Element Solver **NGSolve**
 - high-order elements for scalar and vector fields
 - Elasticity, Maxwell, Navier-Stokes, ...
 - shared memory and/or distributed memory parallelization
 - intensive object oriented, compile-time polymorphism in C++11

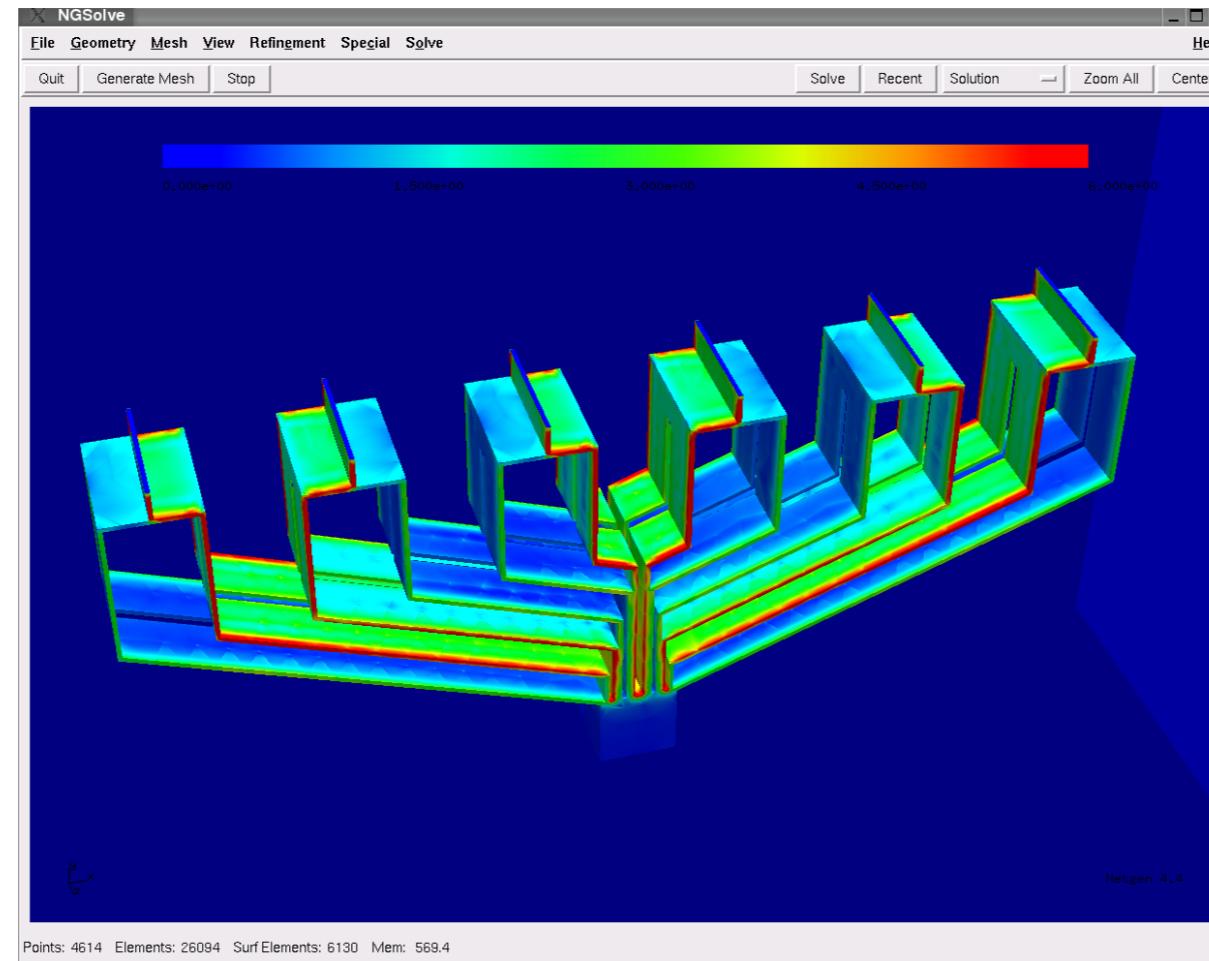
Open source (LGPL license), available from sourceforge

Magnetic field simulation in Power Transformers (with Siemens Austria)



