

Preconditioning of GMRES for Helmholtz problems with quasimodes

Victorita Dolean^[0000-0002-5885-1903],
Pierre Marchand^[0000-0002-2522-6837],
Axel Modave^[0000-0002-9145-6585],
Timothée Raynaud^[0009-0007-1780-4135]

1 Introduction

Time-harmonic scalar wave propagation can be modeled by the Helmholtz equation:

$$-\Delta u - k^2 u = f,$$

where $k > 0$ is the wave number, u the unknown field, and f a given source. Boundary conditions (and/or radiation conditions if the problem is defined in an unbounded domain) are imposed to ensure a well-posed problem. For numerical solutions, the computational domain is truncated, and the finite element method gives a linear system

$$\mathbf{A}\mathbf{u} = \mathbf{b}, \tag{1}$$

where $\mathbf{A} \in \mathbb{C}^{N \times N}$ is sparse, non-singular, and typically non-Hermitian. The unknown and the given source are $\mathbf{u}, \mathbf{b} \in \mathbb{C}^N$. To guarantee the accuracy of the solution and counteract the pollution effect, the system can be very large.

Krylov solvers such as GMRES are well-suited for solving these large non-Hermitian linear systems. However, the convergence can be extremely slow due to the highly indefinite nature of the problem. Then, well-designed acceleration techniques, such as domain decomposition methods, are often necessary to ensure convergence within a reasonable time frame.

The numerical solution is even more difficult for quasisonant Helmholtz problems. Quasimodes occur when a Helmholtz problem with Sommerfeld radiation

Victorita Dolean
Department of Mathematics and Computer Science, Eindhoven University of Technology, P.O. Box 513, 5600 MB Eindhoven, The Netherlands, e-mail: v.dolean.maini@tue.nl

Pierre Marchand · Axel Modave · Timothée Raynaud
POEMS, CNRS, INRIA, ENSTA, Institut Polytechnique de Paris, 91120 Palaiseau, France, e-mail: pierre.marchand@inria.fr, axel.modave@ensta.fr, and timothee.raynaud@ensta.fr

conditions is solved in an unbounded domain. They refer to an increasing sequence of wave numbers such that the norm of the inverse operator increases quickly [8]:

Definition 1 (Quasimodes [8, Definition 1.1]) Let $H_{\text{loc}}^1(\Omega)$ be the space of functions locally in $H^1(\Omega)$, where $\Omega \subset \mathbb{R}^2$ is unbounded. A family of Dirichlet (or Neumann) quasimodes with quality $\epsilon(k)$ is a sequence $\{(u_j, k_j)\}_{j=1}^\infty \subset H_{\text{loc}}^1(\Omega) \times \mathbb{R}$, where $u_j = 0$ (or $\partial_n u_j = 0$) on Γ , such that $k_j \rightarrow \infty$ as $j \rightarrow \infty$, and there exists a compact set $K \subset \Omega$ such that, for all j ,

$$\text{supp}(u_j) \subset K, \quad \|\Delta u_j - k_j^2 u_j\|_{L^2(\Omega)} \leq \epsilon(k_j), \quad \text{and} \quad \|u_j\|_{L^2(\Omega)} = 1.$$

The effect of quasimodes on the GMRES convergence has been studied [4, 8], but not in the context of domain decomposition methods. In this work, we investigate the impact of quasimodes and the deflation of their approximate expressions on the GMRES convergence, within the framework of domain decomposition.

2 GMRES convergence

Here, we briefly present the Generalized Minimal Residual (GMRES) method [9] and its properties, and we introduce a residual estimate based on the harmonic Ritz (HR) values.

GMRES is an iterative algorithm for solving general linear systems such as (1), which seeks approximate solutions in the Krylov subspaces

$$\mathcal{K}_l(\mathbf{A}, \mathbf{r}_0) := \text{span}\{\mathbf{r}_0, \mathbf{A}\mathbf{r}_0, \mathbf{A}^2\mathbf{r}_0, \dots, \mathbf{A}^{l-1}\mathbf{r}_0\}, \quad \text{for } l < N,$$

where $\mathbf{r}_0 := \mathbf{b} - \mathbf{A}\mathbf{x}_0$ is the initial residual, with \mathbf{x}_0 an initial guess. At each iteration $l < N$, the approximate solution is computed such that the two-norm of the residual $\mathbf{r}_l := \mathbf{b} - \mathbf{A}\mathbf{x}_l$ is minimized:

$$\mathbf{x}_l := \underset{\mathbf{x} \in \mathbf{x}_0 + \mathcal{K}_l(\mathbf{A}, \mathbf{r}_0)}{\text{argmin}} \|\mathbf{b} - \mathbf{A}\mathbf{x}\|_2 \in \mathbf{x}_0 + \mathcal{K}_l(\mathbf{A}, \mathbf{r}_0). \quad (2)$$

