

## Direct solution of larger coupled sparse/dense linear systems using low-rank compression on single-node multi-core machines in an industrial context

36<sup>th</sup> IEEE International Parallel and Distributed Processing symposium

Emmanuel Agullo, <u>Marek Felšöci</u>, Guillaume Sylvand May 30 - June 3, 2022

Inria Bordeaux Sud-Ouest, France Team-project CONCACE

## Introduction

## Industrial context

# AIRBUS

- study the propagation of sound waves emitted by an aircraft
  - acoustic pollution reduction, prototype certification
- discrete model for numerical simulations
  - volume domain v (jet flow)
    - Finite Elements Method (FEM) [10, 8]
  - surface domain **s** (surface of the aircraft and the volume domain)
    - Boundary Elements Method (BEM) [5, 12]



An acoustic wave (blue arrow) emitted by the aircraft's engine, reflected on the wing and crossing the jet flow. Real-life case [11] (left) and a numerical model example (right). Global linear system coupling [6, 7] the FEM and the BEM unknowns:

$$\left[\begin{array}{ccc} A_{vv} & A_{sv}^{\mathsf{T}} \\ A_{sv} & A_{ss} \end{array}\right] \times \left[\begin{array}{c} x_{v} \\ x_{s} \end{array}\right] = \left[\begin{array}{c} b_{v} \\ b_{s} \end{array}\right]$$



• symmetric coefficient matrices:

- sparse parts volume domain v discretization with FEM (A<sub>vv</sub>), surface/volume domain interaction (A<sub>sv</sub>)
- a dense part surface domain sind discretization with BEM (A<sub>ss</sub>)
- $\bullet \ \ \text{finer} \ \ \text{model} \ \rightarrow \ \ \text{larger} \ \ \text{system}$
- direct solution using Schur complement [13]



### **Direct solution**

#### Schur complement

 $\bullet\,$  reduce the problem on boundaries  $\rightarrow\,$  simplify the system to solve

$$\begin{array}{c|c} R_{1} \\ R_{2} \end{array} \begin{bmatrix} A_{vv} & A_{sv}^{T} \\ A_{sv} & A_{ss} \end{bmatrix} \times \begin{bmatrix} x_{v} \\ x_{s} \end{bmatrix} = \begin{bmatrix} b_{v} \\ b_{s} \end{bmatrix}$$

#### **Computation steps**

1. eliminate  $x_v$  from the second equation  $\rightarrow$  Schur complement *S* 

$$\begin{array}{c} R_{\mathbf{1}} \\ R_{\mathbf{2}} \leftarrow R_{\mathbf{2}} - A_{sv}A_{vv}^{-1} \times R_{\mathbf{1}} \\ \begin{bmatrix} A_{vv} & A_{sv}^{T} \\ 0 & \underbrace{A_{ss} - A_{sv}A_{vv}^{-1}A_{sv}^{T}}_{S} \end{bmatrix} \times \begin{bmatrix} x_{v} \\ x_{s} \end{bmatrix} = \begin{bmatrix} b_{v} \\ b_{s} - A_{sv}A_{vv}^{-1}b_{v} \end{bmatrix}$$

2. solve the reduced Schur complement system

$$Sx_s = b_s - A_{sv}A_{vv}^{-1}b_v$$

3. determine  $x_v$  using  $x_s$ 

$$x_{v} = A_{vv}^{-1}(b_{v} - A_{sv}^{T}x_{s})$$

## Numerical computation

#### Properties of the input linear system

- A<sub>vv</sub> and A<sub>ss</sub> are symmetric
- $A_{vv}$  and  $A_{sv}$  are sparse

Ideal computation of  $S = A_{ss} - A_{sv}A_{vv}^{-1}A_{vs}$ 

• symmetric factorization of  $A_{vv} \rightarrow L_{vv}L_{vv}^T$ : <u>fill-in</u>

$$S = A_{ss} - A_{sv} (L_{vv} L_{vv}^T)^{-1} A_{sv}^T$$

• computation of the Schur complement

$$S = A_{ss} - \underbrace{(A_{sv}(L_{vv}^{T})^{-1})}_{\text{triangular solute}} \underbrace{(A_{sv}(L_{vv}^{T})^{-1})^{T}}_{\text{implicitly known}}$$

triangular solve implicitly known



Initial state of A



A after computing S

- coupling of a sparse direct and a dense direct solver
  - fully-featured community solvers implement appealing functionalities we want to take advantage of
    - low-rank compression, out-of-core computation, distributed memory parallelism
- two different schemes depending on the choice of sparse solver's building blocks
  - baseline coupling and advanced coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



#### baseline coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- A as a whole
- sparse facto.+Schur
- dense facto., dense solve



#### baseline coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- A as a whole
- sparse facto.+Schur
- dense facto., dense solve



#### baseline coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- A as a whole
- sparse facto.+Schur
- dense facto., dense solve



#### baseline coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- A as a whole
- sparse facto.+Schur
- dense facto., dense solve