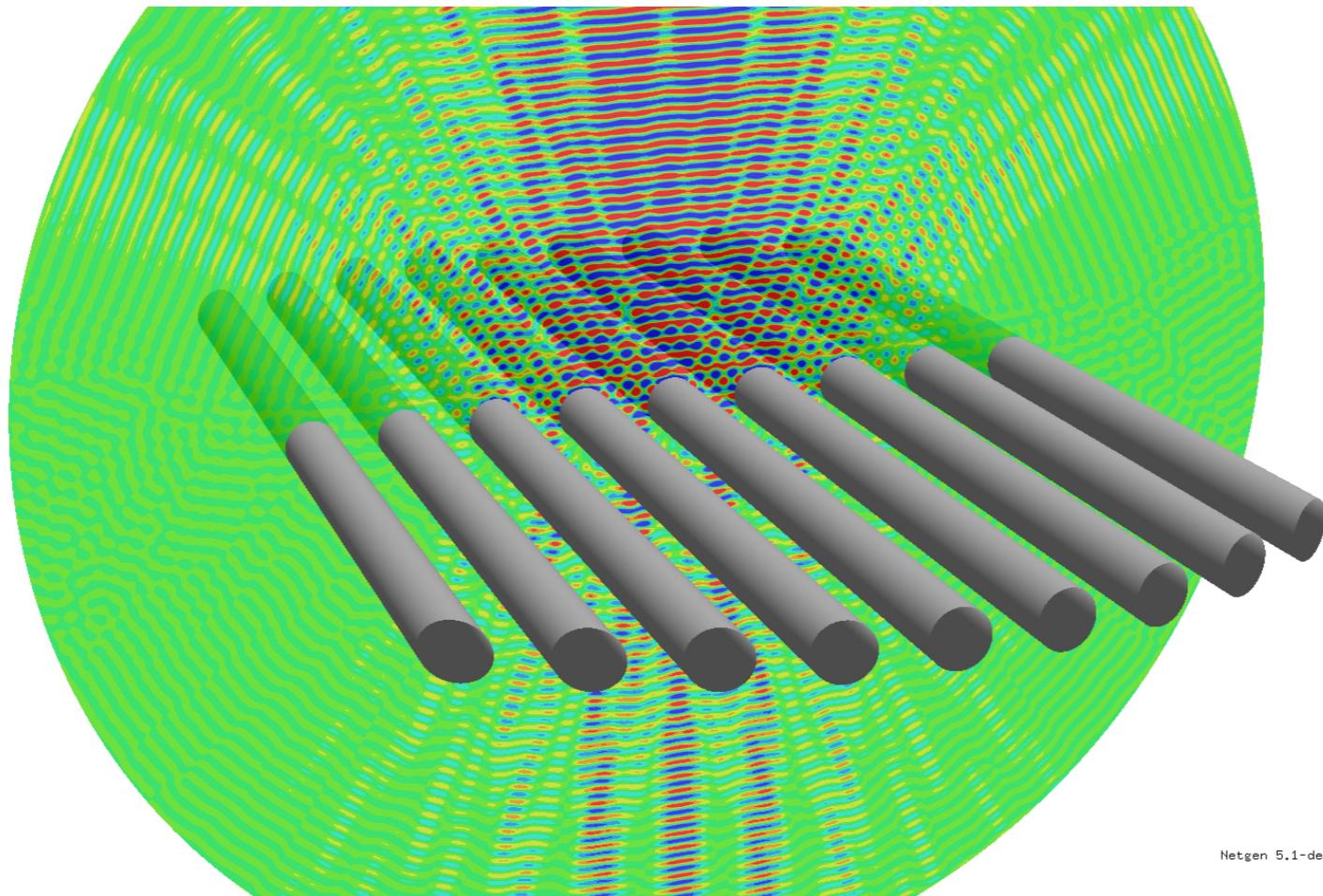
- Nonlinear material, anisotropic structures, homogenization in layered media, time harmonic
- Nedelec elements order 2 or 3, 1-3 M complex unknowns, 10-30 min on 8-core PC

Component simulation : Eddy currents in a bus bar



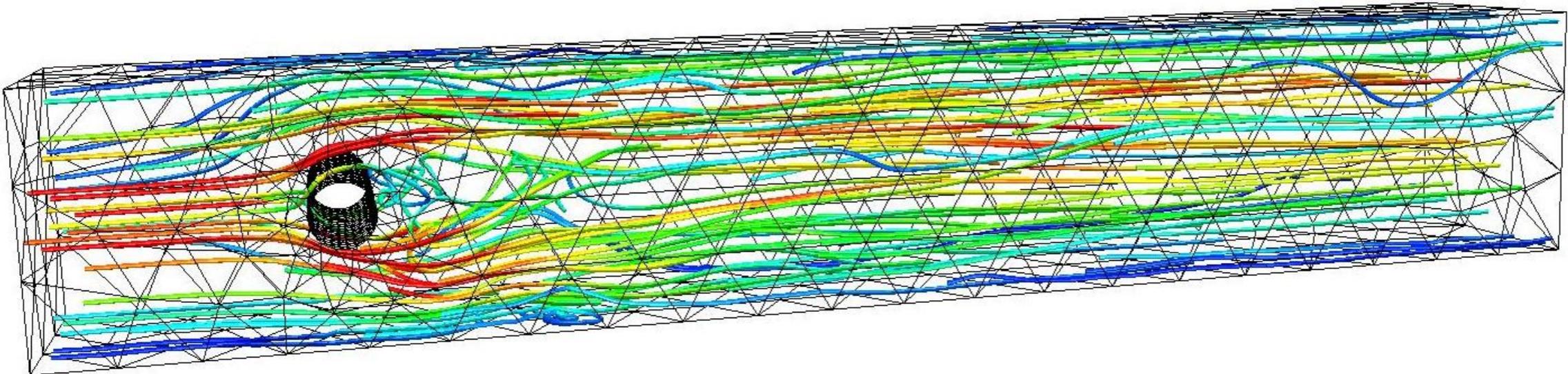
- Linear time harmonic simulation, isotropic mesh
- Order 3-5, 1-10 M complex unknowns, 50-100 k Unknowns per sec on 8-core PC

Acoustic and electromagnetic Scattering



- Helmholtz, hybrid DG, order 4, Robin-type domain decomposition
- $D = 80\lambda$, 536M complex unknowns, 2000 cores on VSC2, 413 its, 15 min [PhD Huber, 2013]

