A coarse space for heterogeneous Helmholtz problems based on the Dirichlet-to-Neumann operator

Lea Conen, Victorita Dolean, Rolf Krause, Frédéric Nataf

A coarse space for heterogeneous Helmholtz problems based on the Dirichlet-to-Neumann operator

Lea Conen[a,*], Victorita Dolean[b], Rolf Krause[a], Frédéric Nataf[c]

[a] *Università della Svizzera italiana, Institute of Computational Science, Via G. Buffi 13, 6900 Lugano, Switzerland*
[b] *Université de Nice-Sophia Antipolis, Laboratoire J.-A. Dieudonné, Parc Valrose, 060108 Nice Cedex 2, France*
[c] *Université Pierre et Marie Curie, Laboratoire J.L. Lions, Tour 15-25 Bureau 319, 4 place Jussieu, 75005 Paris, France*

Abstract

The Helmholtz equation governing wave propagation and scattering phenomena is difficult to solve numerically. Its discretization with piecewise linear finite elements results in typically large linear systems of equations. The inherently parallel domain decomposition methods constitute hence a promising class of preconditioners. An essential element of these methods is a good coarse space. Here, the Helmholtz equation presents a particular challenge, as even slight deviations from the optimal choice can be devastating.

In this paper, we present a coarse space that is based on local eigenproblems involving the Dirichlet-to-Neumann operator. Our construction is completely automatic, ensuring good convergence rates without the need for parameter tuning. Moreover, it naturally respects local variations in the wave number and is hence suited also for heterogeneous Helmholtz problems. The resulting method is parallel by design and its efficiency is demonstrated on 2D homogeneous and heterogeneous numerical examples.

Keywords: Helmholtz equation, domain decomposition, coarse space, Dirichlet-to-Neumann operator

1. Introduction

The Helmholtz equation

$$-\Delta u - k^2 u = f \tag{1}$$

with suitable boundary conditions and wave number $k > 0$ governs wave propagation and scattering phenomena arising in a wide range of engineering applications, such as aeronautics, underwater acoustics, and geophysical seismic

*Corresponding author. Phone number: +41 (0)58 666 4975
Email addresses: lea.conen@usi.ch (Lea Conen), dolean@unice.fr (Victorita Dolean), rolf.krause@usi.ch (Rolf Krause), nataf@ann.jussieu.fr (Frédéric Nataf)

imaging. Its discretization with piecewise linear finite elements results for large wave number k in an indefinite, ill-conditioned linear system of equations. Indefiniteness could be avoided using a non-standard variational formulation [1]. The number of grid points grows rapidly with k in order to maintain accuracy and to avoid the pollution effect [2]. As the linear system of equations is hard to solve for iterative methods [3, 4], carefully designed methods are necessary.

There has been a vast amount of research on this topic, including incomplete factorization methods [5–8] and the "sweeping preconditioner" [9]. Preconditioning with a shifted, easier problem has been considered e.g. in [10, 11]. The problems encountered when multigrid methods are applied to the Helmholtz equation have been analyzed in detail [12, 13]. Particularly interesting is the wave-ray multigrid [12], where special levels based on plane waves that lie in the kernel of the homogeneous, continuous Helmholtz operator are introduced, designed to represent the oscillatory part of the solution.

In this work, we concentrate on domain decomposition methods (DDMs), see e.g. [14]. As the systems of linear equations resulting from the discretization of Problem (1) are typically large, these methods constitute due to their inherent parallelism an interesting class of preconditioners. DDMs have two main ingredients: the transmission conditions, specifying the information exchange between neighboring subdomains and the coarse space, allowing for global transfer of information. Unfortunately, classical choices for either of these two parts are not effective for the Helmholtz equation. This led to the development of specially adapted methods, for a numerical comparison see [15].

Early work on transmission conditions for the Helmholtz equation was done in [16], where a first order approximation to the Sommerfeld radiation condition is employed. In the sequel, different, more advanced techniques have been used; including PML at the interfaces [17–19], non-local transmission conditions [20], optimized Schwarz methods [21], and others, e.g. [22].

Originally used in the multigrid context [12], plane waves have been successfully employed also as coarse space basis functions for DDMs. Their evaluation at the subdomains' interfaces are used in the FETI(-DP)-H methods [23, 24]. Later, they have also been used in other DDMs [25, 26] and as deflation vectors [27]. Plane waves, to our knowledge, have been employed mainly for homogeneous problems; the extension to heterogeneous ones is not obvious.

In this paper, we concentrate on the development of a coarse space for a restricted additive Schwarz method. We adapt an idea for elliptic problems where the coarse space is based on local functions, the solutions of eigenproblems involving the Dirichlet-to-Neumann operator on the subdomains' interfaces [28, 29]. Our coarse space is robust with respect to heterogeneous coefficients and its construction is completely automatic, refraining from the need for parameter tuning. The latter feature is crucial for indefinite problems as in contrast to the elliptic case, even slight deviations from the optimal choice can be fatal [30].

The paper is organized as follows. The Helmholtz equation is introduced in Sec. 2, the two-level DDM in Sec. 3. In Sec. 4, we motivate and define our coarse space. Emphasis is put on the development of a criterion for choosing the coarse space size automatically. Sec. 5 elaborates on the effect the second

level has on spectrum and convergence rates. In Sec. 6, we test the DtN coarse space numerically and compare it to the standard one based on plane waves.

2. Helmholtz equation and discretization

We are interested in the interior Helmholtz problem of the following form: Let $\Omega \subset \mathbb{R}^d$, $d = 2, 3$, be a polygonal, bounded domain. Find $u : \Omega \to \mathbb{C}$ s.t.

$$-\Delta u - k^2 u = f \qquad \text{in } \Omega, \tag{2a}$$

$$u = 0 \qquad \text{on } \Gamma_D, \tag{2b}$$

$$\frac{\partial u}{\partial n} + \imath k u = 0 \qquad \text{on } \Gamma_R, \tag{2c}$$

where $\Gamma_D \cup \Gamma_R = \Gamma := \partial\Omega$ is a disjoint partition of $\partial\Omega$. We abbreviate the boundary conditions in the form $\mathcal{C}(u) = 0$ on Γ. The wave number k is given by $k(\vec{x}) = \omega/c(\vec{x})$, where ω is the angular frequency and c is the speed of propagation that might depend on $\vec{x} \in \Omega$. Equation (2c) is designed to approximate an unbounded domain and is a first order approximation of the Sommerfeld radiation condition, c.f. [31].