To study the GMRES convergence, it is usual to consider an equivalent problem for (2) by expressing the residual at iteration l such that:

$$\|\mathbf{r}_l\|_2 = \min_{q_l \in \mathcal{P}_l^1} \|q_l(\mathbf{A})\mathbf{r}_0\|_2, \quad (3)$$

where \mathcal{P}_l^1 is the set of polynomials of degree at most l with the constraint $q_l(0) = 1$ [5]. Here, the notation $p_l \in \mathcal{P}_l^1$ refers to the polynomial such that $\mathbf{r}_l = p_l(\mathbf{A})\mathbf{r}_0$.

Many GMRES convergence bounds are established in the literature, see e.g. [5]. However, most of them fail to capture the non-linear behavior of the convergence rate, which is typically observed for quasiresonant Helmholtz problems [4]. Harmonic Ritz (HR) values, which approach the eigenvalues of \mathbf{A} during iterations [11, Chapter 26], can be used to interpret the non-linear behavior of the GMRES convergence rate [2, 4].

Definition 2 (Harmonic Ritz (HR) values) At iteration $l > 0$, the *harmonic Ritz values* $\{v_j^{(l)}\}_{j=1}^l$ are the roots of the minimizing polynomial $p_l \in \mathcal{P}_l^1$ from (3).

Theorem 1 (See [4, Theorem 2.6]) Let \mathbf{A} be a non-singular and diagonalizable matrix. Let $J \in [1, l]$. Consider:

- a set $N_J^{(l)}$ of J HR values at iteration l ;
- a set Λ_J of J eigenvalues of \mathbf{A} ;
- the set $\Lambda_J^c := \sigma(\mathbf{A}) \setminus \Lambda_J$ of the other eigenvalues of \mathbf{A} ;
- $s_J^l(z) := \prod_{\lambda_j \in \Lambda_J} (1 - z/\lambda_j) \cdot \prod_{v_j^{(l)} \in N_J^{(l)}} (1 - z/v_j^{(l)})^{-1}$.

Then, for any $m > 0$ such that $l + m < N$, the local relative residual satisfies

$$\frac{\|\mathbf{r}_{l+m}\|_2}{\|\mathbf{r}_l\|_2} \leq \left(\sum_{\lambda_i \in \Lambda_J^c} \kappa(\lambda_i) \right) \max_{\lambda_i \in \Lambda_J^c} |s_J^l(\lambda_i)| \min_{q_m \in \mathcal{P}_m^1} \max_{\lambda_i \in \Lambda_J^c} |q_m(\lambda_i)|,$$

with $\kappa(\lambda_i) := \|\hat{\mathbf{v}}_i\|_2 \|\mathbf{v}_i\|_2$, where $\hat{\mathbf{v}}_i$ and $\mathbf{v}_i \in \mathbb{C}^n$ are the left and right eigenvectors associated to $\lambda_i \in \sigma(\mathbf{A})$, with $\hat{\mathbf{v}}_i^* \mathbf{v}_i = 1$ (see [11, Chapter 52]).

As explained in [4], if all the HR values of $N_J^{(l)}$ are close to the eigenvalues in Λ_J at iteration l , then $\max_{\lambda_i \in \Lambda_J^c} |s_J^l(\lambda_i)| \approx 1$. For the following iterations, the convergence bound then depends only on the other eigenvalues in Λ_J^c .

The non-linear GMRES convergence rate behavior can be interpreted in local steps with the ratio $\|\mathbf{r}_{l+m}\|_2 / \|\mathbf{r}_l\|_2$.

3 Domain decomposition preconditioning with deflation

To deal with large-scale Helmholtz problems such as (1), it can be necessary to resort to domain decomposition methods. Also, the presence of quairesonances can hinder the GMRES convergence, and deflation is a useful tool to remove such effect [4]. In this section, we present both methods and their combination.

The computational domain Ω is decomposed into S overlapping subdomains Ω_s ($s = 1, \dots, S$), with n_s the local number of inside dofs.