#### baseline coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- *S* non-compressed, dense
- $A_{sv}^T$  explicitly stored, dense

- A as a whole
- sparse facto.+Schur
- dense facto., dense solve



#### baseline coupling

- separate  $A_{vv}$ ,  $A_{sv}$  and  $A_{ss}$
- sparse facto., sparse solve
- dense facto., dense solve



- *S* non-compressed, dense
- $A_{sv}^T$  explicitly stored, dense

#### advanced coupling

- A as a whole
- sparse facto.+Schur
- dense facto., dense solve



• *S* non-compressed, dense

## Contribution

- keep using fully-featured well optimized community solvers despite limitations in their API
- two new algotihms allowing for low-rank compression of  $S \rightarrow$  processing of larger coupled systems
  - 1. multi-solve based on the baseline coupling
  - 2. multi-factorization based on the advanced coupling

## **Multi-solve**

$$S_{i} = A_{ss_{i}} - A_{sv} \underbrace{(I_{vv} L_{vv}^{T})^{-1} A_{sv_{i}}^{T}}_{I_{vv}}$$

- 1 sparse facto. of the green matrix (symmetric)
- plenty of *sparse solve* involving the orange blocks (result is dense)



WITHOUT compression

## **Multi-solve**

$$S_{i} = A_{ss_{i}} - A_{sv} \underbrace{(L_{vv}L_{vv}^{T})^{-1}A_{sv_{i}}^{T}}_{I_{vv}}$$

- 1 sparse facto. of the green matrix (symmetric)
- plenty of *sparse solve* involving the orange blocks (result is dense)





WITH compression

## **Multi-factorization**

$$S_{ij} = A_{ss_{ij}} - \overbrace{A_{sv_i}(L_{vv}U_{vv})^{-1}A_{sv_j}^T}^{used with Schur API}$$

- multiple *sparse facto.+Schur* of the violet matrix (non-symmetric)
- computation of the Schur complement blocks via API



WITHOUT compression

## **Multi-factorization**

$$S_{ij} = A_{ss_{ij}} - \overbrace{A_{sv_i}(L_{vv}U_{vv})^{-1}A_{sv_j}^T}^{used with Schur API}$$

- multiple *sparse facto.+Schur* of the violet matrix (non-symmetric)
- computation of the Schur complement blocks via API



WITHOUT compression



WITH compression

## Experimental evaluation

### Environment

#### Academic test case

- short pipe (length: 2 m, radius: 4 m) [2]
- systems close enough to real-life models
- FEM/BEM mesh (v and s parts)

#### Configuration

- PlaFRIM [1], 24-core miriel nodes
  - 128 GiB of RAM
- precision parameter  $\epsilon$  set to  $10^{-3}$
- sparse solver with low-rank compression always on
- out-of-core disabled



A short pipe (20,000 unknowns)

#### Implementation

- sparse solver
  - MUMPS [3]
- dense solvers
  - SPIDO
    - (non-compressed S)
  - HMAT
    - (compressed S) [9]

## Solving larger systems



Best computation times of multi-solve and multi-factorization for both solver couplings MUMPS/HMAT and MUMPS/SPIDO. Parallel runs using 24 threads on single *miriel* node.

multi-factorization

### Performance-memory trade-off of multi-solve



multi-solve (compressed)



Comparison between the non-compressed and compressed multi-solve for the MUMPS/SPIDO and MUMPS/HMAT couplings on a coupled FEM/BEM system with 2,000,000 unknowns for varying  $n_c$  and  $n_c$ .

## Performance-memory trade-off of multi-factorization





Comparison between the non-compressed and compressed multi-factorization for the MUMPS/SPIDO and MUMPS/HMAT couplings on a coupled FEM/BEM system with 1,000,000 unknowns for varying  $n_b$ .

multi-factorization (compressed)

## Industrial application

#### Industrial test case

- 2,259,468 unknowns
  - 2,090,638 in the v part
  - 168,830 in the s part
- 32-core machine with 384 GiB of RAM



Model FEM/BEM mesh

#### Results

Algorithm	RAM	Time
	(GiB)	
vanilla advanced coupling	>384	N/A
multi-solve (non-compressed S)	224	15h
multi-factorization (non-compressed $S$ )	275	8h
multi-solve (compressed S)	35	9h
multi-factorization (compressed $S$ )	137	51m



Computed acoustic pressure

Conclusion

#### Contribution

- two algorithms allowing us to:
  - benefit from the most advanced functionalities of fully-featured solvers
  - process larger systems compared to vanilla couplings
    - 9M (multi-solve) and 2.5M (multi-factorization) vs. 1.3M on a single 24-core, 128 GiB RAM workstation
- confirm the advantage of compressing the Schur complement
- validate the algorithms on a real-life industrial case

### Summary

#### Limitations

- multi-solve: explicit storage of an intermediate result in a non-compressed dense matrix
- multi-factorization: superfluous re-factorizations of the sparse submatrix  $A_{\nu\nu}$

#### Ongoing work

- extension to out-of-core and distributed memory cases
- task-based solver coupling based on the StarPU runtime [4]

#### Acknowledgement

- Projet Région Nouvelle-Aquitaine 2018-1R50119 « HPC scalable ecosystem »
- PlaFRIM experimental testbed [1] (supported by Inria, CNRS LABRI and IMB, Université de Bordeaux, Bordeaux INP and Conseil Régional d'Aquitaine)

## Thank you for attending!

 PlaFRIM: Plateforme fédérative pour la recherche en informatique et mathématiques.

https://plafrim.fr/.