Incompressible Navier Stokes



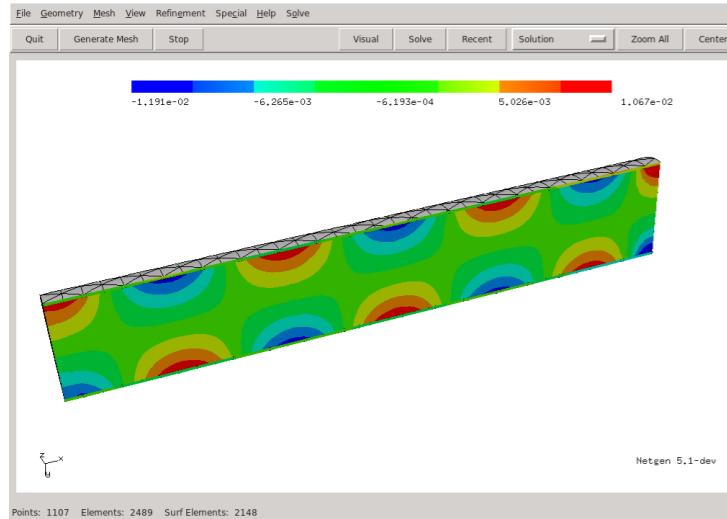
- Hybridized version of Cockburn-Kanschat-Schötzau DG
- Glowinski splitting, Stokes implicit / convection explicit on smaller time-steps
- order 4, 64 cores, 2M unknowns, 3 sec per time-step

[Thesis C. Lehrenfeld, 2010]

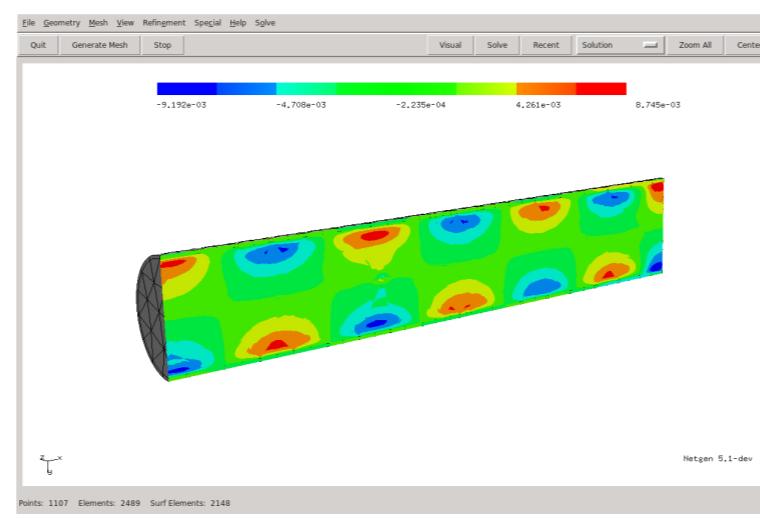
Coriolis mass flow meter (Endress-Hauser)

Monolithic fluid structure interaction

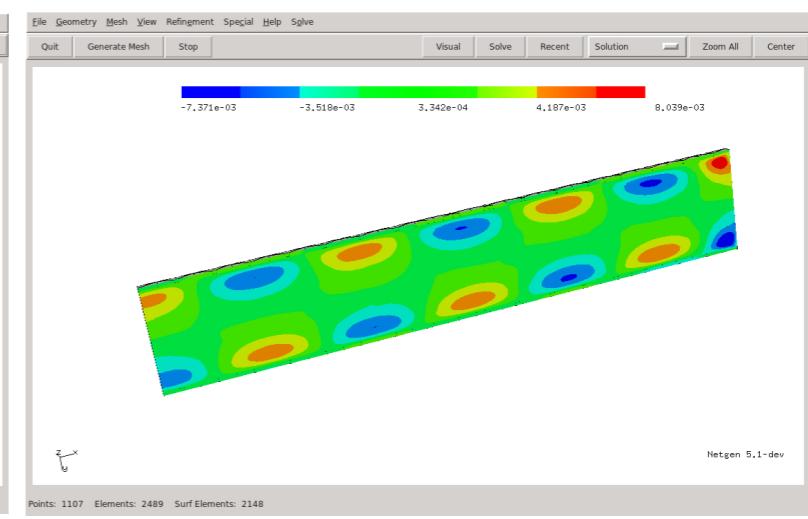
velocity x -component in fluid for different viscosities:



$$\nu = 0.1$$



$$\nu = 10$$

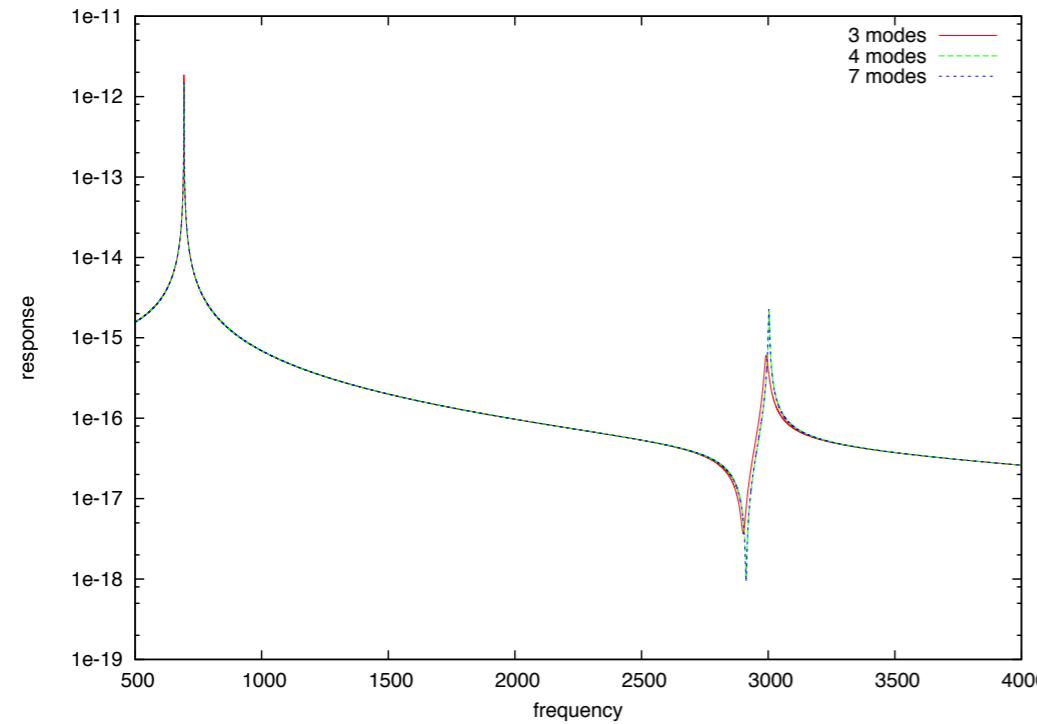


$$\nu = 100$$

- HDG in fluid: robustness for small viscosities (added mass limit)
- HDG in solid: robustness for thin 3D elements (anisotropic estimates [A.Sinwell+JS for TDNNS elements])

$p = 2 \dots 4$, $N = 200k \dots 500k$, Time = 20 sec ... 2 min

Reduced Basis Method (RBM)



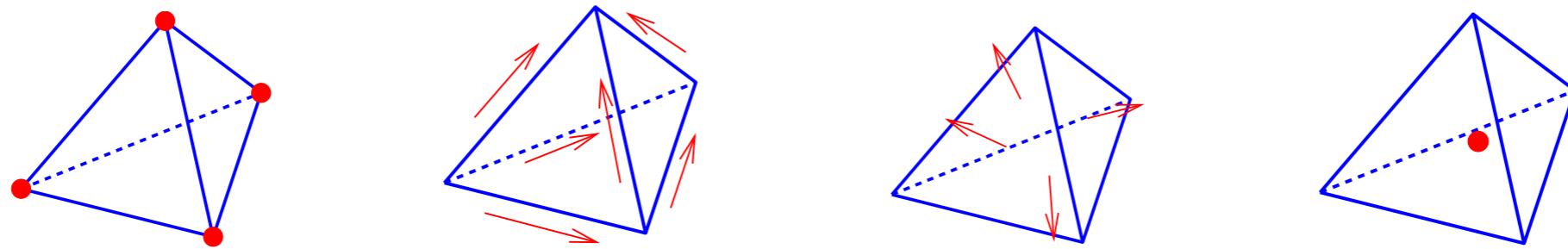
Eigenvalues from RBM:

modes	EV 1	EV 2
3	693 - i 0.136	2992 - i 6.62
4	693 - i 0.170	3003 - i 1.85
7	693 - i 0.166	3002 - i 1.73

with Anna Zechner

The de Rham Complex

$$\begin{array}{ccccccc}
 H^1 & \xrightarrow{\nabla} & H(\text{curl}) & \xrightarrow{\text{curl}} & H(\text{div}) & \xrightarrow{\text{div}} & L^2 \\
 \cup & & \cup & & \cup & & \cup \\
 W_h & \xrightarrow{\nabla} & V_h & \xrightarrow{\text{curl}} & Q_h & \xrightarrow{\text{div}} & S_h
 \end{array}$$



Used for constructing high order finite elements [JS+S. Zaglmayr, '05, Thesis Zaglmayr '06]

$$\begin{aligned}
 W_{hp} &= W_{\mathcal{L}_1} + \text{span}\{\varphi_{h.o.}^W\} \\
 V_{hp} &= V_{\mathcal{N}_0} + \text{span}\{\nabla \varphi_{h.o.}^W\} + \text{span}\{\varphi_{h.o.}^V\} \\
 Q_{hp} &= Q_{\mathcal{R}T_0} + \text{span}\{\text{curl } \varphi_{h.o.}^V\} + \text{span}\{\varphi_{h.o.}^Q\} \\
 S_{hp} &= S_{\mathcal{P}_0} + \text{span}\{\text{div } \varphi_{h.o.}^S\}
 \end{aligned}$$

Hybrid Discontinuous Galerkin (HDG) Method

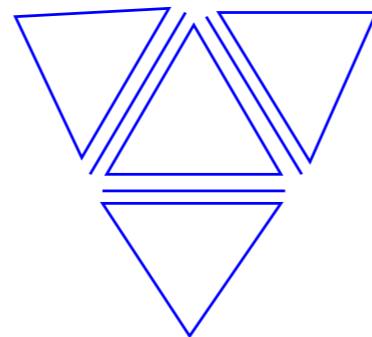
[Cockburn+Gopalakrishnan+Lazarov]

conforming mesh consisting of elements and facets (= edges in 2D and faces in 3D)

$$\mathcal{T} = \{T\} \quad \mathcal{F} = \{F\}$$

Approximate u with piece-wise polynomials on elements and additional polynomials on facets:

$$u_N \in P^p(\cup T) \quad \hat{u}_N \in P^p(\cup F)$$



$$\sum_T \left\{ \int_T \nabla u \nabla v - \int_{\partial T} \frac{\partial u}{\partial n} (v - \hat{v}) - \int_{\partial T} \frac{\partial v}{\partial n} (u - \hat{u}) + \frac{\alpha p^2}{h} \int_{\partial T} (u - \hat{u})(v - \hat{v}) \right\} = \int_{\Omega} f v$$