For $H^1(\Omega)$ the usual Sobolev space of order 1 on Ω, the variational formulation of Problem (2) is: Find $u \in V := \left\{u \in H^1(\Omega) : \ u = 0 \text{ on } \Gamma_D\right\}$ s.t.

$$a(u, v) = F(v) \qquad \forall v \in V, \tag{3}$$

where $a(.,.) : V \times V \to \mathbb{C}$ and $F : V \to \mathbb{C}$ are defined by

$$a(u, v) = \int_\Omega \left(\nabla u \nabla v - k^2 uv\right) \, \mathrm{d}x + \int_{\Gamma_R} \imath k uv \, \mathrm{d}s, \quad F(v) = \int_\Omega fv \, \mathrm{d}x.$$

It is well-posed if $\Gamma_R \neq \emptyset$, c.f. [32, Chapter 2]. We consider a discretization of the variational problem (3) using piecewise linear finite elements on a uniform triangular mesh $\mathcal{T}_h$ of Ω. Denoting by $V_h \subset V$ the corresponding finite element space, it reads: Find $u_h \in V_h$ such that

$$a(u_h, v_h) = F(v_h) \qquad \forall v_h \in V_h. \tag{4}$$

With $\{\phi_k\}_{k=1}^n$ the nodal linear finite element basis for V_h, $n := \dim(V_h)$, we rewrite (4) in matrix form:

$$A\vec{u} = \vec{f}, \tag{5}$$

where the coefficients of the stiffness matrix $A \in \mathbb{C}^{n\times n}$ and the right-hand side $\vec{f} \in \mathbb{C}^n$ are given by $A_{k,l} = a(\phi_l, \phi_k)$ and $\vec{f_k} = F(\phi_k)$. The resulting matrix is indefinite and without the Sommerfeld boundary condition possibly singular. It is complex symmetric when $\Gamma_R \neq \emptyset$.

3. Two-level restricted additive Schwarz method

This section defines the domain decomposition method (DDM) that we use as a preconditioner for the Helmholtz equation (5). It is a two-level restricted additive Schwarz (RAS) method [14] with suitable transmission conditions.

We partition the domain Ω into a set of non-overlapping subdomains $\{\Omega_j'\}_{j=1}^N$ resolved by the mesh $\mathcal{T}_h$. The overlapping subdomains Ω_j are then defined by adding one or several layers of mesh elements to Ω_j' in the following sense:

Definition 3.1. Given a subdomain $D' \subset \Omega$, which is resolved by the finite element mesh $\mathcal{T}_h$, the extension D of D' by one layer of elements is

$$D = \operatorname{Int}\left(\bigcup_{\operatorname{supp}(\phi_k)\cap D' \neq \emptyset} \operatorname{supp}(\phi_k)\right),$$

where $\operatorname{Int}(\cdot)$ denotes the interior of a domain. Extensions by more than one layer are defined recursively.

This gives an overlapping partition $\{\Omega_j\}$ of Ω. Let $V_h(\Omega_j) = \left\{v|_{\Omega_j} : v \in V_h\right\}$, $1 \leq j \leq N$, denote the space of functions in V_h restricted to the subdomain Ω_j. Let $n := |\operatorname{dof}(\Omega)|$ and $n_j := |\operatorname{dof}(\Omega_j)|$, $1 \leq j \leq N$, where for $D \subseteq \Omega$ we define $\operatorname{dof}(D) := \left\{k : \operatorname{supp}(\phi_k) \subset \overline{D}\right\}$.

For $1 \leq j \leq N$, we define a restriction operator $\mathcal{R}_j : V_h \to V_h(\Omega_j)$ by injection, i.e. for $u \in V_h$ we set $(\mathcal{R}_j u)(\vec{x}_i) = u(\vec{x}_i)$ for all $\vec{x}_i \in \Omega_j$. We denote the corresponding matrix in $\mathbb{R}^{n_j \times n}$ that maps coefficient vectors of functions in V_h to coefficient vectors of functions in $V_h(\Omega_j)$ by R_j. Let $D_j \in \mathbb{R}^{n_j \times n_j}$ be a diagonal matrix corresponding to a partition of unity in the sense that $\sum_{i=1}^N \tilde{R}_i^T R_i = I$, where $\tilde{R}_j := D_j R_j$. Then the RAS preconditioner reads

$$M^{-1} := \sum_{j=1}^N \tilde{R}_j^T A_j^{-1} R_j. \tag{6}$$

It remains to define the subdomain matrices A_j. With the classical choice $A_j = R_j A R_j^T$, frequencies in the error smaller than the wave number k are not damped [21]. This might cause slow convergence or even stagnation of the iterative method. Therefore, we define the matrices in A_j in (6) to be the stiffness matrices of the local Robin problems [16, 33]

$$\begin{aligned}
\left(-\Delta - k^2\right)(u_i) &= f && \text{in } \Omega_i,\\
\mathcal{C}(u_i) &= 0 && \text{on } \partial\Omega \cap \partial\Omega_i,\\
\left(\frac{\partial}{\partial n_i} + \imath k\right)(u_i) &= 0 && \text{on } \partial\Omega_i \setminus \partial\Omega.
\end{aligned}$$

More advanced techniques such as optimized boundary conditions [34] are also possible, but not considered here.

