Iterative Trefftz method for three-dimensional electromagnetic waves propagation

Sébastien Pernet, Margot Sirdey, Sébastien Tordeux

▶ **To cite this version:**

Sébastien Pernet, Margot Sirdey, Sébastien Tordeux. Iterative Trefftz method for three-dimensional electromagnetic waves propagation. WAVES 2022 - 15th International Conference on Mathematical and Numerical Aspects of Wave Propagation, Jul 2022, Palaiseau (FR), France. hal-03945383

HAL Id: hal-03945383
https://hal.science/hal-03945383v1

Submitted on 19 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Iterative Trefftz method for three-dimensional electromagnetic waves simulation

Sébastien Pernet[1], Margot Sirdey[1,2], Sébastien Tordeux[2]

[1] MACI/LMA2S ONERA [2] EPC INRIA Makutu, LMAP UMR CNRS 5142, E2S-UPPA

WAVES 2022
Paris-Saclay, France, July 25, 2022

Adimensional Maxwell problem

Studied problem:

$$\begin{cases} \nabla \times \mathbf{H} &= ik\varepsilon_r \mathbf{E} \\ \nabla \times \mathbf{E} &= -ik\mu_r \mathbf{H}, \text{ on } \Omega, \end{cases}$$

where $\Omega \subset \mathbb{R}^3$ is a three-dimensional domain, ε_r and μ_r are the relative permittivity and permeability.

Objective: computation on very large domains

$\rightarrow$ **up to** $\mathbf{200}$ wavelengths $= \mathbf{200}\lambda = \mathbf{64} \times \mathbf{10^6}$ elements in the $\lambda/2$-mesh.

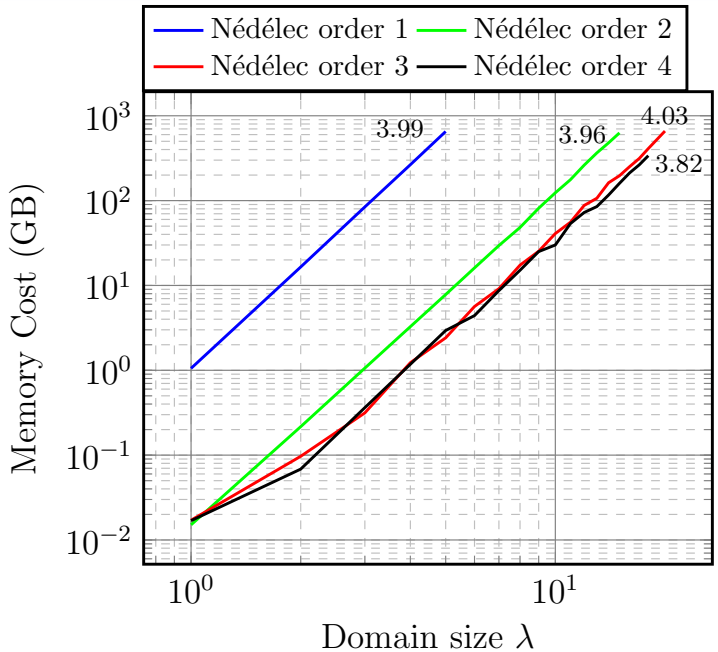
Frigate Surcouf at sea near Toulon. Date: 13 April 2004. Author: Franck Dubey. Source: netmarine.net.

Large domain size with respect to the wavelength $\Leftrightarrow$ $L >> \lambda$

$\Longrightarrow$ **numerical dispersion problem**

Memory limits of a classic solver[i]

- Use of second order Maxwell problem leads to a continuous formulation reducing the number of degrees of freedom
- Finite Element method handles spurious modes



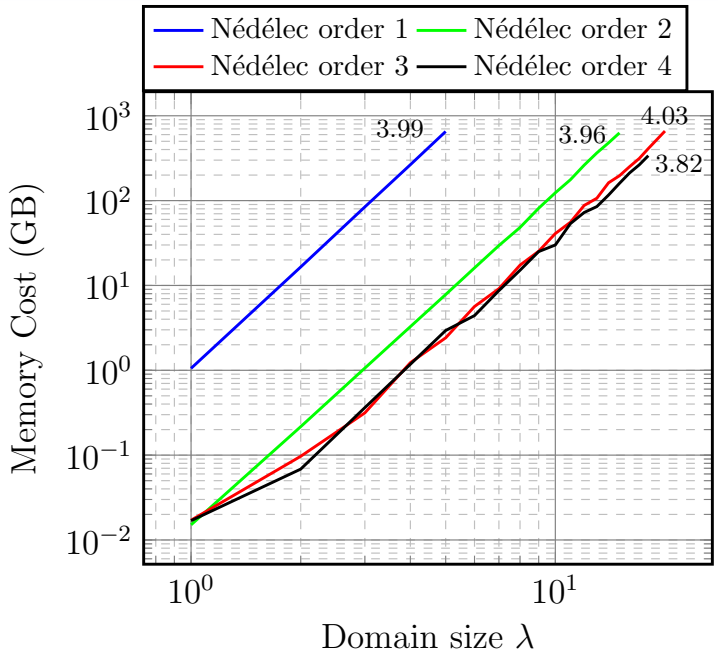
Memory used by the Nédélec solver in function of the size of the domain, memory $= \lambda^4$.

	Memory	Sockets	Domain size	Power	Cost	Hours.Sockets
Laptop	32 GB	1	10λ	200 W	0.034 €/h	0.25
// Computing	1 TB	16	24λ	3.2 KW	0.544 €/h	19.9
HPC	320 TB	5000	100λ	1 MW	170 €/h	> 25000
HPC	5120 TB	80000	200λ	16 MW	2720 €/h	> 800000

[i] F. Ihlenburg and I. Babuska, *Finite element solution of the Helmholtz equation with high wave number part II: the hp version of the FEM, SIAM Journal on Numerical Analysis*, 34 (1), pp. 315–358 (1997).

Memory limits of a classic solver[i]

Efficient optimisation cannot overcome the memory issue → iterative solvers ?



Memory used by the Nédélec solver in function of the size of the domain, memory $= \lambda^4$.

	Memory	Sockets	Domain size	Power	Cost	Hours.Sockets
Laptop	32 GB	1	10λ	200 W	0.034 €/h	0.25
// Computing	1 TB	16	24λ	3.2 KW	0.544 €/h	19.9
HPC	320 TB	5000	100λ	1 MW	170 €/h	> 25000
HPC	5120 TB	80000	200λ	16 MW	2720 €/h	> 800000

[i] F. Ihlenburg and I. Babuska, *Finite element solution of the Helmholtz equation with high wave number part II: the hp version of the FEM, SIAM Journal on Numerical Analysis*, 34 (1), pp. 315–358 (1997).

No convergence with iterative solvers

GMRES iterative solver example on

high order Nédélec Finite Element method (FE)

and

high order Discontinuous Galerkin method (DG)

from first order Maxwell problem and the impedance boundary condition

$$\gamma_t \mathbf{E} + Z_{\partial\Omega}\, \mathbf{n}_{\partial\Omega} \times \gamma_t \mathbf{H} = \mathbf{g} \quad \text{on } \partial\Omega \qquad \text{with} \quad Z_{\partial\Omega} = \frac{1 - R_{\partial\Omega}}{1 + R_{\partial\Omega}}.$$

No convergence with iterative solvers

GMRES iterative solver example on

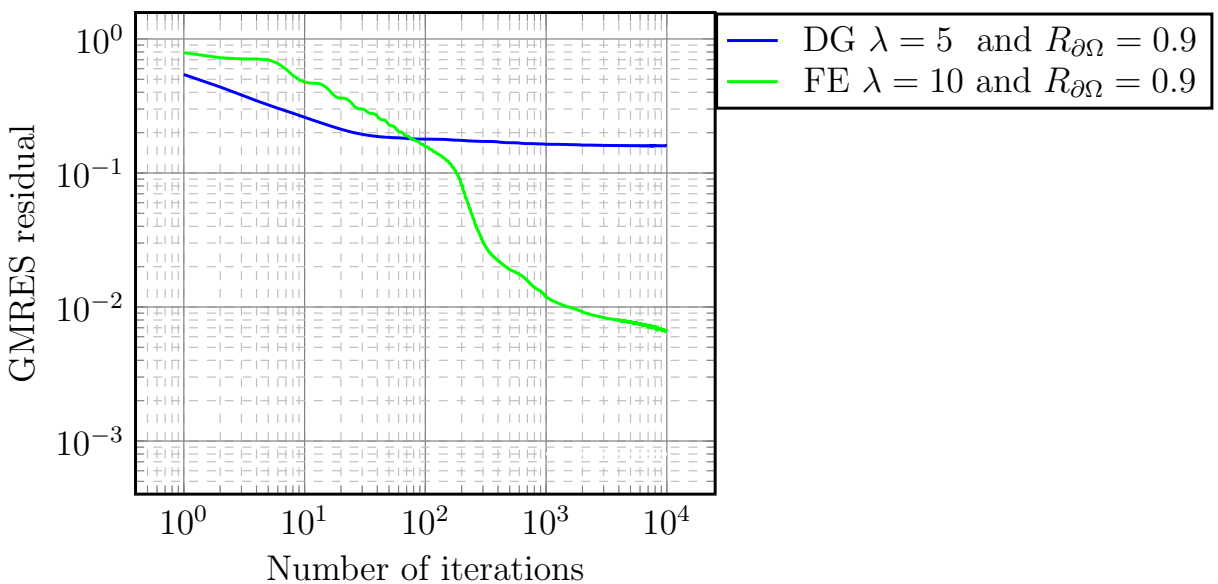
high order Nédélec Finite Element method (FE)

and

high order Discontinuous Galerkin method (DG)

from first order Maxwell problem and the impedance boundary condition

$$\gamma_t \mathbf{E} + Z_{\partial\Omega}\, \mathbf{n}_{\partial\Omega} \times \gamma_t \mathbf{H} = \mathbf{g} \quad \text{on } \partial\Omega \qquad \text{with} \quad Z_{\partial\Omega} = \frac{1 - R_{\partial\Omega}}{1 + R_{\partial\Omega}}.$$



Convergence for GMRES Nédélec FE method and for GMRES DG method.

→ Need to develop **low memory cost methods** based on **domain decomposition**.

Trefftz method construction steps

Trefftz method advantages[ii]

- Non dispersive basis functions
- Adapted to domain decomposition

[ii] R. Hiptmair, A. Moiola and I. Perugia, A survey of Trefftz methods for the Helmholtz equation, In Building bridges: connections and challenges in modern approaches to numerical partial differential equations, pp. 237–279. Springer (2016).

Trefftz method construction steps

Trefftz method advantages[ii]

- Non dispersive basis functions
- Adapted to domain decomposition

1. **Virtual work** principle on an element K

$$W_K = -ik_0 \int_K \varepsilon_r \, \mathbf{E}^K \cdot \overline{\mathbf{E}'^K} + \mu_r \, \mathbf{H}^{\mathbf{K}} \cdot \overline{\mathbf{H}'^{\mathbf{K}}} \mathrm{d}K.$$

2. **The solution and the basis functions satisfy Maxwell** on each element K

$$\begin{cases} W_K &= -\int_K \underbrace{\nabla \times \mathbf{H}^K}_{ik_0 \varepsilon_r \, \mathbf{E}^K} \cdot \overline{\mathbf{E}'^K} - \underbrace{\nabla \times \mathbf{E}^K}_{-ik_0 \mu_r \, \mathbf{H}^K} \cdot \overline{\mathbf{H}'^K} \mathrm{d}K, \\ W_K &= \int_K \mathbf{E}^K \cdot \underbrace{\overline{\nabla \times \mathbf{H}'^K}}_{ik_0 \varepsilon_r \, \mathbf{E}'^K} - \mathbf{H}^K \cdot \underbrace{\overline{\nabla \times \mathbf{E}'^K}}_{-ik_0 \mu_r \, \mathbf{H}'^K} \mathrm{d}K. \end{cases}$$

3. Thanks to the general Stokes formula, a variational formulation is imposed on boundaries: **reciprocity formula**,

$$\sum_K \int_{\partial K} \left(\mathbf{n}_K \times \gamma_t \mathbf{H}^K \right) \cdot \gamma_t \overline{\mathbf{E}'^K} + \gamma_t \mathbf{E}^K \cdot \left(\mathbf{n}_K \times \gamma_t \overline{\mathbf{H}'^K} \right) = 0.$$

[ii] R. Hiptmair, A. Moiola and I. Perugia, A survey of Trefftz methods for the Helmholtz equation, In Building bridges: connections and challenges in modern approaches to numerical partial differential equations, pp. 237–279. Springer (2016).

