Improving AMG interpolation through energy minimization

Due Giorni di Algebra Lineare Numerica e Applicazioni
14-15 Febbraio 2022, Centro Congressi Federico II, via Partenope, Napoli, Italy

Carlo Janna, Andrea Franceschini and Giovanni Isotton
carlo.janna@unipd.it

February 15, 2022

Università
degli Studi
di Padova

Outline

Motivation

AMG is a very popular solution algorithm which is used in several applications:

- it is very flexible as it only requires little information beyond the matrix itself;
- it can be parallelized up to millions of cores;
- as preconditioner for a Krylov method often guarantees performance independent of the grid size;

AMG works almost perfectly for Poisson problems, however, it is less effective in more difficult problems such as those arising in structural mechanics:

- classical or extended interpolation usually give poor results;
- aggregation-based AMG is generally preferred over Classical AMG;
- the near kernel space is larger than the standard constant vector (which is fine for Poisson);

Motivation

In the context of structural problems (or other difficult problems):

- more powerful smoothers than Jacobi or Gauss-Seidel are needed;
- the near kernel is constructed starting from the Rigid Body Modes (RBM);
- it is often necessary to improve the near kernel through smoothing or an eigensolution;
- prolongation is constructed through a least square fit of the test space;

Main issues:

- coarse basis functions with local support are not orthogonal to high frequency modes;
- the prolongation has to be smoothed to allow for fast convergence (helpful in calssical AMG as well);
- smoothed prolongation however leads to high complexities;

Energy minimization in prolongation set-up

The main idea is, for a given prolongation pattern $\mathcal{P}$, to compute:

$$P = \underset{P \in \mathcal{P}}{\operatorname{argmin}} \left(\operatorname{tr}(P^T A P) \right)$$

or, equivalently, to minimize the energy of every prolongation column:

$$p_i^T A p_i \rightarrow \min \qquad \forall i \in C$$

We consider a **Classical** AMG framework, so that:

$$A = \begin{bmatrix} A_{ff} & A_{fc} \\ A_{fc}^T & A_{cc} \end{bmatrix} \qquad P = \begin{bmatrix} W \\ I \end{bmatrix} \qquad R = \begin{bmatrix} 0 & I \end{bmatrix} \qquad S = \begin{bmatrix} I \\ 0 \end{bmatrix}$$

1. Olson, L. N., Schroder, J. B. & Tuminaro, R. S., *A General Interpolation Strategy for Algebraic Multigrid Using Energy Minimization*, **SIAM Journal on Scientific Computing 33**, pp. 966–991 (2011).
2. Manteuffel, T. A., Olson, L. N., Schroder, J. B. & Southworth, B. S., *A Root-Node-Based Algebraic Multigrid Method*, **SIAM Journal on Scientific Computing 39**, S723–S756 (2017).

Why Energy minimization is important

Suppose we have a C/F partition of the unknowns and an SPD X matrix which is spectrally equivalent to the symmetrized smoother $\widetilde{M} = M^T(M + M^T - A)^{-1}M$.

Defining $A_s = S^T AS$ and $X_s = S^T XS$ and:

$$k_S \leq \lambda_{min}(X_S^{-1}A_S) \leq \lambda_{max}(X_S^{-1}A_S) \leq c_2$$

then, if $\|PR\|_A$ is bounded:

$$K_{TG} \leq \frac{c_2}{k_S}\|PR\|_A^2 \qquad \text{where } K_{TG} \text{ is such that} \|E_{TG}\|_A \leq 1 - \frac{1}{K_{TG}}$$

In practice, **the smaller $\|PR\|_A$, the faster convergence!**

Falgout R.D. and Vassilevski P. S., *On Generalizing the Algebraic Multigrid Framework*, **SIAM Journal on Numerical Analysis 42**, pp. 1669–1693 (2004).

The minimum in energy can be simply found by deriving with respect to W and setting to zero:

$$\frac{\partial}{\partial W}\left(\text{tr}(P^T A P)\right) = 0 \Rightarrow A_{ff} W = -A_{fc}$$

which is exactly what is defined *ideal* prolongation $W_{id} = -A_{ff}^{-1} A_{fc}$.

Note that it is also the target that is sought with a 2x2 block FSAI approximation:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22} \end{bmatrix}, \quad G = \begin{bmatrix} I & 0 \\ F & I \end{bmatrix}, \quad GAG^T \simeq I \qquad \Rightarrow \qquad F = -A_{11} A_{12}$$

Janna, C., Ferronato, M. & Gambolati, G., *A Block FSAI-ILU Parallel Preconditioner for Symmetric Positive Definite Linear Systems*, **SIAM Journal on Scientific Computing 32**, pp. 2468–2484 (2010).