DDMs as above do not show convergence rates independent of the number of subdomains [14]. In order to achieve independence, one possibility is to add a coarse space to the iteration procedure (6) via the balancing Neumann-Neumann (BNN) method [35]. For non-symmetric systems, it reads [36]:

$$P_{\mathrm{BNN}} = QM^{-1}P + ZE^{-1}Y^{\dagger}, \tag{7}$$

where $\dagger$ denotes the conjugate transpose, M^{-1} is the one-level preconditioner (6), Z and Y are rectangular matrices with full column rank, $E = Y^{\dagger}AZ$ is the coarse grid matrix, $\Xi = ZE^{-1}Y^{\dagger}$ is the coarse grid correction matrix, and $P = I - A\Xi$ and $Q = I - \Xi A$ are projection matrices. We only consider the case $Z = Y$ and will define Z in Sec. 4.

Following this notation, Z implicitly defines what is called the coarse space W, i.e. the columns of Z represent the basis vectors of W. The choice of Z, and hence of the coarse space W itself, influence the convergence speed of the resulting two-level method (7) significantly. For indefinite systems, the correctness of the coarse space is especially important, as, contrarily to the elliptic case, any deviation from the optimal setting might be fatal, c.f. Sec. 5.1 and [30]. Hence particular emphasis has to be put on the design of W.

4. Dirichlet-to-Neumann coarse space for the Helmholtz equation

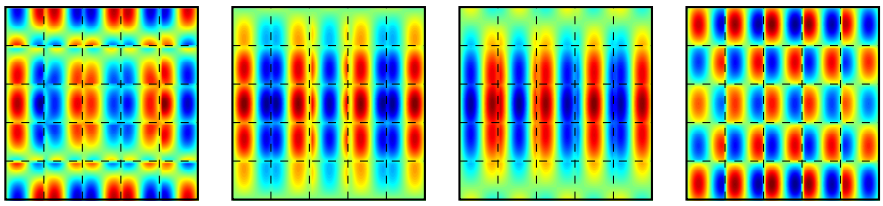
Whereas for certain elliptic problems choices of the coarse space W are known that turn one-level DDMs into optimal solvers with subdomain independent convergence rates [14], for the Helmholtz problem the situation is much more complicated. Various approaches can be found in the literature that aim at designing efficient two-level methods [12, 23–27], often using plane waves as basis functions for the coarse space. Although very elegant by design, this construction does not cover the case of varying coefficients and requires an a priori choice of certain parameters. Here, we therefore aim at constructing a coarse space with the following properties: On the one hand, for constant coefficients it behaves similarly as the one based on plane waves. On the other hand, it is also efficient for heterogeneous coefficients and can be constructed in an automatic, parameter-free fashion. The construction is based on local eigenproblems involving the Dirichlet-to-Neumann operator (DtN), c.f. [29].

4.1. What should the coarse space look like?

As a first step towards the design of the coarse space W, we investigate numerically the properties that it should have. While Fourier analysis detects the flaws of the one-level method (6) in a simplified setting [37], we here aim at getting a better understanding of how the coarse space should be designed by looking at the functions that it contains in the optimal case directly.

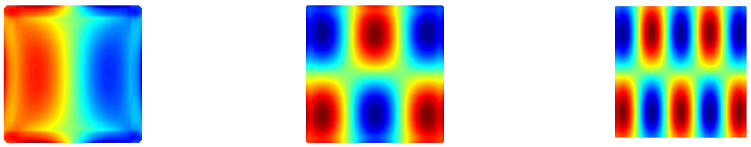
To compute the optimal coarse space functions, we solve the eigenproblem

$$\text{Find } (v_i, \lambda_i) \in \mathbb{C}^n \times \mathbb{C},\ 1 \leq i \leq n, \text{ such that } \left(I - M^{-1}A\right) v_i = \lambda_i v_i, \tag{8}$$



(a) $\mathrm{Re}\,(v_1)$ (b) $\mathrm{Re}\,(v_2)$ (c) $\mathrm{Re}\,(v_3)$ (d) $\mathrm{Re}\,(v_4)$

Figure 1: Real part of optimal coarse space function v_i associated to eigenvalue λ_i, $|\lambda_i| \geq |\lambda_{i+1}|$ for all i, of (8) on Ω. 5×5 subdomains, $n_{\mathrm{loc}} = 40$, $k = 29.3$.



(a) $n_{\mathrm{loc}} = 40$, $k = 29.3$ (b) $n_{\mathrm{loc}} = 80$, $k = 46.5$ (c) $n_{\mathrm{loc}} = 160$, $k = 73.8$

Figure 2: Real part of optimal coarse space function associated to largest eigenvalue of (8) on central subdomain in a 5×5 subdomain decomposition.

and then choose those functions v_i for which the modulus of the associated eigenvalue $|\lambda_i|$ is maximal, since for these functions the preconditioner is not efficient. Even though the v_i are global functions, they have a subdomain structure that is introduced by the preconditioner M^{-1}, see Figure 1. Moreover, one can check numerically that they solve the local problems in the interior of each subdomain away from the overlap. In Fig. 2, we show the dependence of the optimal coarse space functions on the wave number. In accordance with the results obtained with Fourier analysis [21, 37], for the RAS preconditioner (6) with Robin transmission conditions, Fourier frequencies close to the wave number k converge only slowly and hence should be present in the coarse space.

We conclude three guidelines for the design of the coarse space W: *1.* Fourier frequencies close to the wave number k are important. *2.* In the interior of each subdomain, the coarse space functions should lie in the kernel of the discrete Helmholtz operator. *3.* The structure of the optimal functions suggests a subdomain based construction, avoiding the need for global functions.

4.2. Definition of the coarse space based on the Dirichlet-to-Neumann operator

In this section, we introduce our coarse space for the Helmholtz equation. It is based on eigenproblems involving local DtN maps and incorporates naturally the possible heterogeneity in the wave number. The underlying idea originates from work on elliptic problems [28, 29, 38]. This construction respects the main principles outlined in Sec. 4.1: Apart from including all the important modes,

in the interior of each subdomain the coarse functions lie in the kernel of the Helmholtz operator and the construction is based on local problems only. The latter makes it possible to construct the coarse space efficiently in parallel.