The one-level method used here is the Optimized Restricted Additive Schwarz (ORAS) preconditioner [3]. It requires restriction matrices $\mathbf{R}_s \in \mathbb{R}^{n_s \times N}$, which restrict a global vector to the subdomain Ω_s , and the transpose \mathbf{R}_s^T that extends a local vector by zero outside Ω_s . Define diagonal matrices $\mathbf{D}_s \in \mathbb{R}^{n_s \times n_s}$ such that

$$\mathbf{I} = \sum_{s=1}^S \mathbf{R}_s^T \mathbf{D}_s \mathbf{R}_s \in \mathbb{R}^{N \times N}.$$

Then, by noting $\mathbf{B}_s \in \mathbb{C}^{n_s \times n_s}$ the local matrices of discretized physical problem, with impedance boundary conditions (BC) on the subdomains interfaces, the ORAS preconditioner reads

$$\mathbf{M}_{\text{oras}}^{-1} := \sum_{s=1}^S \mathbf{R}_s^T \mathbf{D}_s \mathbf{B}_s^{-1} \mathbf{R}_s.$$

Deflation removes some well-chosen eigenvalues of \mathbf{A} by using projection operators, see e.g. [6]. We assume that we have an exact or approximate expression of $n_{\text{def}} \ll N$ eigenvectors, which we store in the columns of a matrix $\mathbf{Z} \in \mathbb{C}^{N \times n_{\text{def}}}$.

Definition 3 (Deflation matrices) Let $\mathbf{Z} \in \mathbb{C}^{N \times n_{\text{def}}}$ be a full rank matrix such that $\ker(\mathbf{Z}^*) \cap \text{range}(\mathbf{AZ}) = \{0\}$, and let $\mathbf{E} := \mathbf{Z}^* \mathbf{AZ} \in \mathbb{C}^{n_{\text{def}} \times n_{\text{def}}}$ and $\mathbf{Q} := \mathbf{Z} \mathbf{E}^{-1} \mathbf{Z}^*$. The *deflation matrices* are defined as

$$\mathbf{P}_{\text{def}} := \mathbf{I} - \mathbf{AQ} \quad \text{and} \quad \mathbf{Q}_{\text{def}} := \mathbf{I} - \mathbf{QA}.$$

The deflation matrices are projectors onto the complement of the deflated subspace. Applying \mathbf{P}_{def} or \mathbf{Q}_{def} to \mathbf{A} removes the components associated with the deflated eigenvectors.

To combine deflation with domain decomposition preconditioning, we consider here the adapted deflation preconditioner \mathbf{P}_{adef} defined as $\mathbf{P}_{\text{adef}} := \mathbf{M}_{\text{oras}}^{-1} \mathbf{P}_{\text{def}} + \mathbf{Q}$.

The deflated and preconditioned problem to solve is then

$$\left| \text{Find } \mathbf{x} \in \mathbb{C}^N \text{ such that } \mathbf{AP}_{\text{adef}} \mathbf{u} = \mathbf{b}, \quad \mathbf{x} = \mathbf{P}_{\text{adef}} \mathbf{u}. \right. \quad (4)$$

To ensure the non-singularity of (4), the required conditions [6, Theorem 3.10] are

$$\ker(\mathbf{Z}^*) \cap \text{range}(\mathbf{AZ}) = \{0\} \quad \text{and} \quad \ker(\mathbf{Z}^*) \cap \text{range}(\mathbf{M}_{\text{oras}} \mathbf{Z}) = \{0\}. \quad (5)$$

If the columns of \mathbf{Z} are eigenvectors of \mathbf{A} , then the hypothesis of Definition 3 holds and \mathbf{E} is invertible. In practice, we assume that the second condition of (5) is satisfied.

4 Numerical investigations

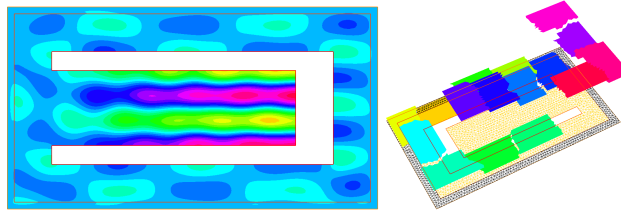


Fig. 1 Left: real part of the solution of (6). Right: domain decomposition ($S = 16$) with the mesh.