Adding constraints to energy minimization

First of all, a dense prolongation is impractical, so we enforce a given non-zero pattern. We consider the pattern $\overline{P}_0$ of a tentative prolongation P_0, and we extend it using the strength of connection matrix S:

$$\overline{P} = S^k \overline{P}_0$$

for a small power k.

Then we minimize the energy of every column of P restricted to the non-zero pattern $\overline{P}$ by solving the following sequence of n_c dense linear systems:

$$A(\mathcal{I}_i, \mathcal{I}_i)\widetilde{p}_i = -A(\mathcal{I}_i, i) \qquad \forall i \in C$$

with $\mathcal{I}_i$ the set of indices of the non-zeroes in the i-th column of $\overline{P}$.

Another important condition to satisfy is to ensure that the prolongation is able to accurately prepresent some prototypes of the near kernel of A. By denoting with V the matrix collecting these prototypes, this condition reads:

$$V \subseteq \text{range}(P), \qquad \text{with} \qquad V = \begin{bmatrix} V_f \\ V_c \end{bmatrix}$$

This condition becomes:

$$W\, V_c = V_f \qquad \Rightarrow \qquad V_c(\mathcal{J}_i, :)^T \widetilde{w}_i = v_i \qquad \forall i \in \mathcal{F}$$

where $\mathcal{J}_i$ is the set of column indices of the prescribed nonzeros of the i-th row of W. Note that this is another sequence of n_f small and dense, generally **underdetermined**, linear systems.

By collecting all the non-zero entries of W into the vector w and using Lagrange multipliers, the problem can be stated as the solution of the following saddle-point system:

$$\begin{bmatrix} K & B \\ B^T & 0 \end{bmatrix} \begin{bmatrix} w \\ \lambda \end{bmatrix} = \begin{bmatrix} f \\ g \end{bmatrix}$$

Observations:

- The size of K is typically much larger than the size of A (about 20 times larger in 3D mechanics);
- With a proper unknown ordering both K and B are block diagonal matrices;
- If w is numbered following the columns of P then K is block diagonal;
- If w is numbered following the rows of P then B is block diagonal;
- Unfortunately, they cannot be block diagonal at the same time;

The solution algorithm

As in the original paper, we solve the minimization problem through a **restricted conjugate gradient**, which is a nullspace method.
Define the projection orthogonal to B, $\Pi_B = I - B(B^T B)^{-1} B^T$, so that:

$$B^T y = 0 \qquad \forall y = \Pi_B v$$

Then, starting from a tentative prolongation W_0 already satisfying the constraint, $B^T w_0 = g$, we apply the **conjugate gradient** method to the system:

$$\Pi_B K \Pi_B \Delta w = \Pi_B (f - K w_0)$$

to find a correction ΔW to be applied to W_0:

$$W = W_0 + \Delta W$$

The solution algorithm

During CG iteration, it is possible to monitor the decrease in energy and stop when a plateu is reached.

Some issues related to the proposed solution methd:

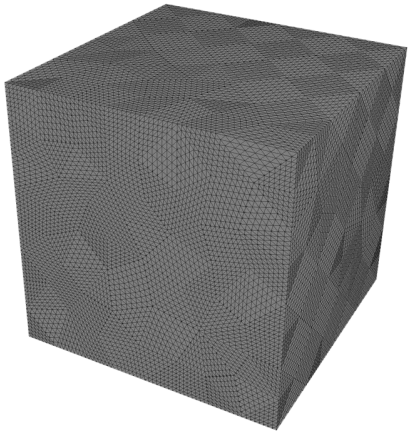
- The energy minimization is just a piece of the AMG set-up, so we cannot spend to much effort in it $\Rightarrow$ it must converge quickly!
- Memory occupation is a big issue. K and B require a significant amount of memory and sometimes it is unavoidable to proceed in matrix-free mode;
- Preconditioning is crucial:
  - projected Jacobi, $\Pi_B D_K \Pi_B$, is enough for many problems;
  - incomplete Cholesky with very low fill-in converges faster but it is more difficult to apply, if K cannot be stored;

Linear elasticity Benchmark:
Regular discretization of an elastic cube with tetrahedral FE.