The coarse space functions are local functions that are eigenfunctions of the DtN operator on the boundary of a subdomain. We first consider the continuous formulation of the DtN eigenproblems. On each interface $\Gamma_i := \partial\Omega_i \setminus \partial\Omega$ we solve the eigenproblem: Find $(u_{\Gamma_i}, \lambda) \in V(\Gamma_i) \times \mathbb{C}$ such that

$$\mathrm{DtN}_{\Omega_i}(u_{\Gamma_i}) = \lambda u_{\Gamma_i}. \tag{9}$$

For the definition of the DtN operator, we need to extend a function from the boundary of a subdomain to its interior:

Definition 4.1 (Helmholtz extension)**.** Let $D \subset \Omega$, let $\Gamma_D = \partial D \setminus \partial\Omega$. Let $v_{\Gamma_D} : \Gamma_D \to \mathbb{C}$. The extension $u : \Omega_i \to \mathbb{C}$ of v with respect to the Helmholtz operator is defined by

$$\begin{aligned} -\Delta u - k^2 u &= 0 && \text{in } D, \\ \mathcal{C}(u) &= 0 && \text{on } \partial\Omega \cap \partial D, \\ u &= v_{\Gamma_D} && \text{on } \Gamma_D. \end{aligned}$$

The DtN operator is then defined as follows:

Definition 4.2 (Dirichlet-to-Neumann operator)**.** Let $D \subset \Omega$, let $\Gamma_D = \partial D \setminus \partial\Omega$. Let $v_{\Gamma_D} : \Gamma_D \to \mathbb{C}$. Then

$$\mathrm{DtN}_{\mathrm{D}}(v_{\Gamma_D}) = \left.\frac{\partial u}{\partial n}\right|_{\Gamma_D},$$

where $u : D \to \mathbb{C}$ is the extension of v_{Γ_D} in the sense of Def. 4.1.

We choose $m_i \in \mathbb{N}$ eigenfunctions for each subdomain Ω_i according to the following criterion:

Criterion 4.3 (Choice of DtN eigenfunctions)**.** On each subdomain Ω_i, we choose all eigenfunctions v of the DtN eigenproblem (9), for which the associated eigenvalue λ satisfies

$$\mathrm{Re}(\lambda) < k_i.$$

Here $k_i := \max_{\vec{x} \in \Omega_i} k(\vec{x})$ is the maximum wave number on Ω_i. If no eigenvalue satisfies this condition, the eigenvalue with smallest real part is chosen.

This criterion provides a way to automatically construct the coarse space W without the need to tune its dimension, a crucial parameter for the convergence of the two-level method. Numerical evidence that Criterion 4.3 has the desired properties is provided in Sec. 4.3.

We proceed with the discrete formulation of the DtN eigenproblem and explain how to construct the matrix $Z \in \mathbb{C}^{n \times \sum_{j=1}^{N} m_i}$ spanning the coarse space. The columns $1 + \sum_{j=1}^{i-1} m_j, \ldots, \sum_{j=1}^{i} m_j$ of Z are set to $R_i^T W_i$ for $1 \leq i \leq N$.

Here $W_i \in \mathbb{C}^{n_i \times m_i}$ is a rectangular matrix associated to subdomain Ω_i, which is given by Algorithm 4.1. Z is a rectangular, block-diagonal matrix with blocks W_i that may have overlapping rows due to the overlap in the domain decomposition. We now define the single components of Algorithm 4.1.

Algorithm 4.1 Construction of the block W_i of the DtN coarse matrix.

1: Solve the discrete DtN eigenproblem (11) on subdomain Ω_i.
2: Choose m_i eigenvectors $\vec{g}_i^k \in \mathbb{C}^{n_{\Gamma_i}}$, $1 \leq k \leq m_i$ by the discrete analogue of Criterion 4.3.
3: **for** $k \leftarrow 1$ to m_i **do**
4: Compute the extension $\vec{u}_i^k \in \mathbb{C}^{n_i}$ of $\vec{g}_i^k$ according to Def. 4.4.
5: **end for**
6: Define the matrix $W_i \in \mathbb{C}^{n_i \times m_i}$ as $W_i := \left(D_i \vec{u}_i^1, \quad \ldots, \quad D_i \vec{u}_i^{m_i}\right)$.

For Line 1 of Algorithm 4.1, we need the discrete formulation of the DtN eigenproblem (9): Let I and Γ_i be the sets of indices corresponding to the interior and boundary degrees of freedom, respectively. Let n_I and n_{Γ_i} be their cardinalities. We define $a_i : H^1(\Omega_i) \times H^1(\Omega_i) \to \mathbb{R}$,

$$a_i(v, w) = \int_{\Omega_i} \left(\nabla v \cdot \nabla w - k^2 v w\right) \mathrm{d}x.$$

Using the finite element basis $\{\phi_k\}$ for $V(\Omega)$, let $A^{(i)}$ be the coefficient matrix of a Neumann boundary value problem on Ω_i, $A^{(i)}_{kl} = a_i\left(\phi_k, \phi_l\right)$, with boundary conditions defined by $\mathcal{C}$ on $\partial\Omega_i \cap \partial\Omega$. With the usual block notation, the subscripts I and Γ_i for the matrices A and $A^{(i)}$ denote the entries of these matrices associated to the respective degrees of freedom. Let

$$M_{\Gamma_i} = \left(\int_{\Gamma_i} \phi_k \phi_l \, \mathrm{d}s\right)_{k,l \in \Gamma_i}$$