Trefftz variational formulation

4. Find $\mathbf{x} = (\mathbf{E}, \mathbf{H}) \in \mathbb{X}$, such that for all $\mathbf{x}' = (\mathbf{E}', \mathbf{H}') \in \mathbb{X}$,

$$\mathbf{a}(\mathbf{x}, \mathbf{x}') = \ell(\mathbf{x}'),$$

with $\mathbb{X}$ **a discontinuous space of local solutions of Maxwell system**, and

$$\mathbf{a}(\mathbf{x}, \mathbf{x}') = \sum_K \int_{\partial K} \widehat{\gamma_\times^K \mathbf{H}^K} \cdot \gamma_t \overline{\mathbf{E}'^K} + \widehat{\gamma_t \mathbf{E}^K} \cdot \gamma_\times^K \overline{\mathbf{H}'^K}, \quad \text{with } \gamma_\times \mathbf{H}^K = \mathbf{n}_K \times \gamma_t \mathbf{H}^K.$$

Trefftz variational formulation

4. Find $\mathbf{x} = (\mathbf{E}, \mathbf{H}) \in \mathbb{X}$, such that for all $\mathbf{x}' = (\mathbf{E}', \mathbf{H}') \in \mathbb{X}$,

$$\mathbf{a}(\mathbf{x}, \mathbf{x}') = \ell(\mathbf{x}'),$$

with $\mathbb{X}$ **a discontinuous space of local solutions of Maxwell system**, and

$$\mathbf{a}(\mathbf{x}, \mathbf{x}') = \sum_K \int_{\partial K} \widehat{\gamma_\times^K \mathbf{H}^K} \cdot \gamma_t \overline{\mathbf{E}'^K} + \widehat{\gamma_t \mathbf{E}^K} \cdot \gamma_\times^K \overline{\mathbf{H}'^K}, \quad \text{with } \gamma_\times \mathbf{H}^K = \mathbf{n}_K \times \gamma_t \mathbf{H}^K.$$

5. Clever choices of **numerical fluxes** $\widehat{\gamma_t \mathbf{E}^F}$ and $\widehat{\gamma_\times \mathbf{H}^F}$ ensure a **strict coercivity**

A numerical flux is defined on a face F
either an **interior face or a boundary face.**

Equivalency between

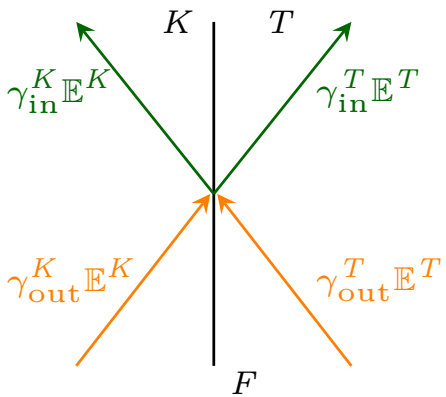
- Upwind schemes,
- UWVF methods,
- Riemann solvers.

Trefftz variational formulation

- **Upwind** schemes

Incoming and outgoing traces

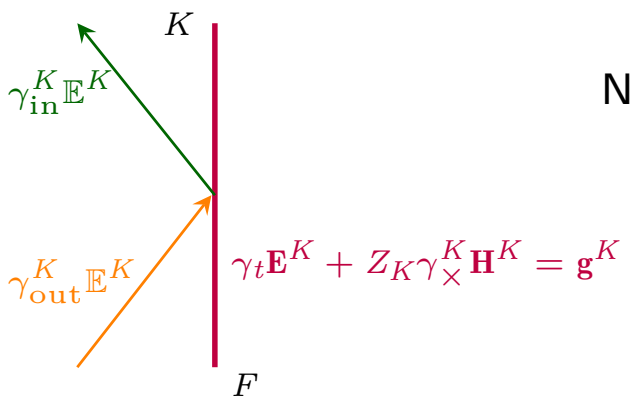
$$\gamma_{\text{in}}^K \mathbb{E}^K := \gamma_t \mathbf{E}^K + Z_K \gamma_\times^K \mathbf{H}^K \quad \text{and} \quad \gamma_{\text{out}}^K \mathbb{E}^K := \gamma_t \mathbf{E}^K - Z_K \gamma_\times^K \mathbf{H}^K.$$



Numerical fluxes on interior face

$$\widehat{\gamma_t \mathbf{E}^F} = \frac{Z_T}{Z_K + Z_T} \gamma_{\text{out}}^K \mathbb{E}^K + \frac{Z_K}{Z_K + Z_T} \gamma_{\text{out}}^T \mathbb{E}^T,$$

$$\widehat{\gamma_\times^K \mathbf{H}^F} = -\frac{1}{Z_K + Z_T} \gamma_{\text{out}}^K \mathbb{E}^K + \frac{1}{Z_K + Z_T} \gamma_{\text{out}}^K \mathbb{E}^T.$$



Numerical fluxes on boundary face

$$\widehat{\gamma_t \mathbf{E}^F} = \frac{Z_{\partial\Omega}}{Z_K + Z_{\partial\Omega}} \gamma_{\text{out}}^K \mathbb{E}^K + \frac{Z_K}{Z_K + Z_{\partial\Omega}} \mathbf{g}^K,$$

$$\widehat{\gamma_\times^K \mathbf{H}^F} = -\frac{1}{Z_{\partial\Omega} + Z_K} \gamma_{\text{out}}^K \mathbb{E}^K + \frac{1}{Z_{\partial\Omega} + Z_K} \mathbf{g}^K.$$

Trefftz variational formulation

- **Riemann** solvers[iii]
  Homogeneous numerical fluxes

$$\widehat{\gamma_t \mathbf{E}^F} = \{\{\gamma_t \mathbf{E}\}\}_F - \frac{[\![\gamma_\times \mathbf{H}]\!]_F}{2}, \quad \text{and} \quad \widehat{\gamma_\times \mathbf{H}^F} = \{\{\gamma_\times \mathbf{H}\}\}_F + \frac{[\![\gamma_t \mathbf{E}]\!]_F}{2}.$$

4. Find $\mathbf{x} = (\mathbf{E}, \mathbf{H}) \in \mathbb{X}$, such that for all $\mathbf{x}' = (\mathbf{E}', \mathbf{H}') \in \mathbb{X}$,

$$\mathbf{a}(\mathbf{x}, \mathbf{x}') = \sum_K \int_{\partial K} \widehat{\gamma_\times^K \mathbf{H}^K} \cdot \gamma_t \overline{\mathbf{E}'^K} + \widehat{\gamma_t \mathbf{E}^K} \cdot \gamma_\times^K \overline{\mathbf{H}'^K} = \ell(\mathbf{x}'),$$

   with $\mathbb{X}$ **a discontinuous space of local solutions of Maxwell system**.

5. Clever choices of **numerical fluxes** $\widehat{\gamma_t \mathbf{E}^F}$ and $\widehat{\gamma_\times \mathbf{H}^F}$ ensure a **strict coercivity**.
6. **Consistency**: accuracy of the best approximation is ensured by the choice of $\mathbb{X}$.
7. Trefftz method is **convergent** and provides **convergence theory for the iterative method** [iv].

$\Rightarrow$ **How to construct $\mathbb{X}$?**

[iii] S. Pernet, N. Serdiuk, M. Sirdey, et S. Tordeux, (2021). Discontinuous Galerkin Method based on Riemann fluxes for the time domain Maxwell System. Research report, INRIA Bordeaux-Sud-Ouest.

[iv] A. Buffa, P. Monk, Error estimates for the ultra weak variational formulation of the Helmholtz equation, Mathematical Modelling and Numerical Analysis 42 (6) (2008) 925940.

Trefftz variational formulation

- **Riemann** solvers[iii]
  Homogeneous numerical fluxes

$$\widehat{\gamma_t \mathbf{E}^F} = \{\{\gamma_t \mathbf{E}\}\}_F - \frac{[\![\gamma_\times \mathbf{H}]\!]_F}{2}, \quad \text{and} \quad \widehat{\gamma_\times \mathbf{H}^F} = \{\{\gamma_\times \mathbf{H}\}\}_F + \frac{[\![\gamma_t \mathbf{E}]\!]_F}{2}.$$

4. Find $\mathbf{x} = (\mathbf{E}, \mathbf{H}) \in \mathbb{X}$, such that for all $\mathbf{x}' = (\mathbf{E}', \mathbf{H}') \in \mathbb{X}$,

$$\mathbf{a}(\mathbf{x}, \mathbf{x}') = \sum_K \int_{\partial K} \widehat{\gamma_\times^K \mathbf{H}^K} \cdot \gamma_t \overline{\mathbf{E}'^K} + \widehat{\gamma_t \mathbf{E}^K} \cdot \gamma_\times^K \overline{\mathbf{H}'^K} = \ell(\mathbf{x}'),$$

   with $\mathbb{X}$ **a discontinuous space of local solutions of Maxwell system**.

5. Clever choices of **numerical fluxes** $\widehat{\gamma_t \mathbf{E}^F}$ and $\widehat{\gamma_\times \mathbf{H}^F}$ ensure a **strict coercivity**.
6. **Consistency**: accuracy of the best approximation is ensured by the choice of $\mathbb{X}$.

7. Trefftz method is **convergent** and provides **convergence theory for the iterative method** [iv].

$$\Rightarrow \textbf{How to construct } \mathbb{X} \textbf{ ?}$$

[iii] S. Pernet, N. Serdiuk, M. Sirdey, et S. Tordeux, (2021). Discontinuous Galerkin Method based on Riemann fluxes for the time domain Maxwell System. Research report, INRIA Bordeaux-Sud-Ouest.

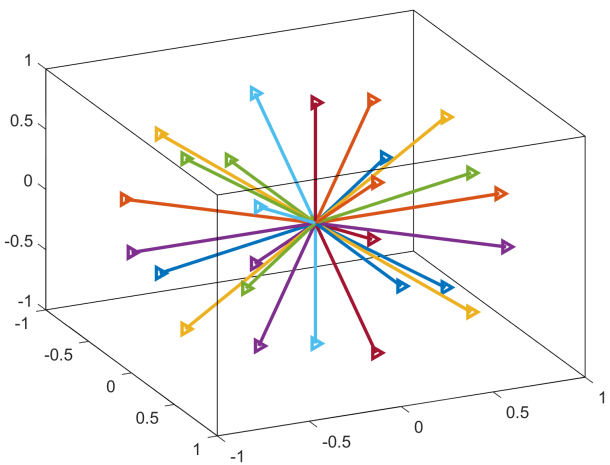
[iv] A. Buffa, P. Monk, Error estimates for the ultra weak variational formulation of the Helmholtz equation, Mathematical Modelling and Numerical Analysis 42 (6) (2008) 925940.

Basis functions space

An electromagnetic plane wave $\mathbf{E} = \overrightarrow{p}\, e^{ik\,\overrightarrow{d}\cdot\overrightarrow{x}}$ is defined by :

- a direction of propagation $\overrightarrow{d} = (d_x, d_y, d_z)$,
- a polarization $\overrightarrow{p} = (p_x, p_y, p_z)$ orthogonal to $\overrightarrow{d}$.

3D discrete space



A plane waves basis in a cube, composed of normalized basis functions $\mathbf{w}^{\mathrm{j}}$, for $\mathrm{j} = 1, N$.

- On an element K, the numerical wave $\mathbf{x}^K := (\mathbf{E}^K, \mathbf{H}^K)$ is

$$\mathbf{E}^K = \sum_{\mathrm{j}=1}^{N} u_{\mathrm{j}} \underbrace{\overrightarrow{p_{\mathrm{j}}}\, e^{ik\overrightarrow{d_{\mathrm{j}}}\cdot\overrightarrow{x}}}_{\mathbf{w}^{\mathrm{j}}},$$

$$\mathbf{H}^K = \sum_{\mathrm{j}=1}^{N} u_{\mathrm{j}} \left(\overrightarrow{p_{\mathrm{j}}} \times \overrightarrow{d_{\mathrm{j}}}\right) e^{ik\overrightarrow{d_{\mathrm{j}}}\cdot\overrightarrow{x}}.$$

- The number of plane waves N is the number of degrees of freedom for each element K.
- **What is the optimal number of plane waves ?**

Non exhaustive references about plane waves basis and alternatives

What is the optimal number of plane waves ?

- Moiola, A. (2011). Trefftz-discontinuous Galerkin methods for time-harmonic wave problems (Doctoral dissertation, ETH Zurich). (Page 156).
- Moiola, A., Hiptmair, R., Perugia, I. (2011). Vekua theory for the Helmholtz operator. Zeitschrift für angewandte Mathematik und Physik, 62(5), 779-807.