Current solvers in NGSolve

High-order / Low-order two level methods

- overlapping multiplicative Schwarz on vertex patches (sequential version)
- element-wise BDDC preconditioner (MPI-parallel version)

Coarse problem = low order system solved with direct solver (MKL-Pardiso, MUMPS)

High order space: $p = 3 \dots 5$, $N = 10^6 \dots 10^8$

Low order space: $N = 10^5 \dots 10^6$

Not scalable for fine meshes !

BDDC timings

- Poisson equation, HDG discretization, BDDC preconditioner
- Unstructured mesh of 496 K tets
- Coarse space: mean value on facets (1.01 M), solved with MUMPS, $\text{eps} = 10^{-8}$
- On 4x10 core server, 512 GB shared memory

order	N	its	time[sec]	N / (time × cores)
2	11.1 M	38	15.4	18 K
3	20.0 M	47	23.9	20 K
4	32.5 M	55	43.6	18 K
5	49.1 M	61	69.7	17 K
6	70.0 M	66	121	14 K
7	96.0 M	70	196	12 K

[JS - C. Lehrenfeld 2012] cond number estimate $O(\log^3 p)$

Matrix-free implementation

- Matrix application :

- Apply matrix vector product on the fly

- Use sum-factorization for high-order basis functions

- Costs per element-matrix-vector product: $O(p^4)$

- Preconditioner:

- Precompute local inverses and extension operators on the reference element, and apply them simultaneous

- Costs per element: $O(p^6)$, but good constant from one dominating BLAS3 operation.

Experiments for a 3D Model Problem

- Unit cube, 3D Poisson equation
- Unstructured mesh of 756 tetrahedral elements

p	N	matrix based its/sec	matrix free, 8 cores its/sec
4	57435	15 / 0.42	54 / 1.32
6	118636	20 / 1.62	61 / 3.32
8	210753	23 / 3.77	67 / 8.41
10	339834	26 / 7.85	71 / 15.36
12	511927	28 / 14.05	75 / 27.51
14	733080	-	78 / 47.50
16	1009341	-	81 / 76.44

Dell Precision T7500, 2 Intel Xeon QC at 2 GHz

Promising approach for current GPU development

Low order coarse grid preconditioner

Goals:

- robustness wrt coefficients
- robustness wrt unknowns
- sub-domains by automatic mesh partitioning
- unified method for elliptic PDEs

operation count is less an issue, since it is used only for the low order coarse grid

Method 1

Elliptic model problem: $-\Delta u = f$

Non-overlapping domain decomposition: $\overline{\Omega} = \cup_k \overline{\Omega}_k$

Sub-domain Schur complement: $S^{\partial\Omega_k} := A_{BB} - A_{BI}A_{II}^{-1}A_{IB}$

Compute Schur complements on **closed** edges E_i :

$$S^{E_i} := S_{E_i, E_i} - S_{E_i, \setminus E_i} S_{\setminus E_i, \setminus E_i}^{-1} S_{\setminus E_i, E_i}$$

$$\widehat{S}^{\partial\Omega_k} := \sum_{E_i \subset \partial\Omega_k} I_{E_i} S^{E_i} I_{E_i}^T$$

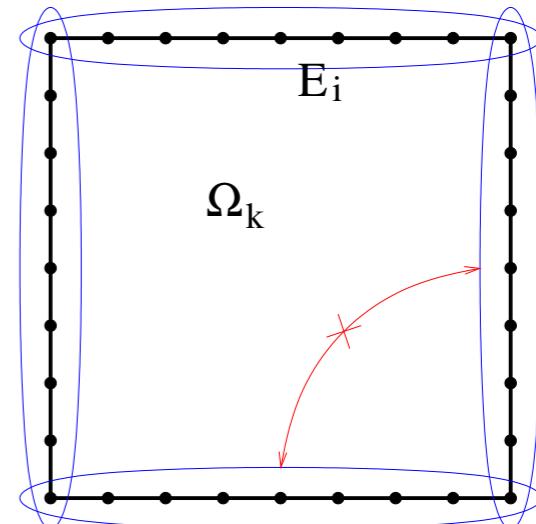
Claim:

$$\widehat{S} := \sum_{\Omega_k} \widehat{S}^{\partial\Omega_k}$$

is a good (i.e. cheap+low cond number) preconditioner for the global Schur complement $S := \sum_{\Omega_k} S^{\partial\Omega_k}$

references ???

3D case: wait 5 slides,



Context

- Classical substructuring methods
Looks similar to Bramble-Pasciak-Schatz, Smith, BJORSTAD, Widlund, ... but it differs
- AMGe (element-wise algebraic multigrid)
P. Vassilevsky et al: small patches, recursively
- BDDC - deluxe scaling:
Dohrmann-Widlund: add up interface Schur complements before inverting, different coarse space
- Not a main-stream method treated in the domain decomposition textbooks (Smith-BJorstad-Gropp, Quarteroni-Valli, Toselli-Widlund, Pechstein)
- direct Schur complement computation was too expensive in the 80s.
Local flops are free - philosophy is relatively young.