- [2] test\_FEMBEM, a simple application for testing dense and sparse solvers with pseudo-FEM or pseudo-BEM matrices. https://gitlab.inria.fr/solverstack/test\_fembem.
- [3] P. R. Amestoy, I. S. Duff, and J.-Y. L'Excellent, MUMPS multifrontal massively parallel solver version 2.0, (1998).
- [4] C. Augonnet, S. Thibault, and R. Namyst, StarPU: a Runtime System for Scheduling Tasks over Accelerator-Based Multicore Machines, Rapport de recherche RR-7240, INRIA, Mar. 2010.
- [5] P. K. Banerjee and R. Butterfield, Boundary element methods in engineering science, vol. 17, McGraw-Hill London, 1981.
- [6] F. Casenave, Méthodes de réduction de modèles appliquées à des problèmes d'aéroacoustique résolus par équations intégrales, PhD thesis, Université Paris-Est, 2013.
- [7] F. Casenave, A. Ern, and G. Sylvand, Coupled BEM-FEM for the convected Helmholtz equation with non-uniform flow in a bounded domain, Journal of Computational Physics, 257 (2014), pp. 627–644.
- [8] A. Ern and J.-L. Guermond, Theory and practice of finite elements, vol. 159, Springer Science & Business Media, 2013.
- B. Lizé, Résolution Directe Rapide pour les Éléments Finis de Frontière en Électromagnétisme et Acoustique : H-Matrices. Parallélisme et Applications Industrielles, PhD thesis, Université Paris 13, 2014.
- [10] P. Raviart and J. Thomas, A mixed finite element method for 2-nd order elliptic problems, in Mathematical Aspects of Finite Element Methods, I. Galligani and E. Magenes, eds., vol. 606 of Lecture Notes in Mathematics, Springer Berlin Heidelberg, 1977, pp. 292-315.
- [11] Sebaso, Jet engine airflow during take-off. https://commons.wikimedia.org/wiki/File: 20140308-Jet\_engine\_airflow\_during\_take-off.jpg.
- [12] A. Wang, N. Vlahopoulos, and K. Wu, Development of an energy boundary element formulation for computing high-frequency sound radiation from incoherent intensity boundary conditions, Journal of Sound and Vibration, 278 (2004), pp. 413–436.
- [13] F. Zhang, The Schur complement and its applications, vol. 4, Springer Science & Business Media, 2006.

## Appendix

### Numerical assessment



Relative error  $E_{rel}$  for the runs of multi-solve and multi-factorization having the best execution times and for both solver couplings MUMPS/HMAT and MUMPS/SPIDO. Parallel runs using 24 threads on single *miriel* node.

## Multi-solve

$$S_{i} = A_{ss_{i}} - \underbrace{A_{sv}}_{Z_{i}} \underbrace{(L_{vv}L_{vv}^{T})^{-1}A_{sv_{i}}^{T}}_{Z_{i}}$$

- 1 sparse facto. of the green matrix (symmetric)
- plenty of *sparse solve* involving the orange blocks (result is dense)



WITHOUT compression

## Multi-solve

$$S_{i} = A_{ss_{i}} - \underbrace{A_{sv} \underbrace{A_{sv} \left(L_{vv} L_{vv}^{\mathsf{T}}\right)^{-1} A_{sv_{i}}^{\mathsf{T}}}_{Z_{i}}}_{Z_{i}}$$

- 1 sparse facto. of the green matrix (symmetric)
- plenty of *sparse solve* involving the orange blocks (result is dense)





WITH compression

WITHOUT compression

## **Multi-factorization**

$$S_{ij} = A_{ss_{ij}} - \overbrace{A_{sv_i}(L_{vv}U_{vv})^{-1}A_{sv_j}^T}^{used with Schur API 
ightarrow X_{ij}}$$

- multiple *sparse facto.+Schur* of the violet matrix (non-symmetric)
- computation of the Schur complement blocks via API



## **Multi-factorization**

$$S_{ij} = A_{ss_{ij}} - \overbrace{A_{sv_i}(L_{vv}U_{vv})^{-1}A_{sv_j}^T}^{used with Schur API 
ightarrow X_{ij}}$$

- multiple *sparse facto.+Schur* of the violet matrix (non-symmetric)
- computation of the Schur complement blocks via API