In this presentation: Basis reduction strategy → Conditioning improvement

- T. Luostari, T. Huttunen, P. Monk, (2013). Improvements for the ultra weak variational formulation, Int. J. Numer. Meth. Engng 94, 598624.
- S. Congreve, J. Gedicke, I. Perugia, (2019). Numerical investigation of the conditioning for plane wave discontinuous Galerkin methods, Vol. 126 of Lecture Notes in Computational Science and Engineering, Springer.
- H. Barucq, A. Bendali, J. Diaz, S. Tordeux, (2021). Local strategies for improving the conditioning of the plane-wave Ultra-Weak Variational Formulation. Journal of Computational Physics, 441, 110449.

Alternatives could be used for the basis functions space

- Evanescent modes
  - E. Parolin, D. Huybrechs, A. MOIOLA, (2022). Stable approximation of Helmholtz solutions by evanescent plane waves. arXiv preprint arXiv:2202.05658.
- Quasi-Trefftz methods
  - L. M. Imbert-Gérard, A. Moiola, P. Stocker, (2020). A space-time quasi-Trefftz DG method for the wave equation with piecewise-smooth coefficients. arXiv preprint arXiv:2011.04617.
  - H. S. Fure, S. Pernet, M. Sirdey, S. Tordeux, (2020). A discontinuous Galerkin Trefftz type method for solving the two dimensional Maxwell equations. SN Partial Differ. Equ. Appl. 1, 23.

General matrix shape

$$\mathbf{A}[\mathbf{x}] = \mathbf{F},$$

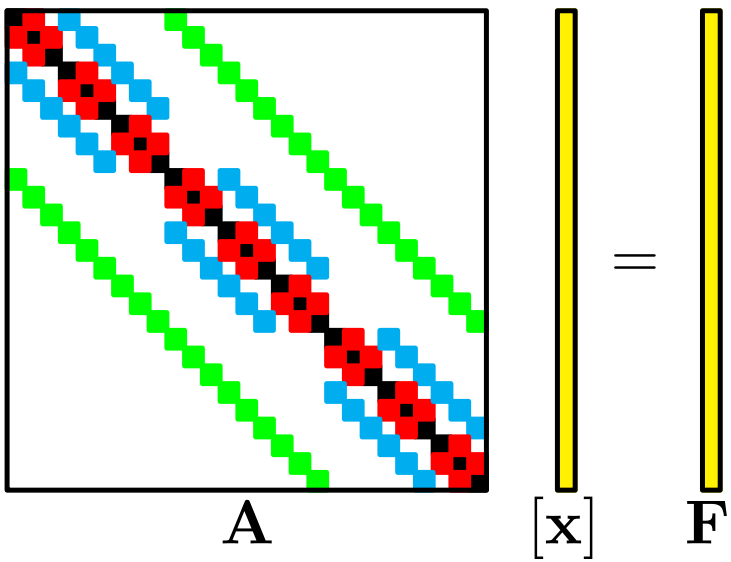
where

$$\mathbf{A}_{i,j} := \mathbf{a}(\mathbf{w}^j, \mathbf{w}^i)$$

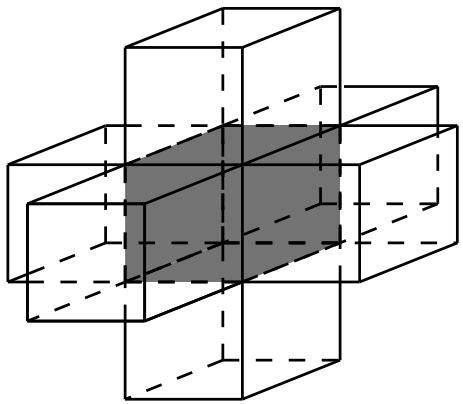
$$\mathbf{F}_i := \mathbf{l}(\mathbf{w}^i) \quad \text{for} \quad i, j = 1, N_{\text{dof}}.$$

$$N_{\text{dof}} := N \times N_{\text{elem}}$$

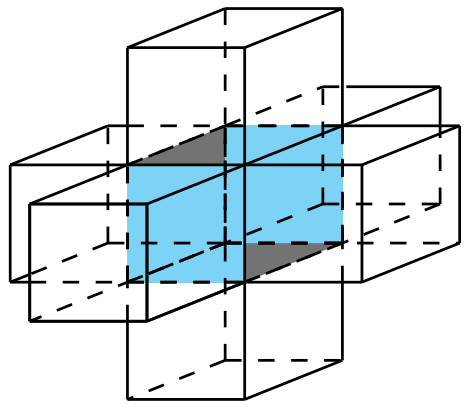
$$\text{size(A)} = N^2 \times 7 \times N_{\text{elem}}$$



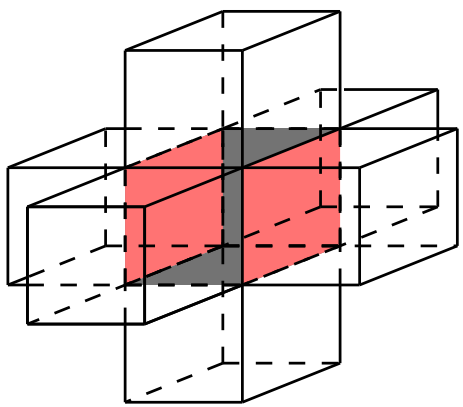
Element assembling of $\mathbf{A}$ for $N_{\text{elem}} = 27 = 3^3$ cubes.



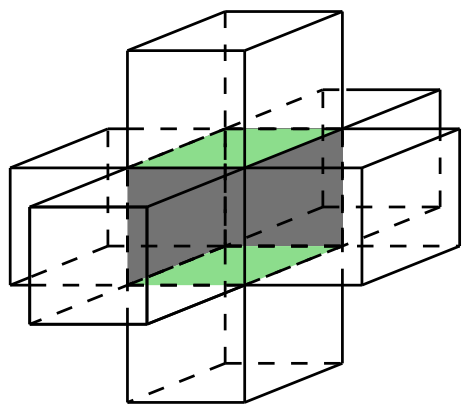
Self-interaction



Front and back interactions

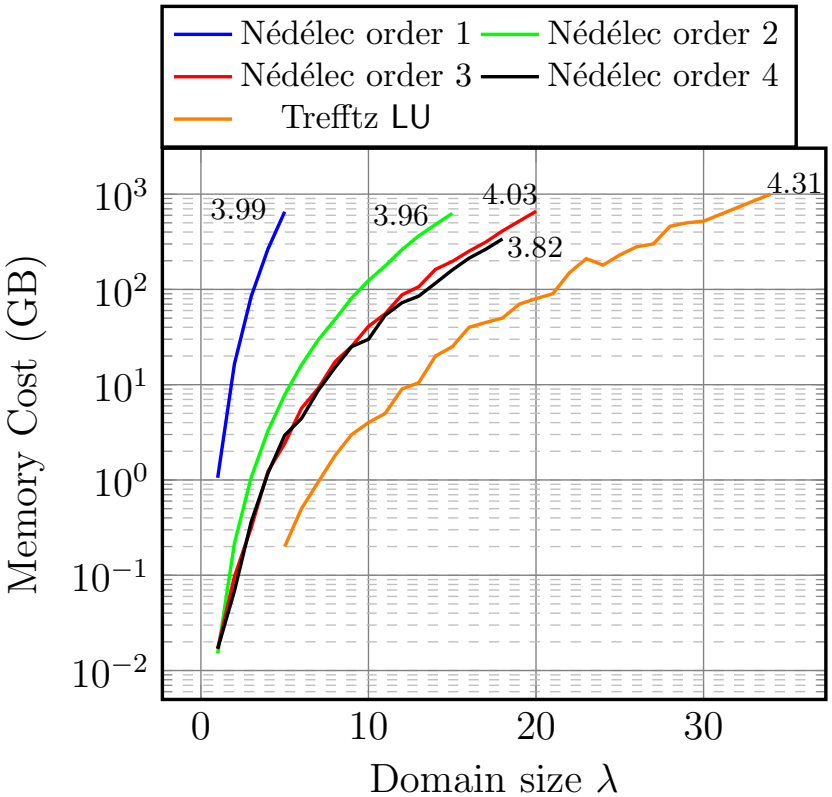


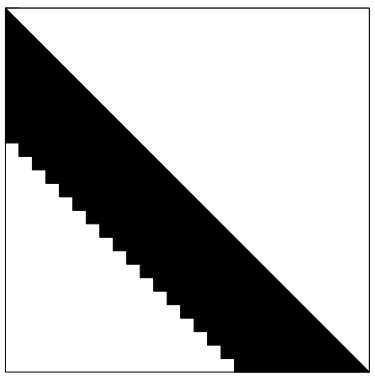
Left and right interactions



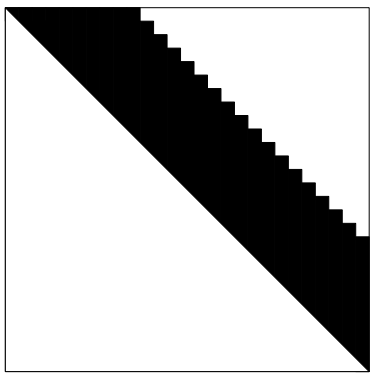
Bottom and top interactions

Memory limit of the direct solver





L



U

Domain size	5λ	10λ	20λ	100λ	200λ
N_{elem}	1000	8000	64000	8×10^6	64×10^6
LU memory	4.3 GB	69 GB	1.1 TB	690 TB	11040 TB

Table: LU factorization memory in function of the domain size.

GMRES method alternative

Coercivity property $\mathrm{Re}(\mathbf{a}(\mathbf{x},\mathbf{x}')) > 0$ $\rightarrow$ suitable framework for GMRES.

$$[\mathbf{x}]^{N_{\mathrm{kry}}} = \mathrm{GMRES}(\mathbf{A},\mathbf{F},N_{\mathrm{kry}})$$

General Minimal RESidual (GMRES) approximates the solution $\mathbf{x}^{N_{\mathrm{kry}}} \in \mathbb{K}^{N_{\mathrm{kry}}}$ of

$$\text{Find } \mathbf{x}^{N_{\mathrm{kry}}} \in \mathbb{K}^{N_{\mathrm{kry}}},\ \mathbf{a}(\mathbf{x}^{N_{\mathrm{kry}}},\mathbf{x}') = \mathbf{l}(\mathbf{x}'),\ \forall \mathbf{x}' \in \mathbb{K}^{N_{\mathrm{kry}}}$$

$$\iff \quad \text{Find } [\mathbf{x}^{N_{\mathrm{kry}}}] \in [\mathbb{K}^{N_{\mathrm{kry}}}],\quad [\mathbf{x}']^{*}\mathbf{A}[\mathbf{x}^{N_{\mathrm{kry}}}] = [\mathbf{x}']^{*}\mathbf{F},\quad \forall [\mathbf{x}'] \in [\mathbb{K}^{N_{\mathrm{kry}}}],$$

GMRES method alternative

Coercivity property $\mathrm{Re}(\mathbf{a}(\mathbf{x},\mathbf{x}')) > 0$ $\rightarrow$ suitable framework for GMRES.

$$[\mathbf{x}]^{N_{\mathrm{kry}}} = \mathrm{GMRES}(\mathbf{A},\mathbf{F},N_{\mathrm{kry}})$$

General Minimal RESidual (GMRES) approximates the solution $\mathbf{x}^{N_{\mathrm{kry}}} \in \mathbb{K}^{N_{\mathrm{kry}}}$ of

$$\text{Find } \mathbf{x}^{N_{\mathrm{kry}}} \in \mathbb{K}^{N_{\mathrm{kry}}},\ \mathbf{a}(\mathbf{x}^{N_{\mathrm{kry}}},\mathbf{x}') = \mathbf{l}(\mathbf{x}'),\ \forall \mathbf{x}' \in \mathbb{K}^{N_{\mathrm{kry}}}$$

$$\iff \quad \text{Find } [\mathbf{x}^{N_{\mathrm{kry}}}] \in [\mathbb{K}^{N_{\mathrm{kry}}}],\quad [\mathbf{x}']^{*}\mathbf{A}[\mathbf{x}^{N_{\mathrm{kry}}}] = [\mathbf{x}']^{*}\mathbf{F},\quad \forall [\mathbf{x}'] \in [\mathbb{K}^{N_{\mathrm{kry}}}],$$

where the **Krylov space** of dimension N_{kry} is

$$\left[\mathbb{K}^{N_{\mathrm{kry}}}\right] := \mathrm{span}\underbrace{\left[\begin{pmatrix}\vdots\\ \mathbf{F}\\ \vdots\end{pmatrix}, \begin{pmatrix}\vdots\\ \mathbf{AF}\\ \vdots\end{pmatrix}, \ldots, \begin{pmatrix}\vdots\\ \mathbf{A}^{N_{\mathrm{kry}}-1}\mathbf{F}\\ \vdots\end{pmatrix}\right]}_{N_{\mathrm{kry}}} \Bigg\} N_{\mathrm{dof}},$$

$\Longrightarrow$ **size(Krylov)** $= N_{\mathrm{kry}} \times N_{\mathrm{dof}}$, where N_{dof} is the dimension of $\mathbf{F}$.