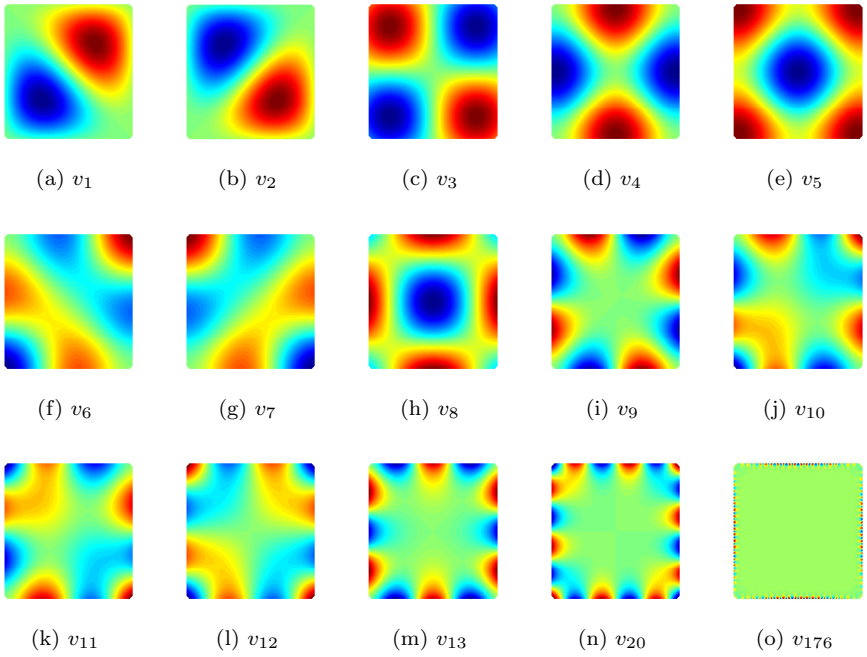
be the mass matrix on the interface of subdomain Ω_i. The discrete formulation of the eigenproblem (9) is [28]: For $1 \leq i \leq N$ find $(\vec{u}, \lambda) \in \mathbb{C}^{n_{\Gamma_i}} \times \mathbb{C}$, s.t.

$$\left(A^{(i)}_{\Gamma_i \Gamma_i} - A_{\Gamma_i I} A_{II}^{-1} A_{I\Gamma_i}\right) \vec{u} = \lambda M_{\Gamma_i} \vec{u}. \tag{11}$$

Now we define the extension operator required in Line 4 of Algorithm 4.1:

Definition 4.4 (Discrete Helmholtz extension)**.** The extension of a vector $\vec{g} \in \mathbb{C}^{n_{\Gamma_i}}$ defined on the interface Γ_i to all degrees of freedom on subdomain Ω_i is the vector $\vec{u} \in \mathbb{C}^{n_i}$ given by $\vec{u} = \left(-A_{II}^{-1} A_{I\Gamma_i} \vec{g}, \quad \vec{g}\right)^T$.

Remark 4.5 (Singular extension). The extensions in definitions 4.1 and 4.4 might give a (numerically) singular problem for subdomains that do not touch the Robin boundary. As those problems are small and are solved directly, strategies for singular systems such as QR decomposition can be employed.



(a) v_1 (b) v_2 (c) v_3 (d) v_4 (e) v_5

(f) v_6 (g) v_7 (h) v_8 (i) v_9 (j) v_{10}

(k) v_{11} (l) v_{12} (m) v_{13} (n) v_{20} (o) v_{176}

Figure 3: Some DtN eigenfunctions. 5×5 subdomains, $n_{\text{loc}} = 40$, $k = 30$.

4.3. How to choose the Dirichlet-to-Neumann coarse space functions

For indefinite systems as those arising from the Helmholtz equation, in contrast to the symmetric positive definite case, increasing the dimension of the coarse space might lead to a deterioration of the convergence rates, c.f. Sec. 5.1 and [30]. An incorrect coarse space might hence be fatal, but it is difficult to decide *a priori* which and how many modes are needed. Thus an appropriate strategy for the coarse space construction is extremely important. In this section, we justify and investigate the coarse space based on the DtN operator introduced in Sec. 4.2 in detail and test Criterion 4.3.

For the tests, we take the setup described in Sec. 6.1 and Problem 1 from Sec. 6.2. Let the domain Ω be decomposed into 5×5 subdomains, $n_{\text{loc}} = 40$, $k = 30$. All experiments in this section are based on this example for the ease of presentation. The conclusions have however also been verified in modified setups. Moreover, they are supported by the numerical experiments in Sec. 6.

We first examine *which* eigenfunctions of the DtN eigenproblem (11) are important. The $n = 176$ eigenvalues on the central subdomain satisfy $\text{Re}\,(\lambda_i) \leq \text{Re}\,(\lambda_j)$ if $i \leq j$, $\text{Re}\,(\lambda_5) < 0 < \text{Re}\,(\lambda_6)$, and $\text{Re}\,(\lambda_{12}) < k < \text{Re}\,(\lambda_{13})$. We show a few of the extensions to the interior of the subdomains of the associated eigenvectors v_i in Fig. 3. From this, our first guess is that eigenfunctions associated to smaller eigenvalues are more useful. We check this numerically by

Choice	# iterations	
	$m_i = 12$	$m_i = 24$
no coarse space	115	115
$\mathrm{Re}(\lambda)$ minimal	16	10
$\lvert\lambda\rvert$ minimal	37	26
$\lvert\lambda - k\rvert$ minimal	77	35
$\lvert\lambda\rvert$ maximal	115	115

Table 1: Iteration numbers for different choices of DtN eigenfunctions.