Algorithmic pattern

1. compute local Schur complements (e.g. using MUMPS)
2. compute edge-wise Schur complements:

$$S^{E_i} = ((S^{-1})_{E_i, E_i})^{-1}$$

if S singular: fix one node, and use low-rank update formulas

3. accumulation of edge-Schur complements between neighbouring sub-domains (direct neighbour communication)
4. local elimination of interior edge nodes \Rightarrow vertex Schur complement, i.e., the coarse grid system
5. global inversion of global vertex system

Analysis

Estimates on sub-domains (good sub-domain shape, good element aspect ratio, locally bounded jumps)

sub-domain Schur complement:

$$u^T S^{\partial\Omega_k} u \simeq |u|_{H^{1/2}(\partial\Omega_k)}^2$$

Schur complement on one edge:

$$u^T S^{E_i} u \simeq |u|_{H^{1/2}(E_i)}^2$$

Energy-norm of assembled inexact Schur complements:

$$u^T \widehat{S}^{\partial\Omega_k} u \simeq \sum_{E_i \subset \partial\Omega_k} |u|_{H^{1/2}(E)}^2$$

Inverse estimates on finite element space:

$$\sum_E |u|_{H^{1/2}(E)}^2 \preceq \|u\|_{H^{1/2}(\partial\Omega_k)}^2 \preceq \log^2(p/h) \sum_E |u|_{H^{1/2}(E)}^2$$

and thus

$$\widehat{S}^{\Omega_k} \preceq S^{\Omega_k} \preceq \log^2(p/h) \widehat{S}^{\Omega_k}$$

Hybrid Discontinuous Galerkin methods

interface basis-functions are piece-wise discontinuous polynomials on edges.

Similar structure: non-conforming P^1 elements

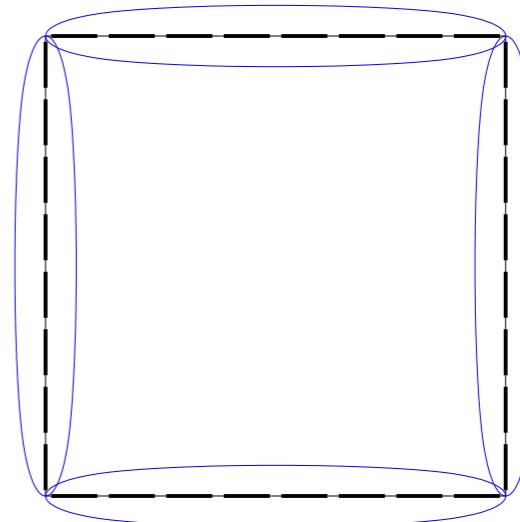
assembled Schur complement matrix

$$\widehat{S}^{\partial\Omega_k} = \sum_{E_i} S^{E_i}$$

is now block-diagonal.

$$\dim \text{Kern}(\widehat{S}^{\partial\Omega_k}) = 4 > \dim \text{Kern } S^{\partial\Omega_k} = 1$$

System falls appart, no coarse system



Augmented system - rank 1 stabilization

Let $b^{\partial\Omega_k}(u)$ be the (some) mean value on the sub-domain. Then solving

$$\text{find } u_h \in V_h : \quad A(u_h, v_h) = f(v_h) \quad \forall v_h$$

is equivalent to

$$\text{find } (u_h, \bar{u}) \in V_h \times \mathbb{R}^M : \quad A(u_h, v_h) + \sum_{k=1}^M (b(u_h) - \bar{u}_k)(b(v_h) - \bar{v}_k) = f(v_h) \quad \forall (v_h, \bar{v}) \in V_h \times \mathbb{R}^M$$

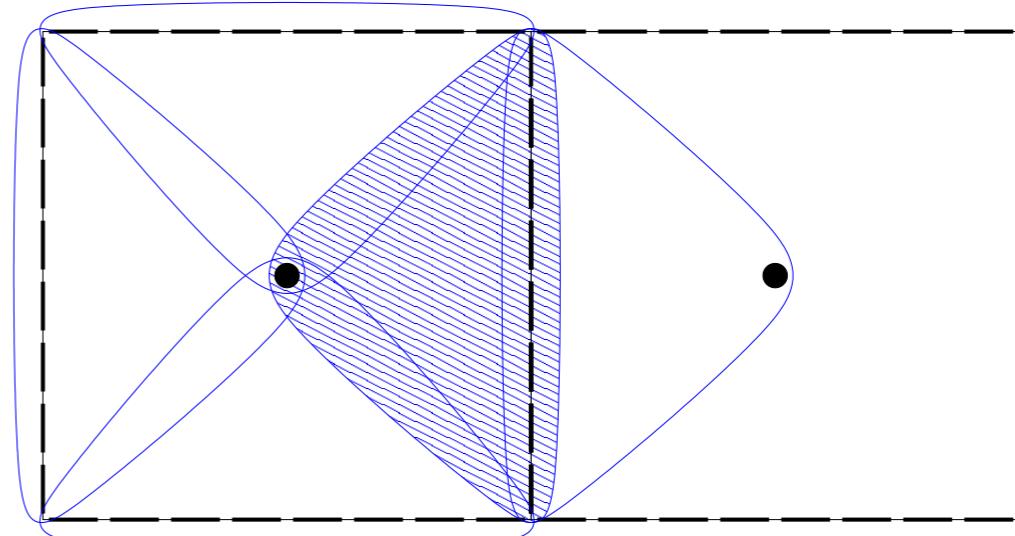
Form Schur complements w.r.t.

$$E_i \times \mathbb{R}$$

Global system has block-diagonal edge-block.