Trefftz GMRES method

- **GMRES convergence depends on the spectrum of the matrix $\mathbf{A}$.**

  Classic theorem for the residual norm of a GMRES solver is given by[v]

$$\|[\mathbf{x}]^{m+1} - [\mathbf{x}]^{m}\| = \|r^{m+1}\| \leq \|\|\mathbf{P}\|\|\|\mathbf{P}^{-1}\|\| \, \|r^{0}\| \, \varepsilon^{m} \, ,$$

  with

$$\varepsilon^{m} = \min_{\substack{p \in \mathcal{P}^m \\ p(0)=1}} \max_{\lambda_i \in \rho(\mathbf{A})} |p(\lambda_i)| \qquad \text{and} \qquad \mathbf{P} \text{ the eigenvectors matrix.}$$

  More accurate studies do exist[vi]
  But classic theorem is enough when $\mathbf{A}$ has good properties

- **The consideration of huge cases depends on the memory cost of the method**

$$\text{size(Krylov)} + \text{size(A)} = N_{\text{kry}} \times N \times N_{\text{elem}} + N^2 \times 7 \times N_{\text{elem}}$$

We propose strategies to

1. **Remove the spectrum of $\mathbf{A}$ from 0.**
2. **Reduce the memory cost of the Trefftz GMRES method.**

[v] Saad, Youcef and Schultz, Martin H,GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems,SIAM,7,3,856–869,1986

[vi] J. Liesen, P. Tichý, (2012). The field of values bound on ideal GMRES. Preprint arXiv:1211.5969.

Trefftz GMRES method

- **GMRES convergence depends on the spectrum of the matrix $\mathbf{A}$.**

  Classic theorem for the residual norm of a GMRES solver is given by[v]

  $$\|[\mathbf{x}]^{m+1} - [\mathbf{x}]^{m}\| = \|r^{m+1}\| \leq \|\|\mathbf{P}\|\|\|\mathbf{P}^{-1}\|\| \, \|r^{0}\| \, \varepsilon^{m} \, ,$$

  with

  $$\varepsilon^{m} = \min_{\substack{p \in \mathcal{P}^m \\ p(0)=1}} \max_{\lambda_i \in \rho(\mathbf{A})} |p(\lambda_i)| \qquad \text{and} \qquad \mathbf{P} \text{ the eigenvectors matrix.}$$

  More accurate studies do exist[vi]
  But classic theorem is enough when $\mathbf{A}$ has good properties

- **The consideration of huge cases depends on the memory cost of the method**

  $$\text{size(Krylov)} + \text{size(A)} = N_{\text{kry}} \times N \times N_{\text{elem}} + N^2 \times 7 \times N_{\text{elem}}$$

We propose strategies to

1. **Remove the spectrum of $\mathbf{A}$ from 0.**
2. **Reduce the memory cost of the Trefftz GMRES method.**

[v] Saad, Youcef and Schultz, Martin H,GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems,SIAM,7,3,856–869,1986

[vi] J. Liesen, P. Tichý, (2012). The field of values bound on ideal GMRES. Preprint arXiv:1211.5969.

GMRES method with a restart strategy

GMRES converges in theory in N_{dof} iterations, *ie* $N_{\text{kry}} = N_{\text{dof}}$.

- N_{elem} the number of elements
- $N = 52$ the number of plane waves
- $N_{\text{dof}} = N \times N_{\text{elem}}$ the number of degrees of freedom

$$\rightarrow \textbf{size(Krylov)} = N_{\text{kry}} \times N_{\text{dof}} = 10^6 \times 52 \times 10^6 = \mathbf{43264}\ \textbf{TB}.$$

Problems when N_{dof} becomes large, *ie* when N_{elem} becomes large

- Huge memory cost to store $\mathbb{K}^{N_{\text{kry}}}$
- Rounding errors accumulated at each step

GMRES method with a restart strategy

GMRES converges in theory in N_{dof} iterations, *ie* $N_{\text{kry}} = N_{\text{dof}}$.

- N_{elem} the number of elements
- $N = 52$ the number of plane waves
- $N_{\text{dof}} = N \times N_{\text{elem}}$ the number of degrees of freedom

$$\rightarrow \textbf{size(Krylov)} = N_{\text{kry}} \times N_{\text{dof}} = 10^6 \times 52 \times 10^6 = \mathbf{43264\ TB}.$$

Problems when N_{dof} becomes large, *ie* when N_{elem} becomes large

- Huge memory cost to store $\mathbb{K}^{N_{\text{kry}}}$
- Rounding errors accumulated at each step

Idea: **truncate Krylov space** $\mathbb{K}^{N_{\text{kry}}}$ into smaller spaces of size $N_{\text{kry}}^{restart} < N_{\text{kry}}$

$$[\mathbf{x}]^{\text{k}} = [\mathbf{x}]^{\text{k}-1} + GMRES(\mathbf{A}, \mathbf{F} - \mathbf{A}[\mathbf{x}]^{\text{k}-1}), \quad [\mathbf{x}]^0 = 0.$$

Example with $N_{\text{kry}}^{restart} = 50$:

$$\textbf{size(Krylov)} = N_{\text{kry}}^{restart} \times N_{\text{dof}} = 50 \times 52 \times 10^6 = \mathbf{41.6\ GB}$$

$$\rightarrow \textbf{Strategy with restart reduces the memory cost}$$

Memory cost with matrix storage

Memory cost to run GMRES method **with matrix storage** is

$$\text{size(Krylov)} + \text{size(A)} = 50 \times N \times N_{\text{elem}} + N^2 \times 7 \times N_{\text{elem}}$$

$$\frac{\text{size(A)}}{\text{size(Krylov)}} \simeq 7$$

With matrix storage: the blocking structure is N_{elem}

Memory cost with matrix storage

Memory cost to run GMRES method **with matrix storage** is

$$\text{size(Krylov)} + \text{size(A)} = 50 \times N \times N_{\text{elem}} + N^2 \times 7 \times N_{\text{elem}}$$

$$\frac{\text{size(A)}}{\text{size(Krylov)}} \simeq 7$$

With matrix storage: the blocking structure is N_{elem}

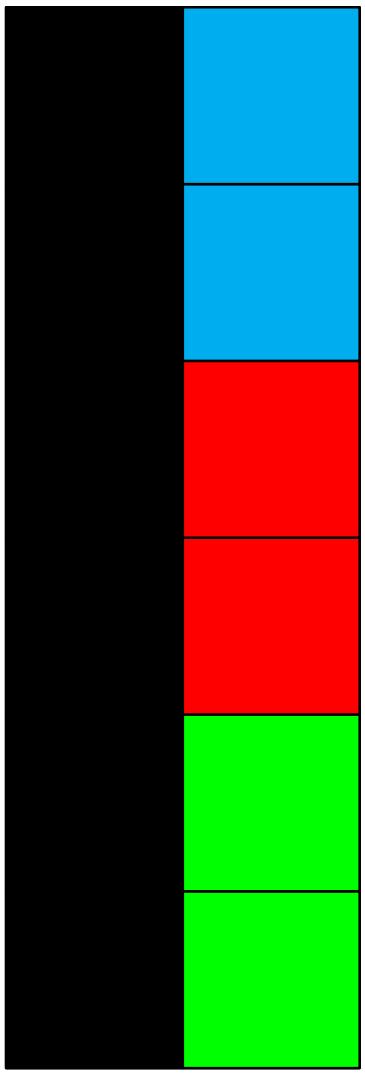
- $N_{\text{kry}}^{restart} = 50$ the dimension of the Krylov space
- N_{elem} the number of elements
- $N = 52$ the number of plane waves
- $N_{\text{dof}} = N \times N_{\text{elem}}$ the number of degrees of freedom

Domain size	5λ	10λ	20λ	100λ	200λ
N_{elem}	1000	8000	64000	8×10^6	64×10^6
With matrix storage	344 MB	2.75 GB	22.04 GB	2.75 TB	2204 TB

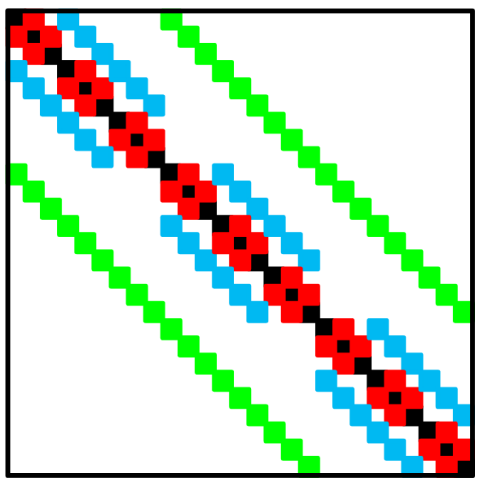
Table: Memory cost for the storage of the matrix $\mathbf{A}$ in function of the domain size.

Solution: run GMRES method without matrix storage

Memory cost for matrix storage $7 \times N^2$



Cartesian structure



$\mathbf{A}$

$\mathbf{A}$ is seeing as actions.

Memory cost without matrix storage

Memory cost to run GMRES method ~~**with matrix storage**~~ without matrix storage is

$$\text{size(Krylov)} + \sout{\text{size(A)}} = 50 \times N \times N_{\text{elem}} + \sout{N^2 \times 7 \times N_{\text{elem}}}$$

$$\text{size(Krylov)} + \text{size(interactions)} = 50 \times N \times N_{\text{elem}} + N^2 \times 7$$

Without matrix storage: only interactions between elements are stored

Memory cost without matrix storage

Memory cost to run GMRES method ~~**with matrix storage**~~ without matrix storage is

$$\text{size(Krylov)} + \cancel{\text{size(A)}} = 50 \times N \times N_{\text{elem}} + \cancel{N^2 \times 7 \times N_{\text{elem}}}$$

$$\text{size(Krylov)} + \text{size(interactions)} = 50 \times N \times N_{\text{elem}} + N^2 \times 7$$

Without matrix storage: only interactions between elements are stored

- $N_{\text{kry}}^{restart} = 50$ the dimension of the Krylov space
- N_{elem} the number of elements
- $N = 52$ the number of plane waves
- $N_{\text{dof}} = N \times N_{\text{elem}}$ the number of degrees of freedom

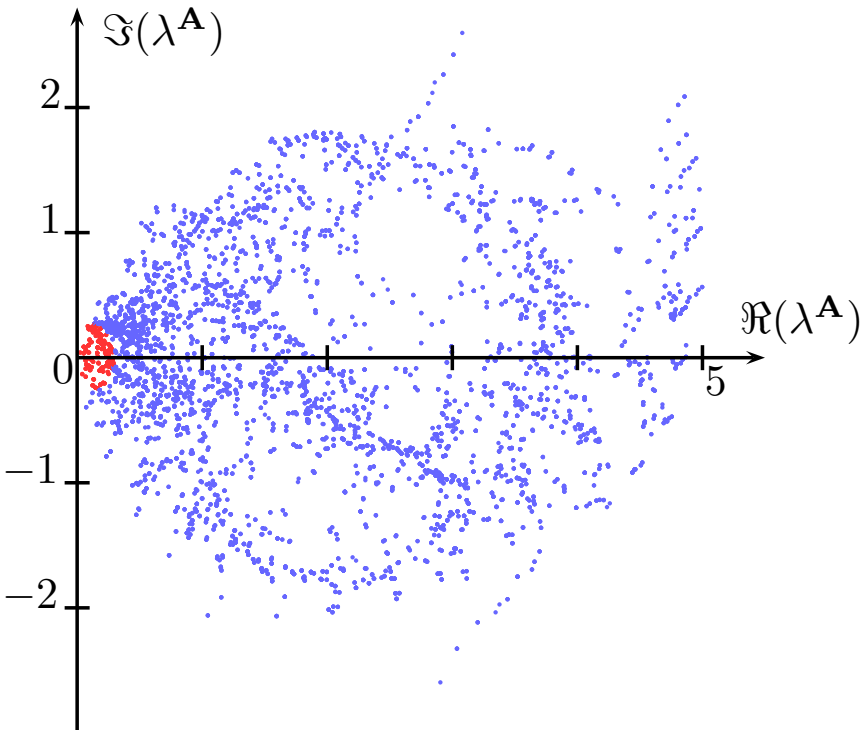
Domain size	5λ	10λ	20λ	100λ	200λ
N_{elem}	1000	8000	64000	8×10^6	64×10^6
With matrix storage	344 MB	2.75 GB	22.04 GB	2.75 TB	2204 TB
Without matrix storage	41.9 MB	0.332 GB	2.66 GB	332 GB	2.66 TB
$\frac{\text{size(interactions)}}{\text{size(Krylov)}}$	7.3E^{-3}	9.1E^{-4}	1.1E^{-4}	9.1E^{-7}	1.1E^{-7}

Table: Comparison of a GMRES method with or without matrix storage.