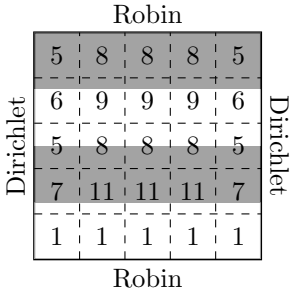


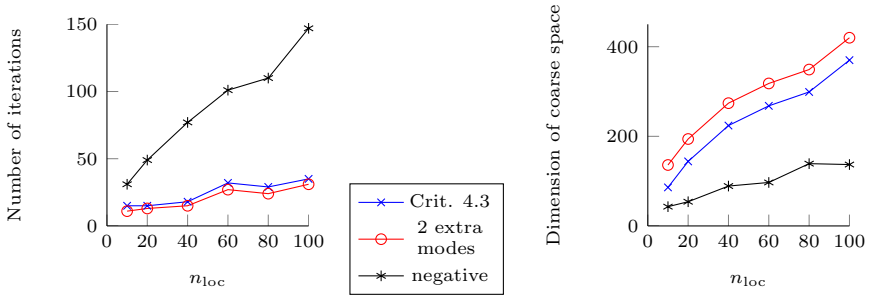
Figure 4: Number of DtN modes per subdomain. Heterogeneous Problem 1, 5×5 subdomains, $n_{\mathrm{loc}} = 40$, $k = 30$.

comparing coarse spaces with 12 modes per subdomain based on the eigenvalues with the smallest real part, the smallest eigenvalues in modulus, the eigenvalues closest to the wave number $k = 30$, and the eigenvalues with the largest modulus. This yields the results in the central columns of Tab. 1. The first alternative, which is in accordance with Criterion 4.3, gives the best results.

To ensure that our findings are not distorted by choosing a too small coarse space, in the last column of Tab. 1 the results for the same experiment with a twice as large coarse space are reported. The best strategy seems robust in terms of the number of iterations; it hardly changes compared to the case with smaller coarse space. Also in this setting, Criterion 4.3 performs best.

In the next step, we examine *how many* modes should be chosen. The more important part of this problem is that we should not choose too few modes as the convergence rates cannot be expected to depend monotonically on the coarse space size due to the indefiniteness of the system. Nevertheless, choosing too many modes increases the computational costs and is not desirable either. The number of modes is controlled by Criterion 4.3. In Fig. 4 we show the resulting number of modes per subdomain for a heterogeneous example. They are influenced both by the boundary conditions and the heterogeneity.

In Fig. 5, we examine whether the number of modes resulting from Criterion 4.3 is sensible. It yields convergence rates that are almost independent of grid width/wave number at the cost of an increasing coarse space size. We investigate whether we can do significantly better by adding the next two eigenvectors on each subdomain – where the eigenvalues are ordered by their real parts – to the coarse space. Fig. 5a shows that this is not the case, it only yields a slight improvement. Moreover, we test whether we could achieve the same behavior with a significantly smaller coarse space. Therefore, we choose another "natural" bound, taking only eigenvectors that are associated to eigenvalues with real part smaller than 0, denoted by "negative" in the legends. For this choice, the number of iterations significantly increases with n_{loc} and hence with k.



(a) Number of iterations (b) Dimension of the coarse space

Figure 5: Varying the coarse space dimension. $k^3h^2 \approx \frac{2\pi}{10}$, Problem 1, 5×5 subdomains.

Additionally, we test whether Criterion 4.3 is independent of the diameters of the subdomains. For that purpose, we take a square $\Omega = [0, L]^2$ decomposed into 5×5 subdomains with $n_{\text{loc}} = 40$. For $L = 1, 5, 10$ and k such that $kL = 30$ is constant, the DtN shows exactly the same behavior: The number of modes that are chosen (224 in total) and the number of iterations (24) do not change with L. Criterion 4.3 consequently provides a useful strategy.

5. Sensitivity of the correction to the coarse space

In order to design and test a coarse space for the Helmholtz equation, it is indispensable to understand how it influences the eigenvalues of the preconditioned operator and hence the convergence behavior of the iterative method. For symmetric positive definite (spd) matrices, this question has been examined extensively e.g. in [39, 40]. In particular, the coarse space does not make the effective condition number of the preconditioned matrix worse for any choice of Z [40, Theorem 2.1]. In this section, we examine to what extent these results apply to indefinite systems. We will see that contrarily to the spd case, for indefinite matrices there seems to be no way to ensure that using a two-level method with an *arbitrarily* chosen coarse space always accelerates the convergence. This is why choosing the right, problem dependent coarse space is important for indefinite systems as those arising from the Helmholtz equation.

5.1. Influence of the Dirichlet-to-Neumann coarse space on the spectrum

We compare the convergence rates and the spectrum of the two-level method (7) with DtN coarse space to those of the corresponding one-level RAS preconditioner (6). Apart from providing a more detailed understanding of our setting, these experiments can also be seen as a demonstration of the challenges that arise when designing a coarse space for an indefinite problem: Neither convergence rates nor the spectrum necessarily ameliorate when adding the second level even if it is carefully designed, c.f. [30].

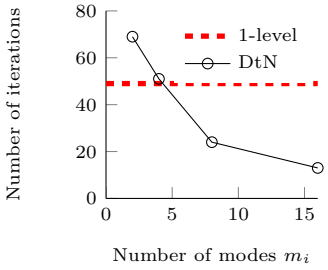
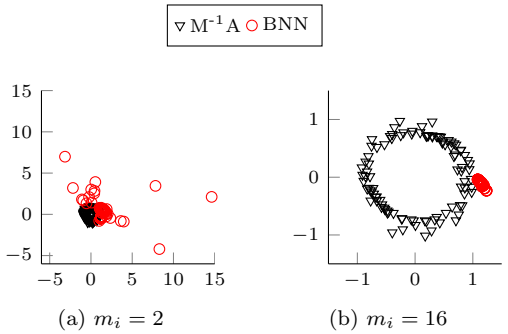


Figure 6: Number of iterations in dependence of m_i. Problem 2, 5×5 subdomains, $n_{\text{loc}} = 40$, $k = 29.3$.



(a) $m_i = 2$ (b) $m_i = 16$

Figure 7: 100 largest eigenvalues for $M^{-1}A$ and BNN in the complex plane. Problem 1, 5×5 subdomains, $n_{\text{loc}} = 40$, $k = 30$.

Firstly, in Fig. 6, we compare the convergence rates of the one- and two-level methods for Problem 2 from Sec. 6.2, using the setup described in Sec. 6.1. Note that in contrast to most of the other experiments, we do not choose Problem 1 to make the problem simpler to solve with the one-level method (6). While using too few coarse space modes gives worse convergence rates than those of the one-level method, employing enough modes resolves the problem.