Coupling to mean values gives global system, 1 variable per sub-domain

Same for $n = 2, 3, \dots$

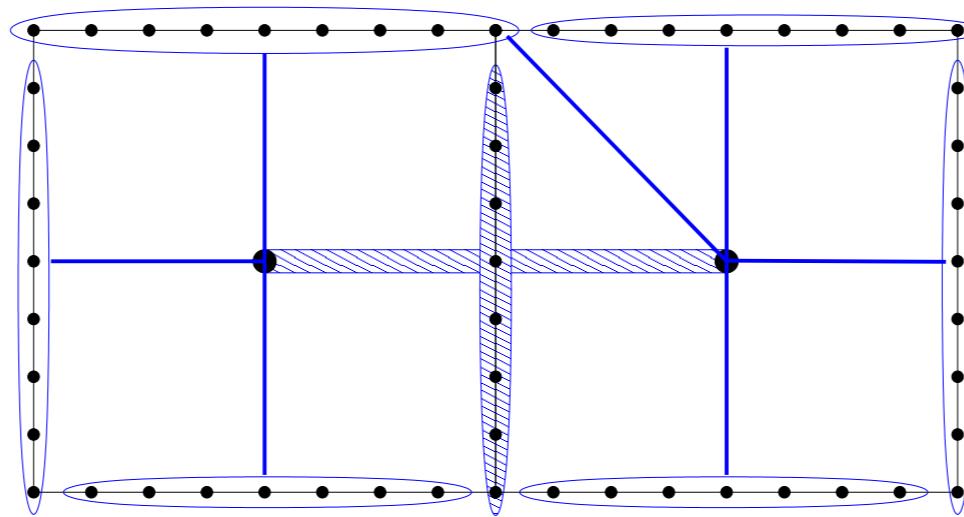


Augmented system for nodal elements

Same can be done for standard nodal elements:

Non-overlapping (compatible !) splitting of boundary nodes:

2 nodes belong to the same block iff lowest 2 sub-domain numbers they belong to are the same



All sub-domain boundary nodes can be locally eliminated. Global system consists of sub-domain mean values.

This is also a good method for 3D.

Computational results

Unit cube, tetrahedral mesh distributed by METIS, polynomial order 3

high-order/low-order BDDC combined with inexact assembled schur complements

error reduction by 10^{-8}

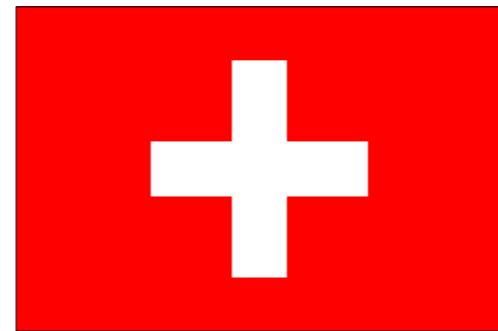
NDOF	14 929 191				123 574 001				
	proc	Nloc	its	T_{solve}	T_{tot}	Nloc	its	T_{solve}	T_{tot}
16	1033961	53	71.4	163.5					
32	510226	55	43.5	79.6					
64	253862	56	22.0	37.6					
128	128969	58	11.5	19.5					
256	65499	59	7.7	12.8	516660	72	63.7	111.2	
512	33872	58	6.1	9.9	261410	74	31.6	52.0	
1024					132876	73	17.7	28.5	
2048					67815	74	20.7	29.6	

On Vienna-Scientific-Cluster 2 (VSC2)

Nodes contain two 8-coreAMD Opteron 6132 HE, 2.2 GHz, 32 GB RAM

16 processes per node.

No miracle for strongly varying coefficients



- strongly varying coefficient inside \Rightarrow no problem
- one strong channel through domain \Rightarrow one bad eigenvalue

A bad function for the strong channel domain is $u = x$

$$\|u\|_S^2 = \min_{\substack{w \in V_h(\Omega_k) \\ w=u \text{ on } \partial\Omega_k}} \|w\|_A^2 \simeq a_{\text{large}}$$

$$\|u\|_{\widehat{S}}^2 = \sum_{E_i \subset \partial\Omega_k} \min_{\substack{w \in V_h(\Omega_k) \\ w=u \text{ on } E_i}} \|w\|_A^2 \simeq a_{\text{small}}$$

Local fix of bad eigenvalues

Solve sub-domain eigenvalue problems (completely local)

$$S^{\partial\Omega_k} u = \lambda \hat{S}^{\partial\Omega_k} u$$

Generic lower bound :

$$\lambda \geq \frac{1}{\text{number of facets}}$$

Upper bound may become large.

Low rank correction of preconditioner:

$$\widehat{\hat{S}} := \hat{S} + BC^{-1}B^T$$

with $B = \text{col}\{Du_j\}$ and $C = \text{diag}\{\lambda_j^{-1}\}$ for bad (large) eigenpairs (u_j, λ_j) .