Cessenat-Després preconditioner

The GMRES method convergence depends on
the **spectrum of $\mathbf{A}$** and on the **eigenvalues close to** 0.

Idea: use a **preconditioned GMRES** method solving $\mathbf{A}^{\#}\mathbf{A}[\mathbf{x}] = \mathbf{A}^{\#}\mathbf{F}$.

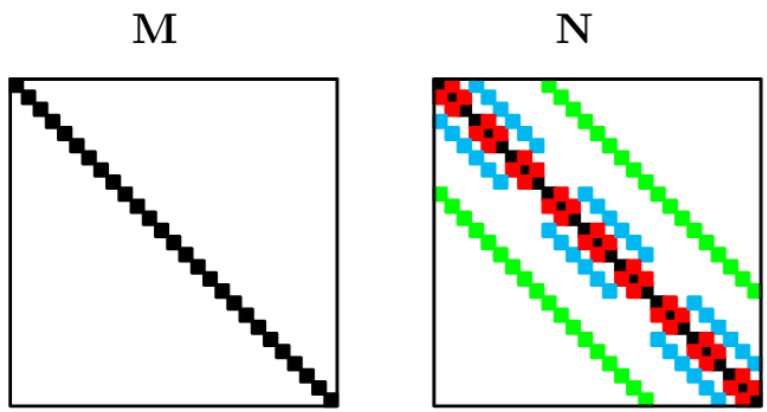


Real and complex parts of the spectrum $\lambda^{\mathbf{A}}$ of matrix $\mathbf{A}$, for a 6λ domain size.

$$\mathbf{A} = \mathbf{M} + \mathbf{N}$$

$$\mathbf{M}^{\mathrm{i,j}}_{\ell,k} = (\mathbf{w}^{\mathrm{i}}_{k}, \mathbf{w}^{\mathrm{i}}_{\ell})_{L^2_t(\partial\mathcal{T})}\, \delta_{\mathrm{i,j}}$$

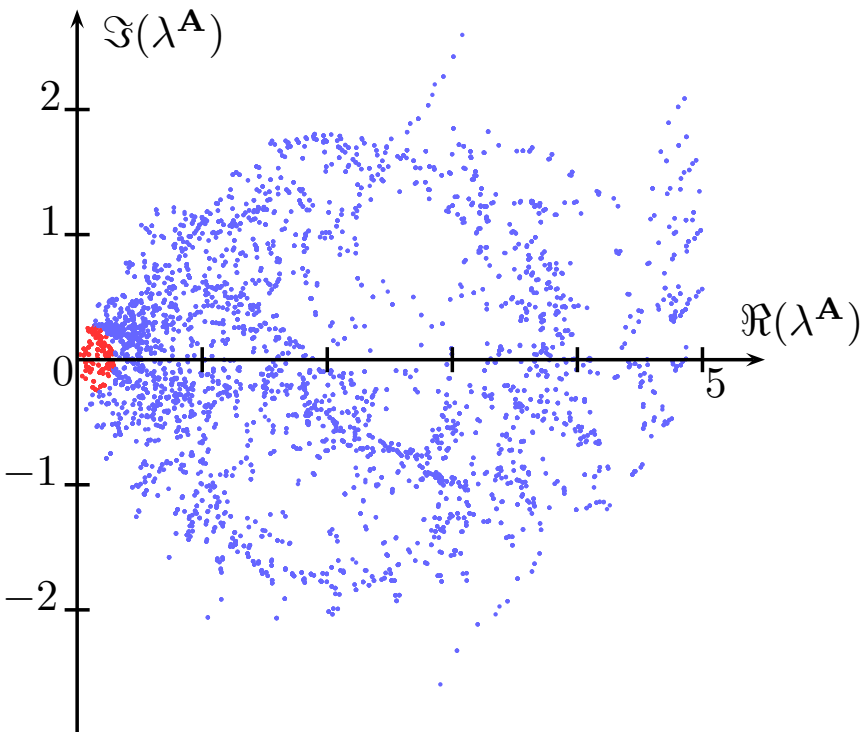
Cessenat-Després decomposition [a]



[a] O. Cessenat and B. Despres, (1998). Application of an ultra weak variational formulation of elliptic PDEs to the two-dimensional Helmholtz problem. SIAM journal on numerical analysis, 35(1), 255-299.

Cessenat-Després preconditioner

The GMRES method convergence depends on
the **spectrum of $\mathbf{A}$** and on the **eigenvalues close to** 0.



Real and complex parts of the spectrum $\lambda^{\mathbf{A}}$ of matrix $\mathbf{A}$,
for a 6λ domain size.

Idea: use a **preconditioned GMRES** method solving $\mathbf{A}^{\#}\mathbf{A}[\mathbf{x}] = \mathbf{A}^{\#}\mathbf{F}$.

$$\begin{aligned} \mathbf{A}^{\#}(\mathbf{M}+\mathbf{N})[\mathbf{x}] &= \mathbf{A}^{\#}\mathbf{F} \\ \mathbf{A}^{\#}\mathbf{M}[\mathbf{x}]^{n+1} &= -\mathbf{A}^{\#}\mathbf{N}[\mathbf{x}]^{n} + \mathbf{A}^{\#}\mathbf{F} \end{aligned}$$

$\rightarrow$ a **preconditioner** can be $\mathbf{A}^{\#} = \mathbf{M}^{-1}$

$\mathbf{M}$ is a scalar product

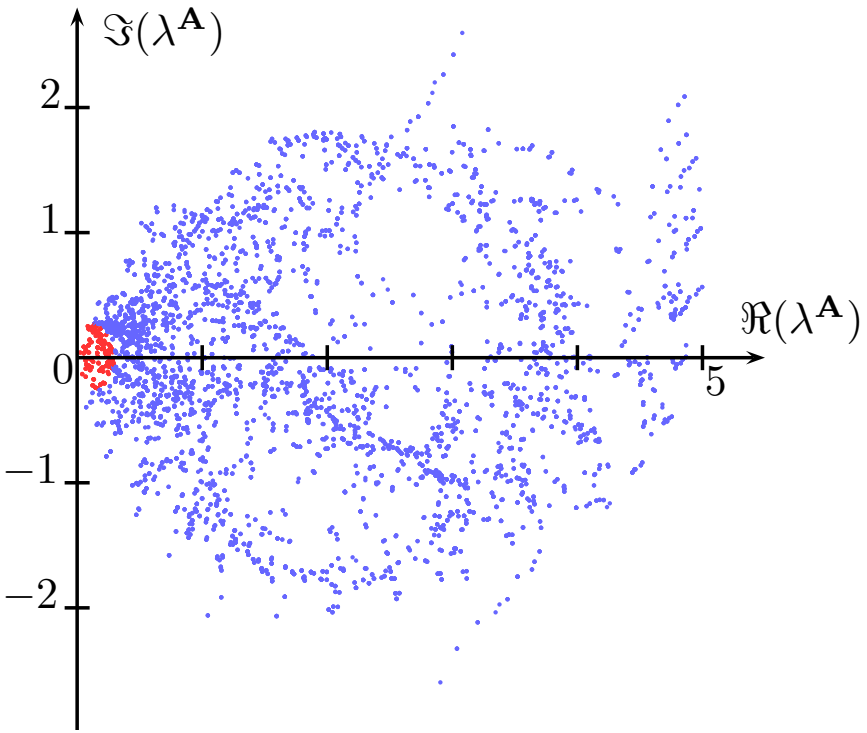
$\rightarrow$ symetric preconditioner

$$\widetilde{\mathbf{A}} = \mathbf{M}^{-\frac{1}{2}}\mathbf{A}\mathbf{M}^{-\frac{1}{2}} = \mathbf{Id} + \mathbf{M}^{-\frac{1}{2}}\mathbf{N}\mathbf{M}^{-\frac{1}{2}}$$

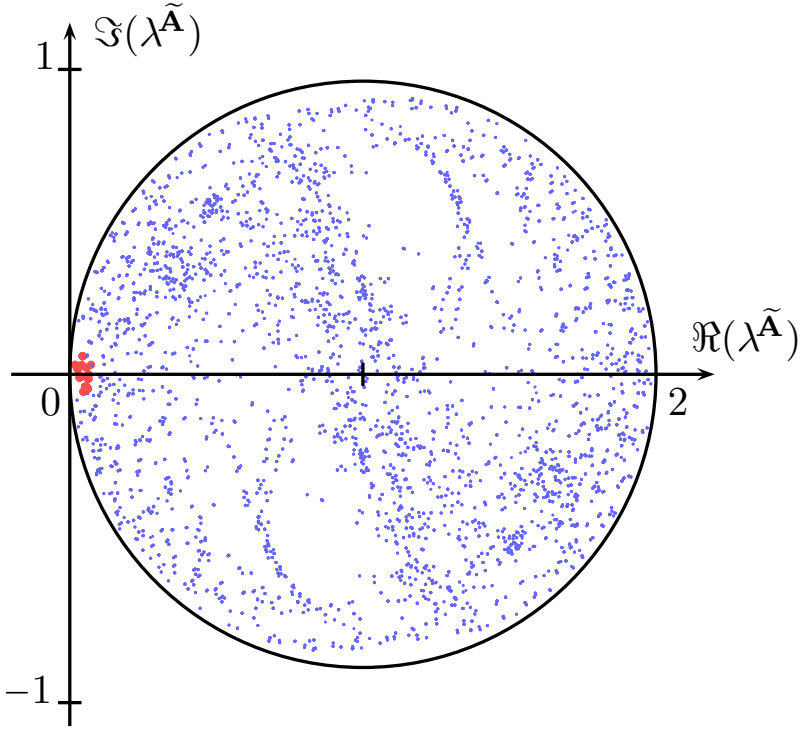
Cessenat-Després preconditioner

The GMRES method convergence depends on
the **spectrum of $\mathbf{A}$** and on the **eigenvalues close to** 0.

$$\tilde{\mathbf{A}} = \mathbf{Id} + \mathbf{M}^{-\frac{1}{2}}\mathbf{N}\mathbf{M}^{-\frac{1}{2}}$$
is contractant



Real and complex parts of the spectrum $\lambda^{\mathbf{A}}$ of matrix $\mathbf{A}$,
for a 6λ domain size.



Real and complex parts of the spectrum $\lambda^{\widetilde{\mathbf{A}}}$ of matrix $\widetilde{\mathbf{A}}$,
for a 6λ domain size.