Matrix	# of rows	# of non-zeroes
C_4820	14,460	556,686
C_35199	105,597	4,079,357
C_246389	739,167	29,610,351
C_1772489	5,317,443	222,268,213

We compare the following prolongation strategies:

- **EXTI** Extended+i (hypre);
- **BAMG** least squares fit;
- **SBAMG** BAMG + smoothing;
- **EMIN** energy minimization;



Numerical results

Matrix	P-type	C_{gd}/C_{op}	n_{it}
C_4820	EXTI	1.062/1.200	91
	BAMG	1.062/1.162	25
	SBAMG	1.062/1.404	17
	EMIN	1.062/1.383	12
C_35199	EXTI	1.068/1.254	209
	BAMG	1.067/1.211	65
	SBAMG	1.066/1.576	33
	EMIN	1.066/1.540	15
C_246389	EXTI	1.066/1.214	510
	BAMG	1.064/1.196	146
	SBAMG	1.064/1.608	55
	EMIN	1.064/1.517	22
C_1772481	EXTI	1.064/1.188	1286
	BAMG	1.062/1.189	270
	SBAMG	1.062/1.590	78
	EMIN	1.062/1.460	33

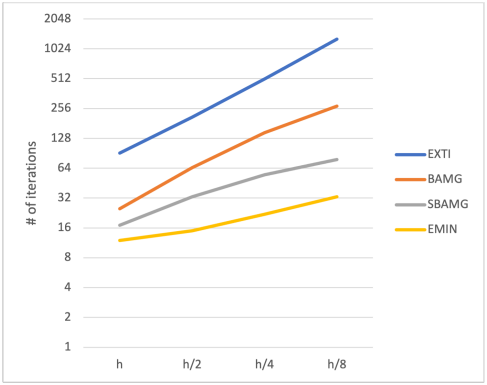


Figure: # of PCG iterations vs. grid size

Numerical results

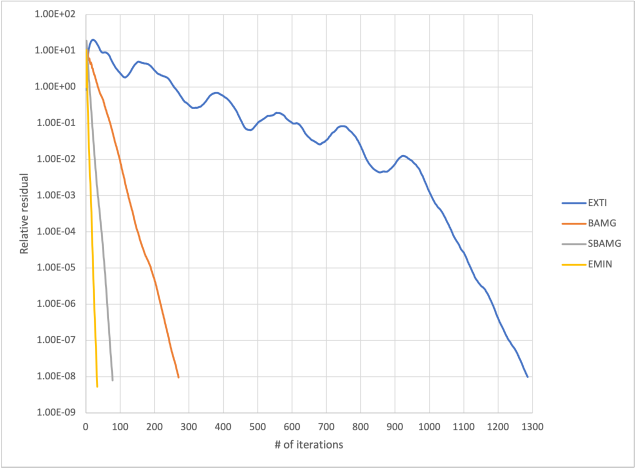


Figure: Relative residual, $\frac{r_k}{r_0}$, vs. iteration count, k

Numerical results

Matrix	# of rows	# of non-zeroes
Pflow_742k	742,793	37,138,461
Gear_167k	167,460	9,861,437
S4A_700k	631,007	26,148,363
Mech_447k	447,703	18,243,793

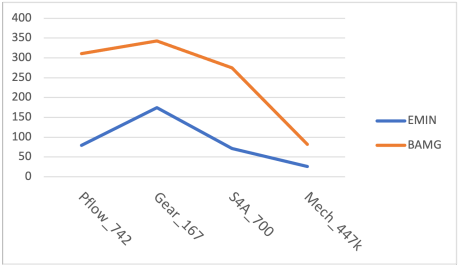


Figure: # of PCG iterations.

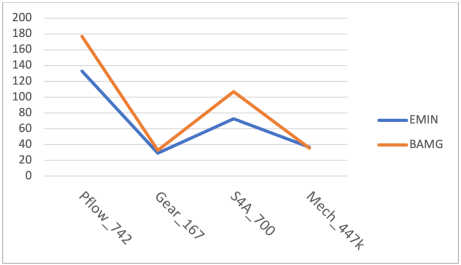


Figure: Total solution time [s].

Conclusions

- Energy minimizing prolongation is a well-known concept in AMG, which, however has little application due to its cost and its difficult implementation;
- Our main contribution was to make energy minimization feasible by means of preconditioning and an effective implementation;
- Our algorithm has proved very effective on linear systems arising from mechanical problems and also in challenging problems arising from different applications.
- Though requiring a slightly more expensive set-up, this additional cost is largely offset by a fast convergence.

Further work:

- Further improve preconditioning of restricted CG;
- Include energy minimization in the HPC AMG solver Chronos.

G. Isotton, M. Frigo, N. Spiezia and C. Janna, *Chronos: A general purpose classical AMG solver for high performance computing*, **SIAM Journal on Scientific Computing 43**, C335-C357, 2021.

Thanks for your attention.

Questions?