We consider the scattering of a time-harmonic plane wave with incident angle θ $u_{\text{inc}}(x, y) = e^{ik(\cos(\theta)x + \sin(\theta)y)}$, by an obstacle \mathcal{O} in \mathbb{R}^2 with Neumann BC on its surface Γ_{obs} .

The computational area is restricted to a rectangular domain $\Omega_{\text{dom}} = (-L_x, L_x) \times (-L_y, L_y) \setminus \mathcal{O}$ and surrounded with perfectly matched layers (PMLs) Ω_{pml} , of thickness L_{pml} [1]. The external boundary is denoted by Γ_{ext} . The PML formulation

consists in a change of variables, with complex-valued stretching functions

$$\gamma_x(x) := 1 + i\sigma_x(x)/k, \quad \gamma_y(y) := 1 + i\sigma_y(y)/k,$$

where $\sigma_x(x) := \mathbb{1}_{\text{pml}}(x)/(L_{\text{pml}} - |x| + L_x)$ and $\sigma_y(y) := \mathbb{1}_{\text{pml}}(y)/(L_{\text{pml}} - |y| + L_y)$. The scattered field u then satisfies:

$$\begin{cases} \partial_x (\gamma_y/\gamma_x \partial_x u) + \partial_y (\gamma_x/\gamma_y \partial_y u) + \gamma_x \gamma_y k^2 u = 0 & \text{in } \Omega_{\text{dom}} \cup \Omega_{\text{pml}}, \\ \partial_n u = -\partial_n u_{\text{inc}} & \text{on } \Gamma_{\text{obs}}, \\ u = 0 & \text{on } \Gamma_{\text{ext}}. \end{cases} \quad (6)$$

The obstacle \mathcal{O} is an open rectangular cavity (see Fig. 1) of length $L_{\mathcal{O}}$ and opening width $l_{\mathcal{O}}$. For this configuration, quasimodes are expected (see Definition 1).

They induce small eigenvalues [4] when the wave numbers are close to the resonance frequencies of a closed rectangular cavity with Neumann BC on all sides except the left edge (Dirichlet), given by

$$k_{n,m} = \pi \sqrt{(m + 1/2)^2/L_{\mathcal{O}}^2 + n^2/l_{\mathcal{O}}^2}, \quad m, n > 0. \quad (7)$$

The eigenvectors associated with the small eigenvalues resulting from the quasimodes are close to the eigenmodes of the corresponding closed configuration.

Discretization by finite elements leads to a linear system $\mathbf{A}\mathbf{u} = \mathbf{b}$, where the matrix \mathbf{A} is complex, has complex-valued eigenvalues, and is non-normal.

The following parameters are chosen: $L_{\mathcal{O}} = 1.3$, $l_{\mathcal{O}} = 0.4$, $\theta = 4\pi/10$, P_3 elements, $\text{tol} = 10^{-6}$. For the first set of experiments, the number of dofs per wavelength is fixed to 10, for a total of about $N \approx 3.10^4$. The second experiments are performed with 5 dofs per wavelength and $N \approx 1.10^5$, at a higher wave number.

For running the numerical experiments, we used the framework FFDDM [10] of FreeFEM [7]. The first experiments illustrate how quasimodes affect GMRES convergence with ORAS preconditioning. The second set of experiments investigates the influence of quasimodes and their deflation when coarse space preconditioners are used.

Influence of quasimodes on the GMRES convergence with ORAS preconditioning

We consider the wave number $k = 23.591$ to be close to $k_{0,3}$ in (7). We investigate how the quasimodes affects the GMRES convergence $\mathcal{S} = 8, 16$ subdomains in Fig. 2 (right).

With $\mathcal{S} = 8$ and 16 (black), a stagnation phase is observed during the mid-iterations, before a fast convergence phase. This behavior is typical for quasiresonant Helmholtz problems [4], and Theorem 1 gives an interpretation with HR values.