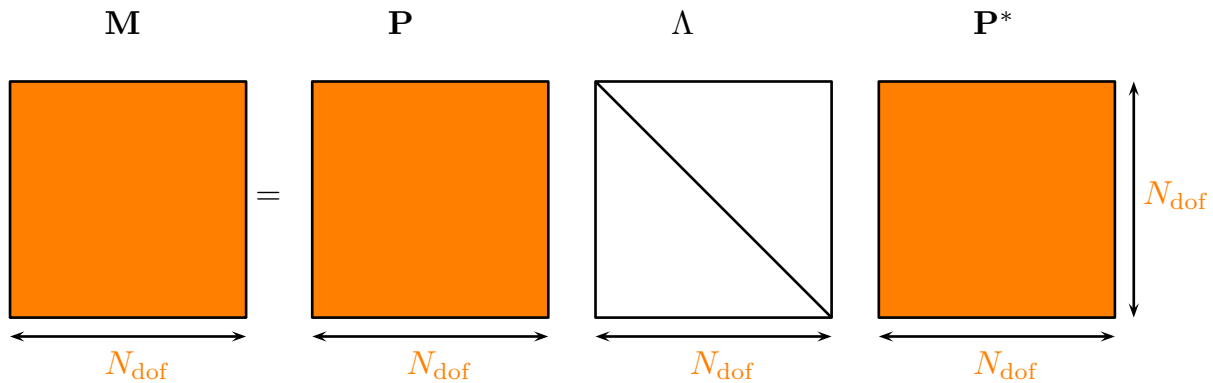
Basis reduction strategy

$\mathbf{M}$ is **symmetric positive definite**

$$\mathbf{M}^{\mathrm{i,j}}_{\ell,k} = (\mathbf{w}^{\mathrm{i}}_k, \mathbf{w}^{\mathrm{i}}_\ell)_{L^2_t(\partial\mathcal{T})}\, \delta_{\mathrm{i,j}}$$

for $\mathrm{i,j} = 1, N_{\mathrm{elem}}$

and $\ell, k = 1, N.$



$\mathbf{P}$ is the orthogonal eigenvector matrix.
Λ is the diagonal eigenvalue matrix satisfying

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N_{\mathrm{dof}}} \quad \text{with } \lambda_{\mathrm{i}} = \Lambda_{\mathrm{i,i}}.$$

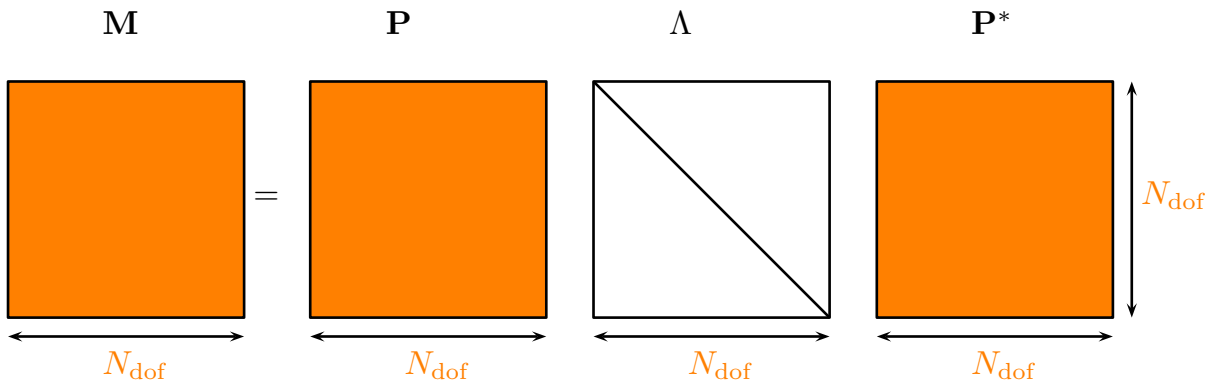
Basis reduction strategy

$\mathbf{M}$ is **symmetric positive definite**

$$\mathbf{M}^{\mathrm{i,j}}_{\ell,k} = (\mathbf{w}^{\mathrm{i}}_{k}, \mathbf{w}^{\mathrm{i}}_{\ell})_{L^2_t(\partial\mathcal{T})}\, \delta_{\mathrm{i,j}}$$

for $\mathrm{i,j} = 1, N_{\mathrm{elem}}$

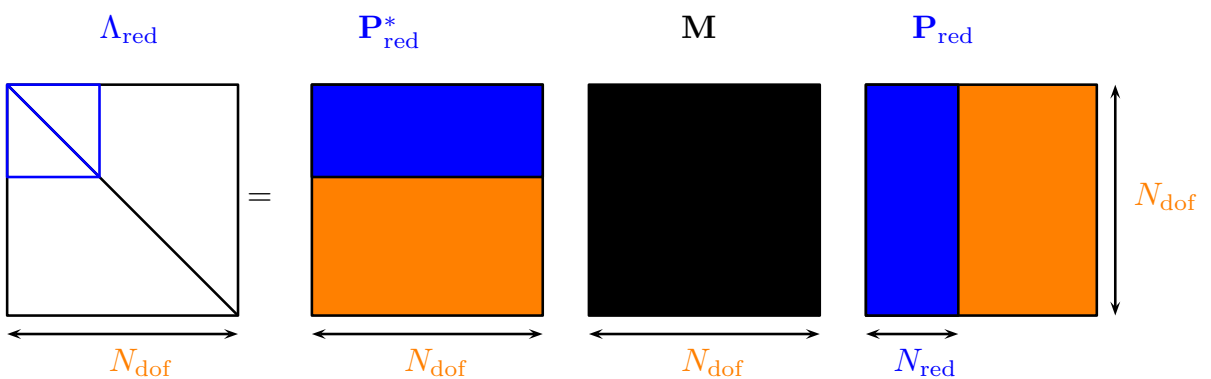
and $\quad \ell, k = 1, N.$



$\mathbf{P}$ is the orthogonal eigenvector matrix.
Λ is the diagonal eigenvalue matrix satisfying

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{N_{\mathrm{dof}}} \quad \text{with } \lambda_{\mathrm{i}} = \Lambda_{\mathrm{i,i}}.$$

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{N_{\mathrm{red}}} \\ \vdots \\ \lambda_{N_{\mathrm{dof}}-1} \\ \lambda_{N_{\mathrm{dof}}} \end{pmatrix} \rightarrow \textbf{Filtering} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{N_{\mathrm{red}}} \end{pmatrix} \qquad \frac{\lambda_{\mathrm{i}}}{\lambda_1} < \varepsilon$$



Normalized reduced basis

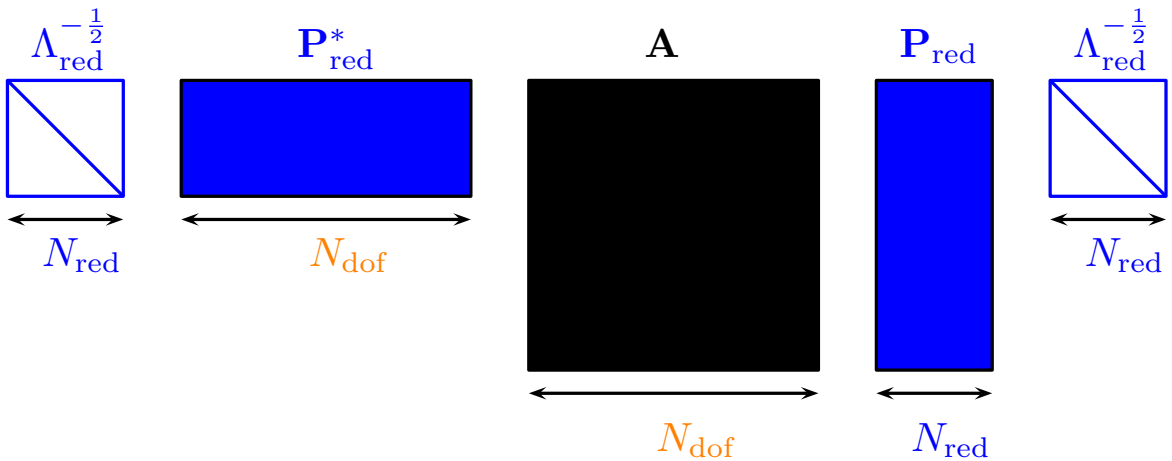
The unknown has a new representation

$$\mathbf{x} = \sum_{i=1}^{N} [\mathbf{x}]_i \mathbf{w}^i = \sum_{i=1}^{N} [\mathbf{y}]_i \widetilde{\mathbf{w}}^i \simeq \sum_{i=1}^{N_{red}} [\mathbf{y}]_i \widetilde{\mathbf{w}}^i, \quad \mathbf{Y}_{red} = [\mathbf{y}]_{i=1,N_{red}}.$$

This approximation **does not lose information** since $\frac{\lambda_i}{\lambda_1} < \varepsilon$, with ε well chosen.

The reduced basis can be normalized and the new matrix is

$$\mathbf{A}_{red} = \mathbf{I}_{red} + \mathbf{N}_{red} =$$



Normalized reduced basis

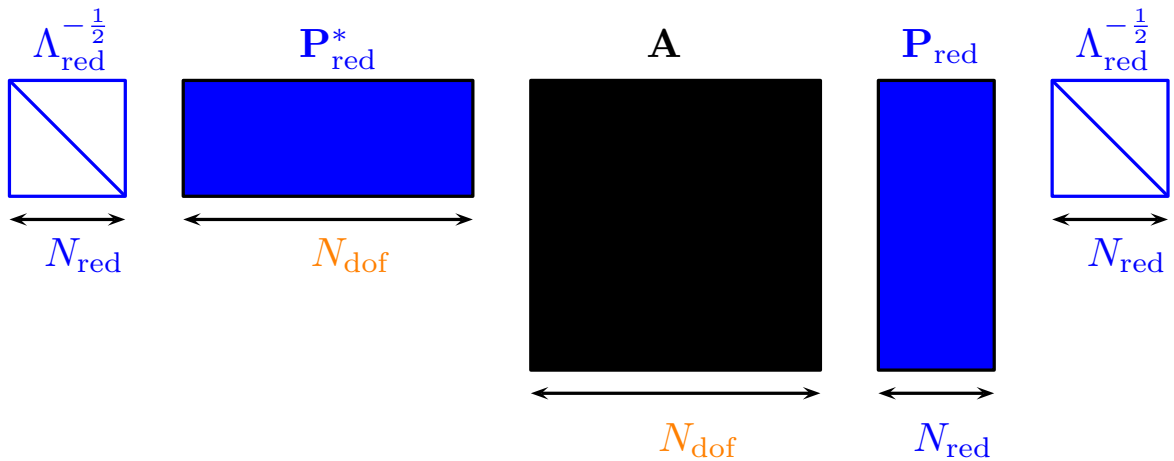
The unknown has a new representation

$$\mathbf{x} = \sum_{i=1}^{N} [\mathbf{x}]_i \mathbf{w}^i = \sum_{i=1}^{N} [\mathbf{y}]_i \widetilde{\mathbf{w}}^i \simeq \sum_{i=1}^{N_{red}} [\mathbf{y}]_i \widetilde{\mathbf{w}}^i, \quad \mathbf{Y}_{red} = [\mathbf{y}]_{i=1,N_{red}}.$$

This approximation **does not lose information** since $\frac{\lambda_i}{\lambda_1} < \varepsilon$, with ε well chosen.

The reduced basis can be normalized and the new matrix is

$$\mathbf{A}_{red} = \mathbf{I}_{red} + \mathbf{N}_{red} =$$

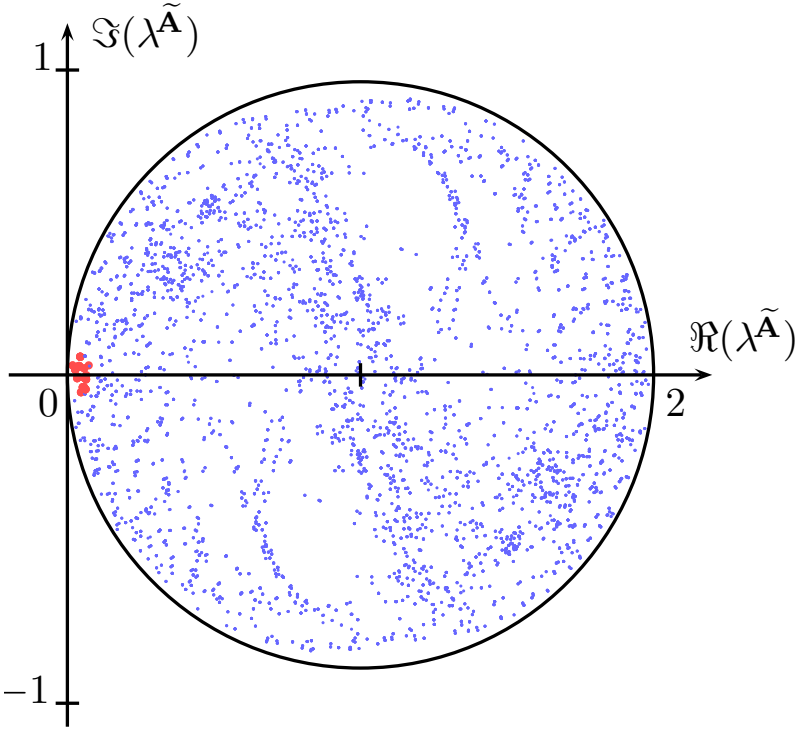


Computation of $\mathbf{M}^{-1}$ is trivial $\rightarrow$ **memory gain**

$$[\mathbf{x}^{n+1}] = -\mathbf{N}_{red}[\mathbf{x}^n] + \mathbf{F}_{red}, \quad \text{with} \quad \mathbf{N}_{red} := \Lambda_{red}^{-\frac{1}{2}} \mathbf{P}_{red}^{*} \mathbf{N} \mathbf{P}_{red} \Lambda_{red}^{-\frac{1}{2}} \quad \text{and} \quad [\mathbf{x}^0] = 0.$$