Dualization

Solve

$$(\widehat{S} + BC^{-1}B^T)u = f$$

via new variables (Lagrange parameters):

$$p = C^{-1}B^T u$$

saddle point system with regular lower right block:

$$\begin{pmatrix} A & B^T \\ B & -C \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}$$

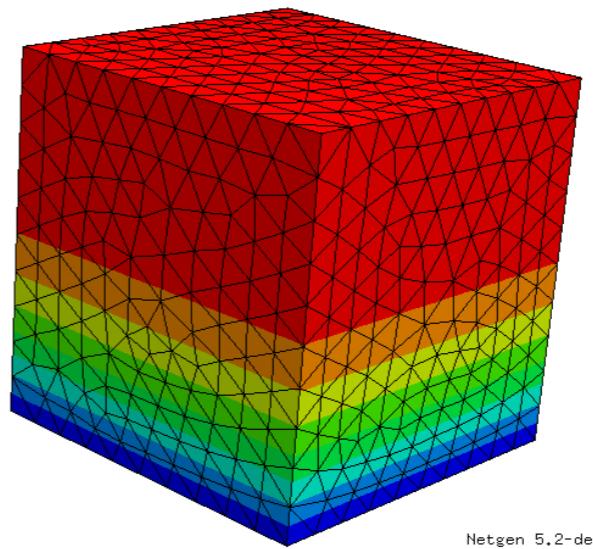
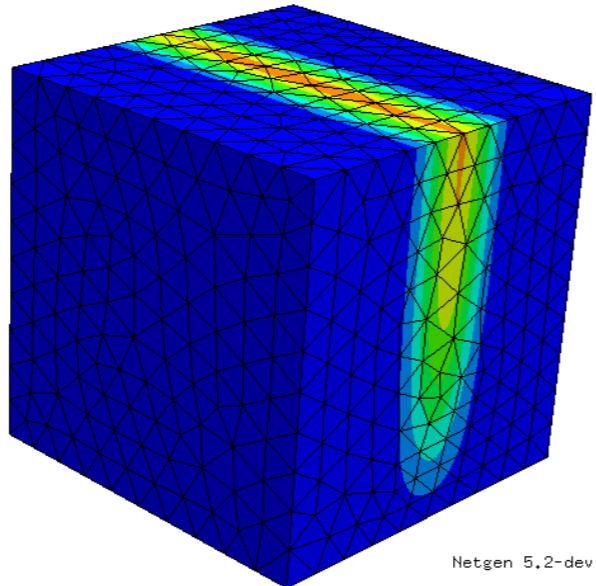
local elimination of facet variables leads to global saddle-point system

Physical meaning of Lagrange parameter: flux through channel

Anisotropic diffusion

$\Omega = (0, 1)^3$, $A(u, v) = \int_{\Omega} a \nabla u \nabla v$ with tensor $a = \text{diag} \{1, 1, 10^{-6}\}$, $f = \exp(-100|x - x_c|^2)$ and space (a) $V = \{v \in H^1 : v = 0 \text{ for } x = 0\}$ and (b) $V = \{v \in H^1 : v = 0 \text{ for } z = 0\}$.

order = 2, ndof = 10 365.



$\lambda_{\min}/\lambda_{\max}$	no local fix	local fix with 19 ev per sub-domain (average)
Case (a)	0.42 / 46.5	0.3 / 5.18
Case (b)	0.11 / 763	0.4 / 5.2

$H(\text{div})$ -case

$$A(u, v) = \alpha (u, v)_{L_2} + \beta (\operatorname{div} u, \operatorname{div} v)_{L_2} \quad \beta \gg \alpha$$

Sub-domain Schur complement:

$$\|u\|_S^2 \simeq \alpha \|u_n\|_{H^{-1/2}(\partial\Omega_k)}^2 + \beta \left(\int_{\partial\Omega_k} u_n \right)^2$$

Assembled facet Schur complements:

$$\|u\|_{\widehat{S}}^2 \simeq \alpha \sum_{F \subset \partial\Omega_k} \|u_n\|_{H^{-1/2}(F)}^2$$

Exactly 1 large eigenvalue per sub-domain. Treated via dualization.

Coarse space: mean value of dual variable (scalar) per sub-domain

$H(\text{curl})$ -case

(work in progress)

$$A(u, v) = \alpha(u, v) + \beta(\text{curl } u, \text{curl } v)$$

Sub-domain Schur complement:

$$\|u\|_S^2 \simeq \alpha \|u_\tau\|_{H^{-1/2}}^2 + \beta \|\text{curl}_\tau u\|_{H^{-1/2}}^2$$

Non-overlapping boundary decomposition leads to many bad eigenvalues.

- Method 1: Keep all edges on sub-domain edges in the coarse space
- Method 2: 2-step inexact Schur complements on faces

Norm on face-boundaries is (up to log-factors)

$$\|u\|_{\partial F}^2 \simeq \alpha \|u_\tau\|_{H^{-1}(\partial F)}^2 + \beta \left(\int_{\partial F} u_\tau \right)^2$$

Breaking into edges gives 1 bad ev $O(\beta)$, and number of edges bad ev $O(h^{-1})$. Dualize or live with it.

Conclusions

- Overview of recent high-order FEM algorithmic development in NGSolve
- Results on High-order / Low-order two level preconditioning
- Discussion of a simple algebraic primal substructuring algorithm

Assembled Inexact Schur Complement Preconditioner

or

Assembled Schur Complement Preconditioner ?