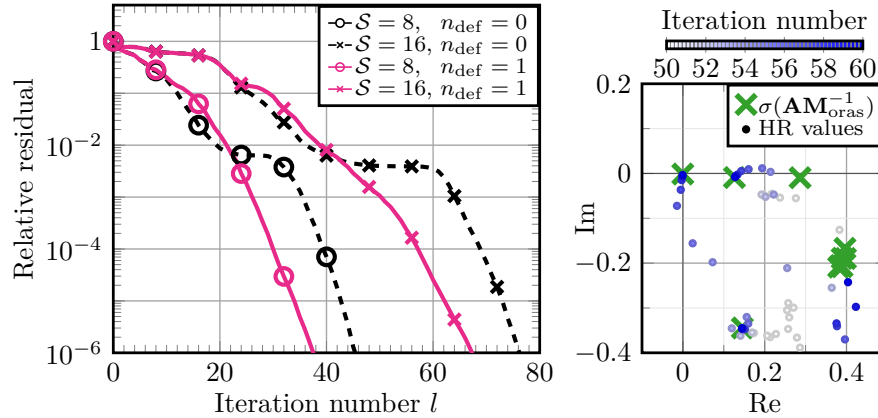


Fig. 2 GMRES convergence with ORAS preconditioning for different number of subdomains S (left). Spectrum of $\mathbf{AM}_{\text{oras}}^{-1}$ and HR values trajectories for $S = 16$ (right).

The spectrum of $\mathbf{AM}_{\text{oras}}^{-1}$ for $S = 16$ (Fig. 2, right) shows that there are a few small eigenvalues with a small imaginary part, which are associated with quasimodes. The positions of the HR values around iteration 50 are plotted in \bullet gray, and they become increasingly blue as the iteration increases up to 60.

We observe that for $S = 16$, the plateau of the residual matches the iterations where the HR values approach the two smallest eigenvalues related to quasimodes. At iteration 60, these small eigenvalues are well approached by HR values. From this iteration, Theorem 1 indicates that the influence of the small eigenvalues on the convergence is removed, so the convergence rate increases.

This analysis reveals that the plateau observed in the middle of the convergence is due to the quasimodes and not only to the number of subdomains. However, it should be noted that when the cavity is divided into a greater number of subdomains, the plateau lasts longer.

In Fig. 2 (left), the same experiments were conducted by deflating the smallest eigenvalue, using an approximation of the closed cavity mode, extended by zero outside the cavity (pink). These experiments confirm that the plateaus are due to quasimodes and disappear when the associated modes are deflated.

Influence of quasimodes with coarse spaces preconditioners

We now consider two-level preconditioners with coarse spaces (CS) built with Dirichlet-to-Neumann (DtN) eigenproblems [3], and we investigate the influence of quasimodes on the GMRES convergence. The same numerical experiments have been carried out with H-GenEO CS [3] and have yielded similar results (not shown here for sake of brevity).

In Fig. 3, the evolution of the relative residual is plotted for $k = 102.11$ (close to $k_{0,13}$) with $S = 16$ and two coarse space sizes (n_{CS}).

In the residual history (green), the first stagnation phase is related to the number of subdomains, and it is significantly reduced by the coarse space. However, several late plateaus remain. By deflating the small eigenvalues associated with the quasimodes, the late stagnation phases disappear (pink). So, these late plateaus are due to the quasimodes. The GMRES convergence preconditioned by this type of coarse space is therefore also hindered by the quasimodes.

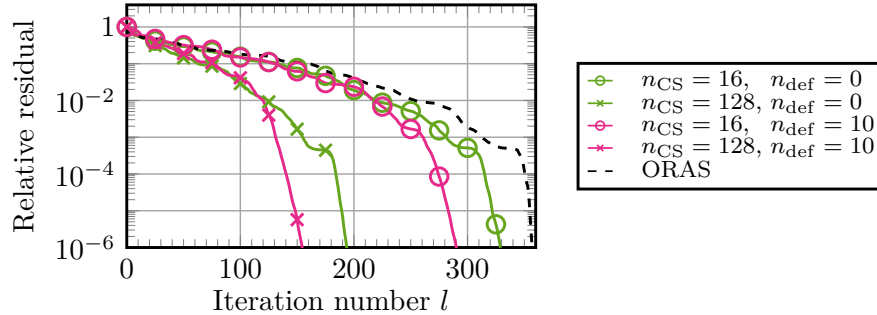


Fig. 3 GMRES convergence with DtN CS, with and without additional deflation contribution for $k = 102.11$ and $S = 16$.

| | $n_{CS} + n_{def} = 32$ | | | $n_{CS} + n_{def} = 144$ | | |
|----------------|-------------------------|----------------|----------------|--------------------------|-----------------|----------------|
| CS composition | $n_{CS} = 32$ | $n_{CS} = 0$ | $n_{CS} = 16$ | $n_{CS} = 144$ | $n_{CS} = 0$ | $n_{CS} = 128$ |
| | $n_{def} = 0$ | $n_{def} = 32$ | $n_{def} = 16$ | $n_{def} = 0$ | $n_{def} = 144$ | $n_{def} = 16$ |
| #iter | 302 | 311 | 273 | 180 | 236 | 141 |

Table 1 Number of GMRES iterations for two second-level sizes comparing the contributions of the DtN CS, deflation, and the combination of the two, for $k = 102.11$, $S = 16$.