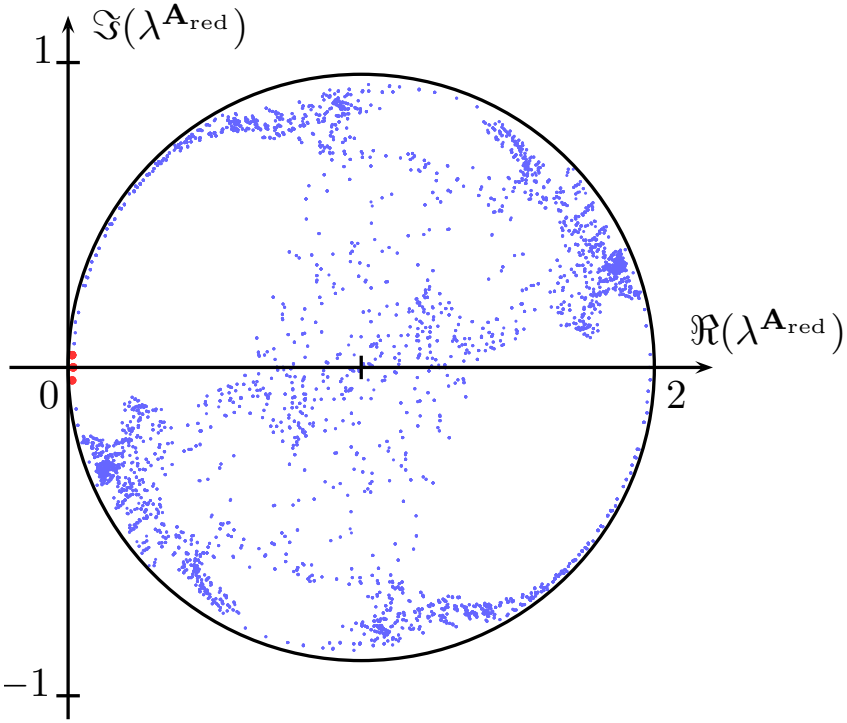
Basis reduction improves matrix conditioning

Unreduced preconditioned matrix $\tilde{\mathbf{A}}$



Real and complex parts of the spectrum $\lambda^{\widetilde{\mathbf{A}}}$ of matrix $\widetilde{\mathbf{A}}$, for a 6λ domain size.

Reduced preconditioned matrix $\mathbf{A}_{\mathrm{red}}$



Real and complex parts of the spectrum $\lambda^{\mathbf{A}_{\mathrm{red}}}$ of $\mathbf{A}_{\mathrm{red}}$, for a 6λ domain size and $\varepsilon = 10^{-5}$.

Basis reduction improves memory burden

$\mathbf{N}_{\text{red}}$ depends on the size of an element K and of the threshold ε.

Size of K \ ε	10^{-13}	10^{-11}	10^{-9}	10^{-7}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
0.25λ	154	126	96	70	48	48	30	24
0.5λ	190	186	174	132	96	84	70	48
1λ	196	196	196	190	180	174	148	114

Table: Values of N_{red} when reducing the basis of size $N = 196$ according to ε and the size of an element K in wavelength.

0.25λ case, $N_{\text{kry}}^{\text{restart}} = 500$.

10^{-15}	ε	10^{-13}	10^{-11}	10^{-9}	10^{-7}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
$\mathbf{N = 196}$	N_{red}	154	126	96	70	48	36	30	16
1.57 GB	Memory cost (GB)	1.23	1.01	0.76	0.56	0.38	0.28	0.24	0.13

Table: Memory cost in GigaBytes (GB) for the storage of the non-reduced matrix $\mathbf{A}$ and of the reduced matrix $\mathbf{A}_{\text{red}}$ according to N_{red}, for a 10λ domain size.

More reduction $\rightarrow$ less noise in the solution $\rightarrow$ improvement of the conditioning
Too much reduction $\rightarrow$ numerical solution deteriorates

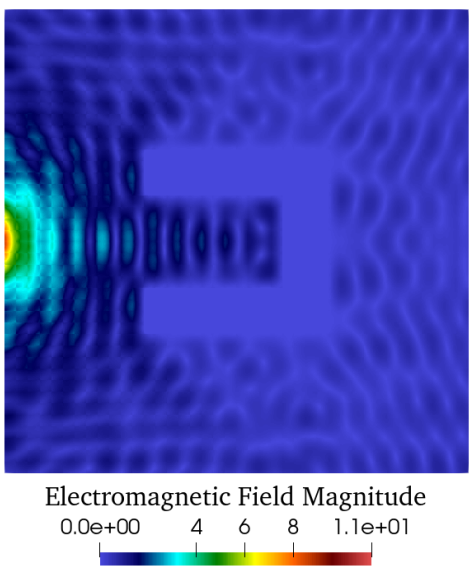
How far can we reduce the plane wave basis ?

ε	10^{-7}	10^{-5}	10^{-4}	10^{-3}	10^{-2}
Trefftz error	N/A	3.4×10^{-2}	9.8×10^{-2}	0.15	0.74
Memory cost (GB)	0.56	0.38	0.28	0.24	0.13
Time (hours)	55	27	11.3	7.23	0.58

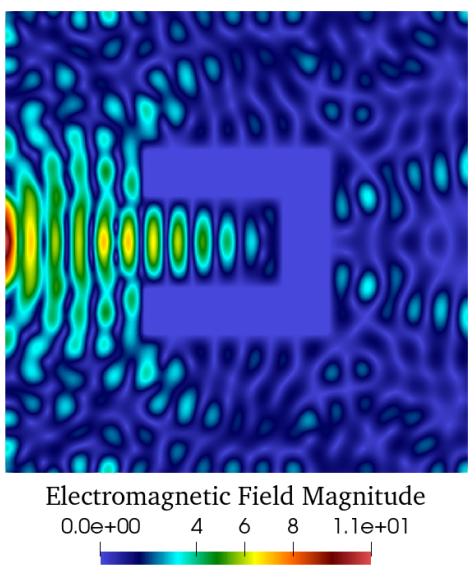
Table: Trefftz error $\frac{\sqrt{\|[\mathbf{x}_\varepsilon - \mathbf{x}_{\text{ref}}]^* \mathbf{M} [\mathbf{x}_\varepsilon - \mathbf{x}_{\text{ref}}]\|_z}}{\sqrt{\|[\mathbf{x}_{\text{ref}}]^* \mathbf{M} [\mathbf{x}_{\text{ref}}]\|_z}}$, computation time and memory cost for a 10^{-8} GMRES residual for $[\mathbf{x}_\varepsilon]$ according to each ε.

GMRES residual $= 10^{-8}$ but poor approximation when ε is not well chosen.

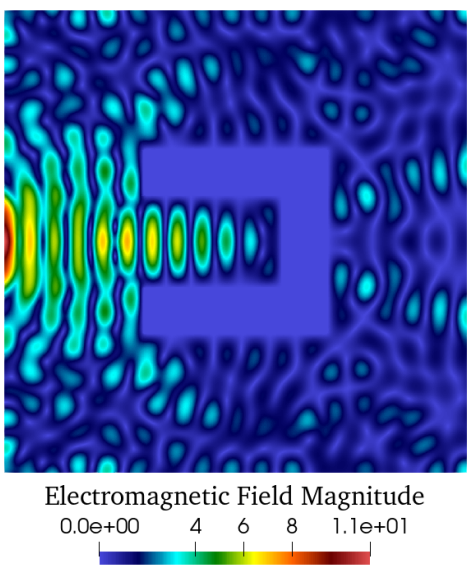
Too much reduction
$N_{\text{red}} = 16$ $\varepsilon = 10^{-2}$



A "optimal" trade-off
$N_{\text{red}} = 36$ $\varepsilon = 10^{-4}$



Solution of reference
$N_{\text{red}} = 70$ $\varepsilon = 10^{-7}$



Global left preconditioner

$\mathbf{A}[\mathbf{x}] = \left(\mathbf{M}^{x/y/z} + \mathbf{N}^{x/y/z}\right)[\mathbf{x}] = \mathbf{F}$, lead to **iterative formulas**

$$\mathbf{M}^{x/y/z}[\mathbf{x}^{n+1}] = \mathbf{F} - \mathbf{N}^{x/y/z}[\mathbf{x}^{n}].$$

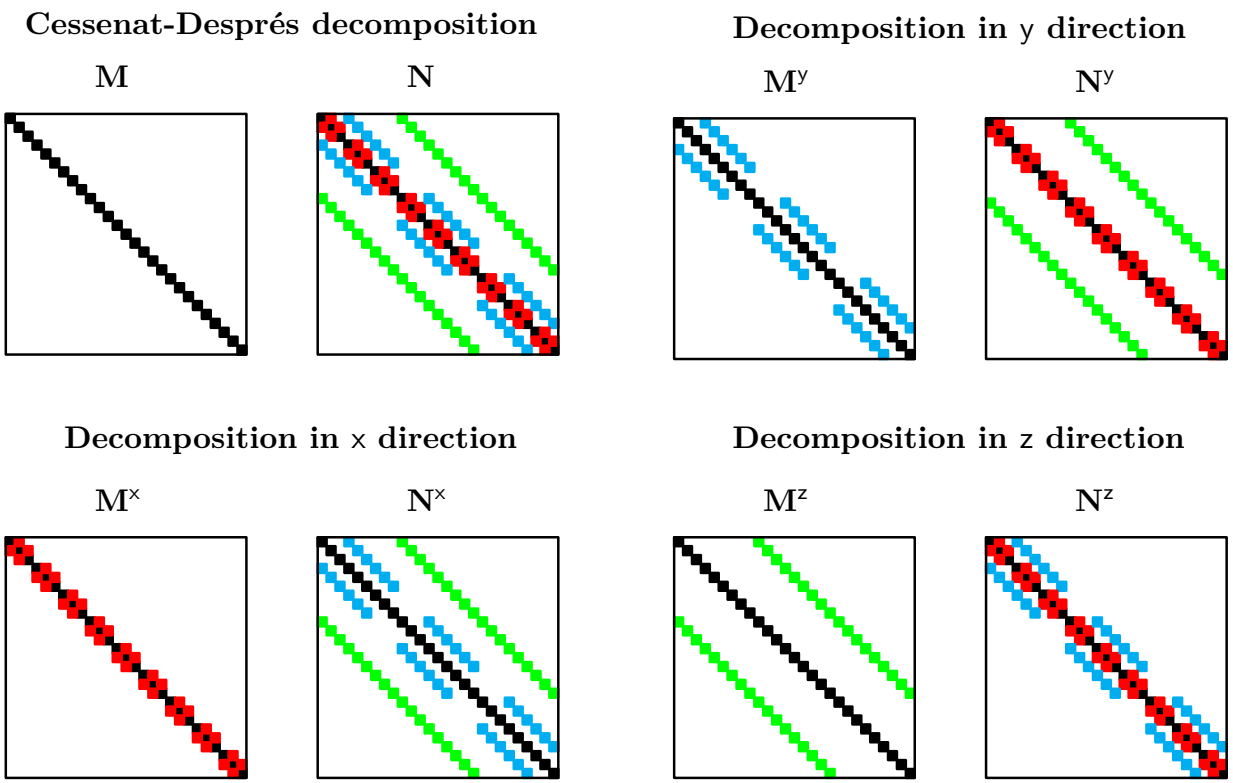


Figure: Matrices structures for the Cessenat-Després decomposition and the decompositions in x, y or z direction.

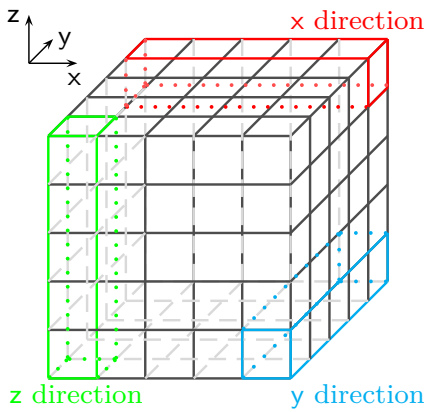


Figure: Global one-dimensional subdomains in a cube.