The next step is to look at the spectrum of the two operators for Problem 1 from Sec. 6.2. We first motivate why it is reasonable to do this in the case of the GMRES method, which does not depend directly on the condition number as CG for spd matrices: The clustering of the eigenvalues is important for the convergence of the GMRES method [41]. As the eigenvalues for RAS all lie within a circle centered at the origin, we compare the largest eigenvalues without and with coarse space. If it decreases for the BNN preconditioner, the clustering is likely to be better. On the other hand, if it increases significantly, this is probably an indication for deterioration.

In Fig. 7, we compare the eigenvalue distribution of the one- and the two-level methods; the largest eigenvalues both of $M^{-1}A$ and the BNN preconditioner are plotted in the complex plane. If the coarse space dimension is small, there is no clear structure in the eigenvalue distribution. While they lie within a circle of radius less than 1 with center $(0, 0)$ for $M^{-1}A$, adding the coarse space with only a few modes has a chaotic effect, scattering the eigenvalues in the complex plane. This changes when adding more modes; the eigenvalues are then clustered near the point $(1, 0)$.

Consequently, choosing an incomplete or incorrect coarse space may have a detrimental effect both on convergence rates and on the spectrum. The coarse space has hence to be designed with the greatest care.

5.2. Understanding the effect of the coarse space for Hermitian matrices

In Sec. 5.1, we have shown that adding a second level to a preconditioner for an indefinite matrix may negatively affect both the convergence rates and

the spectrum of the preconditioned operator. Here we give an explanation of this behavior and identify its causes. This is an extension of the results in [40] to indefinite systems. To be able to provide results, we restrict to a simpler setting: We assume that the matrix A is Hermitian which holds for the matrix associated to the Helmholtz problem defined in (5) if $\Gamma_R = \emptyset$. We examine deflation instead of the balancing preconditioner:

Definition 5.1 (Deflation operator)**.** Let $A \in \mathbb{C}^{n\times n}$ and let $Z \in \mathbb{C}^{n\times r}$, $r < n$. If Z^*AZ, $* \in \{T, \dagger\}$, where T denotes the transpose and $\dagger$ the conjugate transpose, is regular, we define the deflation operator

$$P_D = I - A\Xi, \quad \text{where } \Xi = Z\left(Z^*AZ\right)^{-1} Z^*.$$

The deflation operator is closely related to the balancing preconditioner: If $A = A^*$, then $P_{\mathrm{BNN}}A$ and P_DA have the same eigenvalues except for those that are 0 and 1, respectively [39, Theorem 2.8]. In this section, we use Def. 5.1 with $* = \dagger$ the conjugate transpose.

Let v_i, $1 \le i \le n$ be an orthonormal basis of eigenvectors of A with corresponding eigenvalues $\lambda_i \in \mathbb{R}$. Let $|\lambda_i| \le |\lambda_{i+1}|$ for all $1 \le i < n$. Let $\lambda_{\min} = \min_{i:|\lambda_i|\neq 0} |\lambda_i|$, $\lambda_{\max} = \max_i |\lambda_i|$. We may consider the columns of Z separately in a recursive procedure, using a variant of [42, Theorem 3.2]:

Theorem 5.2. *Let* $P^{(k)} = I - AZ_k\left(Z_k^TAZ_k\right)^{-1}Z_k^T$ *with* $Z_k = \left[\tilde{Z}_1, \tilde{Z}_2, \ldots, \tilde{Z}_k\right]$, *where* $\tilde{Z}_j \in \mathbb{R}^{n\times l_j}$ *has full rank* l_j. *Let* $\tilde{Z}_i^T\tilde{A}_{i-1}\tilde{Z}_i$ *and* $Z_i^TAZ_i$ *be nonsingular for all* $1 \le i \le k$. *Then* $P^{(k)}A = P_kP_{k-1}\ldots P_1A$, *where* $P_{i+1} = I - \tilde{A}_i\tilde{Z}_{i+1}\left(\tilde{Z}_{i+1}^T\tilde{A}_i\tilde{Z}_{i+1}\right)^{-1}\tilde{Z}_{i+1}^T$, $\tilde{A}_i = P_i\tilde{A}_{i-1}$, $\tilde{A}_0 = A$.

The proof is literally the same for our situation. Hence we may restrict to $Z \in \mathbb{C}^{n\times 1}$ with only one column, $Z := \sum_{i\in I}\alpha_iv_i$, where $\alpha_i \in \mathbb{C}$ are coefficients and $I \subseteq \{1, \ldots, n\}$ By computing P_DAv_k for all k, we get the following lemma:

Lemma 5.3 (Structure of P_DA)**.** *W.l.o.g. assume* $I = \{1, 2, \ldots, |I|\}$. P_DA *in the basis of eigenvectors* $(v_i)_{1\le i\le n}$ *is a block diagonal matrix with the two blocks* B *and* D, *where* $B \in \mathbb{C}^{|I|\times|I|}$ *is the block associated to* $(v_i)_{i\in I}$ *and is defined by*

$$B_{ii} = \frac{\sum_{k\neq i}|\alpha_k|^2\lambda_i\lambda_k}{\sum_{k\in I}|\alpha_k|^2\lambda_k}, \quad B_{ij} = -\frac{\alpha_j\bar{\alpha}_i\lambda_j\lambda_i}{\sum_{k\in I}|\alpha_k|^2\lambda_k}, \qquad \forall i, j \in I, i \neq j,$$

and D *is a diagonal matrix with diagonal entries* $\lambda_{|I|+1}, \ldots, \lambda_n$.

The following theorem treats the simple cases in which bounds on the eigenvalues of the deflated operator can be guaranteed.