To deal with large-scale problems with quasimodes, it can be interesting to combine standard coarse space preconditioning with deflation of quasimodes. Tab. 1 shows the number of GMRES iterations required to reach $tol = 10^{-6}$, with two fixed coarse space sizes ($n_{CS} + n_{def} = 32$ or 144), by varying the contributions of the DtN CS and of the deflation. The deflation only is the least efficient strategy, because the local deflated vectors cannot handle the global exchange of information ensured by the coarse spaces (first stagnation phase in Fig. 3). Using only the DtN CS is efficient, but as previously mentioned, it is sensitive to the quasimodes. The third and sixth columns correspond to using both DtN CS and deflation. This combined approach is the most efficient.

For geometries with explicit or approximate expressions of the quasimodes, adding a few deflated vectors to the coarse spaces can thus be worthwhile. Be-

sides, in this case, it is a priori less costly numerically than increasing the size of a coarse space.

5 Conclusion

This work highlighted the impact of quasimodes on the GMRES convergence in the context of domain decomposition methods.

The study with HR values showed that dividing a quairesonant cavity into several subdomains generates small eigenvalues that hinder GMRES convergence. The more the cavity is divided, the longer the stagnation phase.

Preconditioning with DtN coarse spaces is not sufficient to overcome the impact of quasimodes. A combination of these preconditioners and deflation of quasimodes is robust against the number of subdomains (coarse spaces contribution), and the plateaus due to the small eigenvalues related to quasimodes (deflation contribution).

Acknowledgements This work was funded in part by the *ANR JCJC project WavesDG* (research grant ANR-21-CE46-0010) and by the *Agence de l'Innovation de Défense [AID]* through *Centre Interdisciplinaire d'Etudes pour la Défense et la Sécurité [CIEDS]* (project 2022 ElectroMath). The authors would like to thank Pierre-Henri Tournier for his help with FreeFEM implementation.

References

1. Bermúdez, A., Hervella-Nieto, L., Prieto, A., Rodríguez, R.: An exact bounded PML for the Helmholtz equation. *C. R. Math.* **339**, 803–808 (2004)
2. Cao, Z.H.: A note on the convergence behavior of GMRES. *Appl. Numer. Math.* **25**, 13–20 (1997)
3. Dolean, V., Fry, M., Langer, M., Parolin, E., Tournier, P.H.: Achieving wavenumber robustness in domain decomposition for heterogeneous helmholtz equation: an overview of spectral coarse spaces (2025). ArXiv preprint arXiv:2509.02131
4. Dolean, V., Marchand, P., Modave, A., Raynaud, T.: Convergence analysis of GMRES applied to helmholtz problems near resonances (2025). ArXiv preprint arXiv:2505.16345
5. Embree, M.: How Descriptive are GMRES Convergence Bounds? (2022). ArXiv preprint arXiv:2209.01231
6. García Ramos, L., Kehl, R., Nabben, R.: Projections, Deflation, and Multigrid for Nonsymmetric Matrices. *SIAM J. Matrix Anal. Appl.* **41**, 83–105 (2020)
7. Hecht, F.: New development in freefem++. *J. Numer. Math.* **20**, 251–265 (2012)
8. Marchand, P., Galkowski, J., Spence, E.A., Spence, A.: Applying GMRES to the Helmholtz equation with strong trapping: how does the number of iterations depend on the frequency? *Adv. Comput. Math.* **48** (2022)
9. Saad, Y., Schultz, M.H.: GMRES: A Generalized Minimal Residual Algorithm for Solving Nonsymmetric Linear Systems. *SIAM J. Sci. Stat. Comput.* **7**, 856–869 (1986)
10. Tournier, P.H., Jolivet, P., Nataf, F.: FFDDM: Freefem domain decomposition method (2019)
11. Trefethen, L.N., Embree, M.: *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*. Princeton Univ. Press (2005)