Better conditioning with cyclic decomposition

By applying successively each of these 1D preconditioners, we get a global preconditioner associating $[\mathbf{x}^n]$ to $[\mathbf{x}^{n+1}]$ thanks to the iterative scheme

$$\begin{cases} \mathbf{M}^{\mathsf{x}}[\mathbf{x}_0^{n+1}] = \mathbf{F} - \mathbf{N}^{\mathsf{x}}[\mathbf{x}^n], \\ \mathbf{M}^{\mathsf{y}}[\mathbf{x}_1^{n+1}] = \mathbf{F} - \mathbf{N}^{\mathsf{y}}[\mathbf{x}_0^{n+1}], \\ \mathbf{M}^{\mathsf{z}}[\mathbf{x}^{n+1}] = \mathbf{F} - \mathbf{N}^{\mathsf{z}}[\mathbf{x}_1^{n+1}]. \end{cases}$$

Better conditioning with cyclic decomposition

Preconditioned reduced GMRES problem

$$\mathbf{P}_{\mathrm{red}}^{\mathrm{xyz}}\mathbf{A}_{\mathrm{red}}[\mathbf{x}]_{\mathrm{red}} = \mathbf{P}_{\mathrm{red}}^{\mathrm{xyz}}\mathbf{F}_{\mathrm{red}},$$

where $\mathbf{P}_{\mathrm{red}}^{\mathrm{xyz}}$ is the reduced (and consequently symetrized) version of

$$\mathbf{P}^{\mathrm{xyz}} := \left(\mathbf{M}^{\mathrm{z}}\right)^{-1}\mathbf{N}^{\mathrm{z}}\left(\mathbf{M}^{\mathrm{y}}\right)^{-1}\mathbf{N}^{\mathrm{y}}\left(\mathbf{M}^{\mathrm{x}}\right)^{-1}.$$

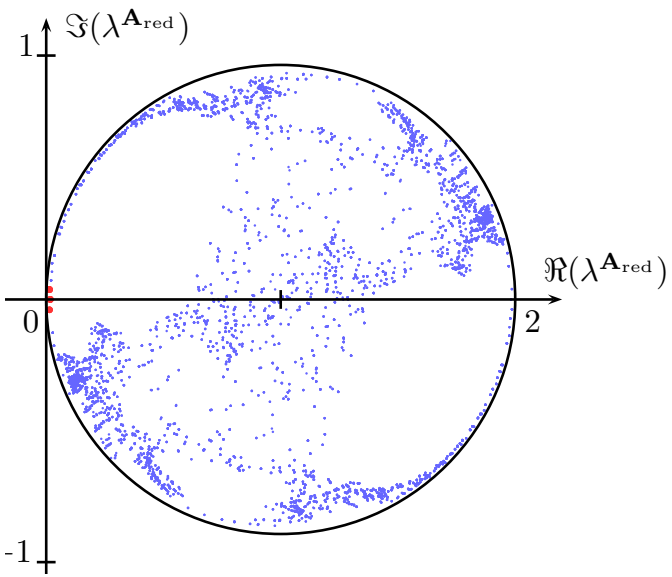


Figure: Real and complex parts of the spectrum $\lambda^{\mathbf{A}_{\mathrm{red}}}$ of $\mathbf{A}_{\mathrm{red}}$.

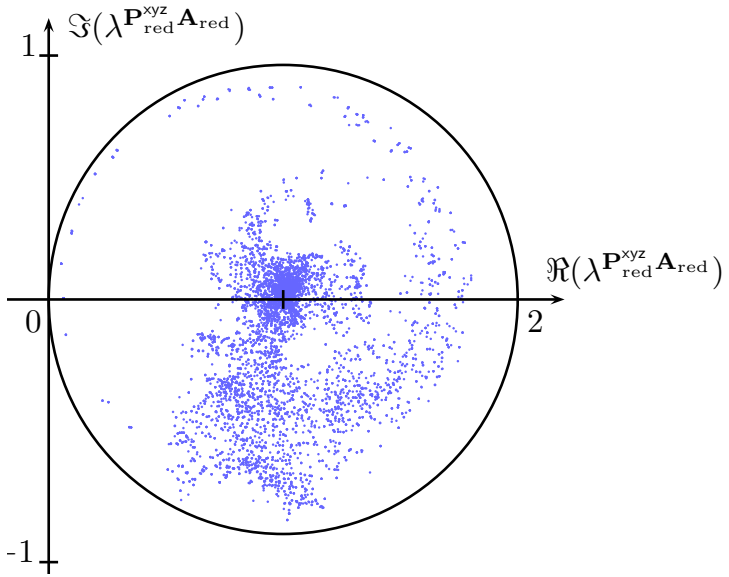


Figure: Real and complex parts of the spectrum $\lambda^{\mathbf{P}_{\mathrm{red}}^{\mathrm{xyz}}\mathbf{A}_{\mathrm{red}}}$ of $\mathbf{P}_{\mathrm{red}}^{\mathrm{xyz}}\mathbf{A}_{\mathrm{red}}$.

Faster convergence with the left global preconditioner

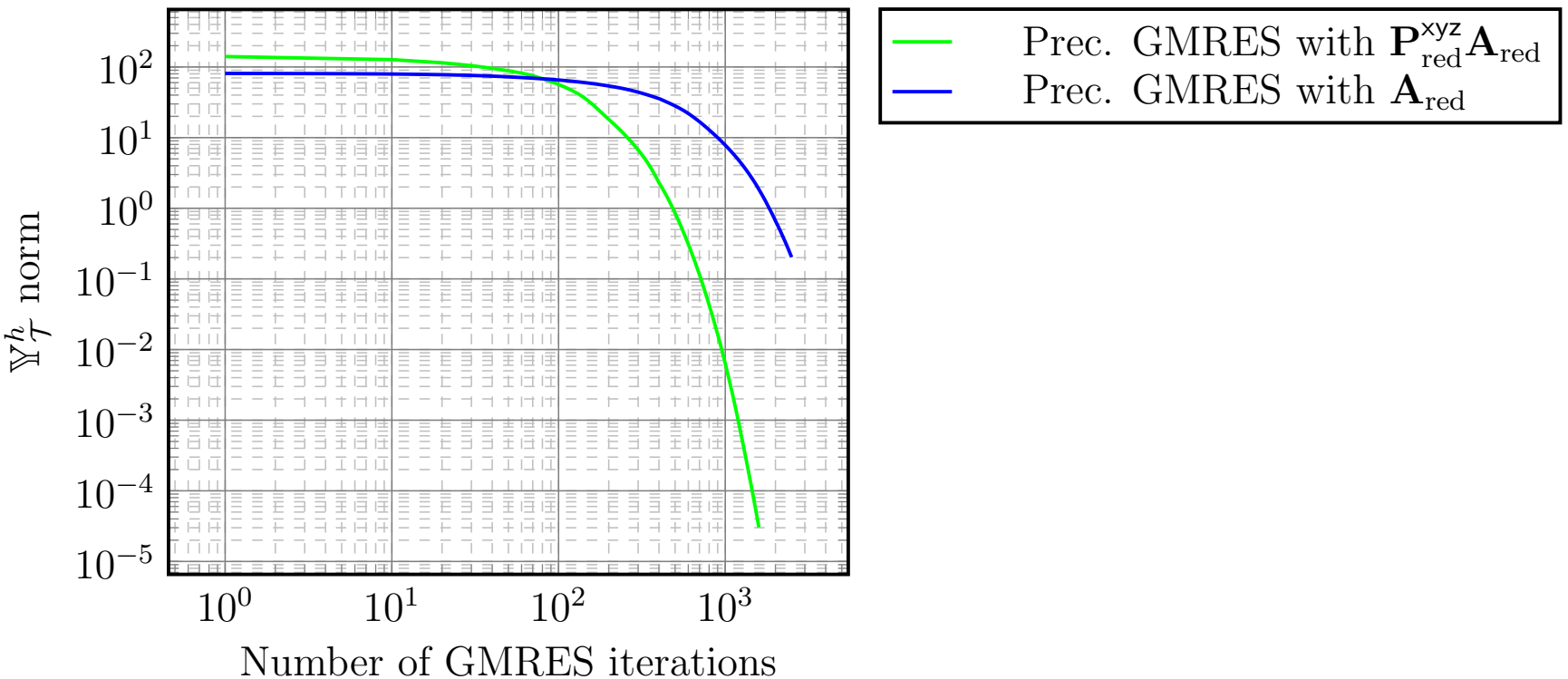
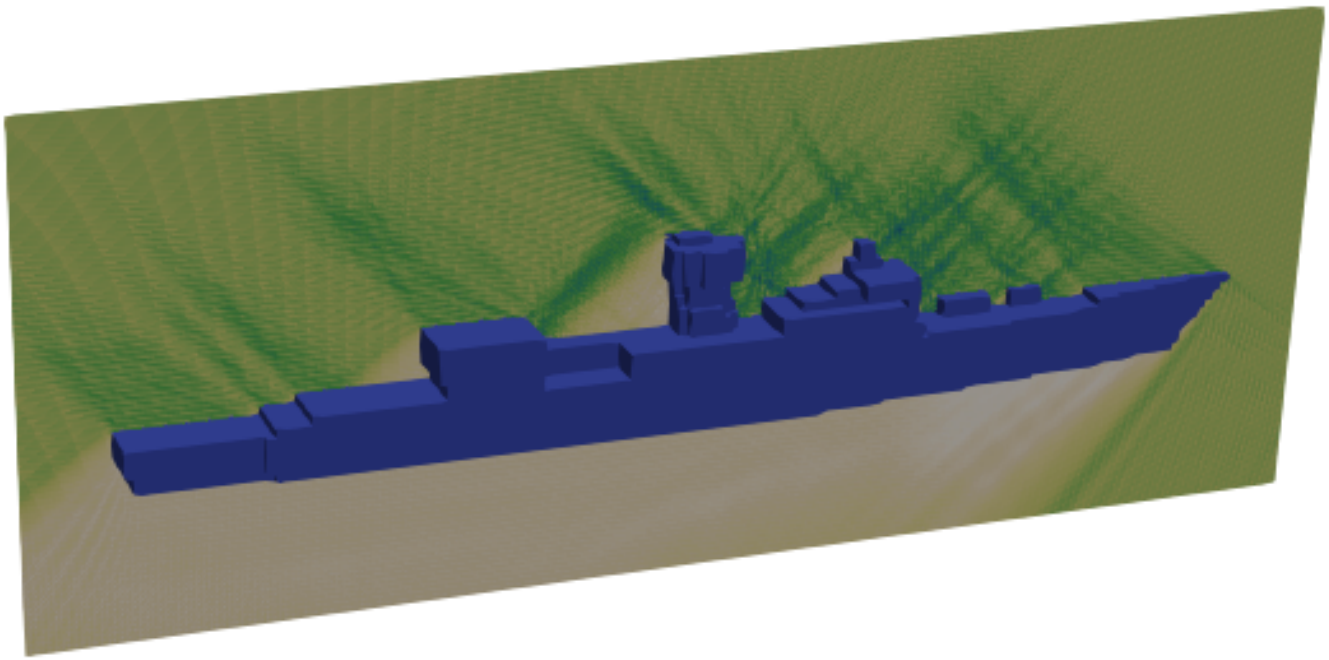


Figure: Comparison of the convergence rate of the preconditioned GMRES UWVF using either $\mathbf{A}_{\mathrm{red}}$ or $\mathbf{P}_{\mathrm{red}}^{\mathrm{xyz}}\mathbf{A}_{\mathrm{red}}$, with $\mathbb{Y}_{\mathcal{T}}^{h}$ the UWVF norm.

Heterogeneous visualization

Interpretation: complex nearby electromagnetic field and shadow under the boat.

Boat size $= 24 \times 61 \times 154$ meters. Cartesian mesh $= 96 \times 244 \times 616$ elements.



Three-dimensional electromagnetic wave striking the top of a boat, for $\lambda = 1$ meter, $\lambda/4$ mesh.

#elem	N_{red}	N	$\#\text{dof}_{\text{red}}$	#dof
$> 14.4 \times 10^6$	46	52	$> 0.663 \times \mathbf{10^9}$	$> 0.75 \times \mathbf{10^9}$

GMRES iterations	Memory cost	Total duration	GMRES residual
800	389 GB	28.6 hours	3.8×10^{-2}

Table: Numerical data and results associated to the boat case.

Thank you for your attention !

References, preprint:

- Fure, H.S., Pernet, S., Sirdey, M. et al. A discontinuous Galerkin Trefftz type method for solving the two dimensional Maxwell equations. SN Partial Differ. Equ. Appl. 1, 23 (2020).
- S. Pernet, N. Serdiuk, M. Sirdey, S. Tordeux. Discontinuous Galerkin Method based on Riemann fluxes for the time domain Maxwell System. [Research Report] RR-9431, INRIA Bordeaux - Sud-Ouest. 2021, pp.55. hal-03396721
- S. Pernet, M. Sirdey, S. Tordeux. Ultra-Weak variational formulation for heterogeneous Maxwell problem in the context of high performance computing. 2022. hal-03642116

Mailing address: margot.sirdey@inria.fr