Theorem 5.4. *If all* λ_i, $i \in I$, *have the same sign, then*

$$\lambda_{\max}(P_DA) \le \lambda_{\max}(A) \quad \textit{and} \quad \lambda_{\min}(P_DA) \ge \lambda_{\min}(A).$$

Proof. Let V be the matrix with columns the eigenvectors v_i, $1 \leq i \leq n$. According to Lemma 5.3, after reordering, $V^\dagger P_D A V$ has block structure with a block B associated to $\{v_i\}_{i \in I}$ and a diagonal block D. As the two blocks are decoupled and all eigenvalues of D are eigenvalues of A, we can consider only B. As by assumption, all λ_i, $i \in I$, have the same sign and eigenvalues are invariant under change of basis, either B or $-B$ is Hermitian positive definite and we can use the result for the real, spd case [40, Theorem 2.1], whose proof is literally the same for complex matrices, to prove the claim. □

Theorem 5.4 shows that if all eigenvalues associated to the eigenvectors that contribute to the vector Z have the same sign, the spectrum of the deflated operator can be bounded by the one of the original operator. Deterioration of the spectrum could thus be avoided if an orthonormal basis of eigenvectors of the global operator was known. This is not feasible in practice.

The remaining question is what happens to the eigenvalues if the λ_i, $i \in I$, have different signs. For simplicity, we restrict to the case $|I| = 2$ and show that different signs of the eigenvalues do cause problems.

Theorem 5.5. *Let $|I| = 2$, i.e. $Z = \alpha_i v_i + \alpha_j v_j$ for some $1 \leq i, j \leq n$. Then we have $\lambda_{\max}(P_D A) > \lambda_{\max}(A)$ if and only if i and j are chosen such that λ_i and λ_j have different signs and*

$$\frac{\left(|\alpha_i|^2 + |\alpha_j|^2\right) |\lambda_i| \, |\lambda_j|}{\left||\alpha_i|^2 \lambda_i + |\alpha_j|^2 \lambda_j\right|} > |\lambda_k| \;\; \forall \; 1 \leq k \leq n. \tag{12}$$

Proof. This follows directly from Theorem 5.4 and Lemma 5.3, observing that the matrix B has the eigenvalues 0 and $\frac{\left(|\alpha_i|^2 + |\alpha_j|^2\right)\lambda_i \lambda_j}{|\alpha_i|^2 \lambda_i + |\alpha_j|^2 \lambda_j}$. □

Theorem 5.5 provides an explanation for the scattering of the eigenvalues observed in Fig. 7: If eigenvectors associated to eigenvalues with different signs enter the coarse space, the eigenvectors of the deflated matrix might grow arbitrarily large. However, the results are not able to explain the clustering of the eigenvalues that occurs when the coarse space dimension gets larger.

5.3. A modified deflation operator

In this section, we examine a variant of the deflation operator used e.g. in [23, 27], where $* = T$ in Def. 5.1. Even though the transpose T seems to be more suitable for complex symmetric matrices than the conjugate transpose $\dagger$, as it is closely related to the structure of the matrix, we show that the situation with this choice is even worse.

We consider a complex symmetric, possibly non-Hermitian matrix A as it arises from the discretization of the Helmholtz equation (2). We assume that A is diagonalizable. It hence has an eigenvector matrix V such that $V^T A V$

is diagonal and $V^TV = I$ [43, Theorem 4.4.13]. Under these assumptions, a modified version of Theorem 5.5 holds, where Condition (12) is substituted by

$$\left| \frac{\left(\alpha_i^2 + \alpha_j^2\right) \lambda_i \lambda_j}{\alpha_i^2 \lambda_i + \alpha_j^2 \lambda_j} \right| > |\lambda_k| \quad \forall 1 \leq k \leq n.$$

Since the coefficients α_i are complex numbers, the denominator $\left|\alpha_i^2 \lambda_i + \alpha_j^2 \lambda_j\right|$ might become arbitrarily large independently of the signs of the eigenvalues. Hence, in contrast to the Hermitian case, no sign restriction on λ_i and λ_j can prevent the eigenvalues from becoming huge:

Example 5.6. Let $A = \begin{pmatrix} 1 & \imath & 0 \\ \imath & 1 & 2\imath \\ 0 & 2\imath & 1 \end{pmatrix}$ with (orthogonal) eigenvectors $v_1 = \begin{pmatrix} 1 & \sqrt{5} & 2 \end{pmatrix}^T$, $v_2 = \begin{pmatrix} 2 & 0 & -1 \end{pmatrix}^T$, $v_3 = \begin{pmatrix} 1 & -\sqrt{5} & 2 \end{pmatrix}^T$ and eigenvalues $\lambda_1 = 1 + \imath\sqrt{5}$, $\lambda_2 = 1$, $\lambda_3 = 1 - \imath\sqrt{5}$. Choosing $Z = \alpha \frac{1}{\|v_1\|} v_1 + \frac{1}{\|v_2\|} v_2$, $\alpha = \sqrt{\frac{\varepsilon - \lambda_2}{\lambda_1}}$ for some real number $0 < \varepsilon < 1$, we get

$$\lambda_{\max}(P_D A) = \left|1 + \imath\varepsilon^{-1}\sqrt{5}\right| > \lambda_{\max}(A) = \left|1 + \imath\sqrt{5}\right|.$$

The theoretical results are clearly in favor of using the conjugate transpose. As the setting has been simplified in order to allow an easy investigation, we compare the matrices Ξ defined in Def. 5.1 for $* \in \{T, \dagger\}$ used in the BNN method (7) with the DtN coarse space for the setting in Tab. 2 with 5×5 subdomains. While the results for $* = \dagger$ are shown in this table, the number of iterations for $* = T$ are not reported because they are almost exactly the same and differ at most by 1. There hence seems to be almost no difference between the two operators. This result is different from [27, Sec. 4.1.2], where the authors conclude that the conjugate transpose outperforms the transpose, using however a different framework and example. We here choose the method using the conjugate transpose as it seems to be more robust. However, in contrast to the spd case, for none of the approaches an arbitrary choice of the coarse space W guarantees a gain compared to the one-level method, neither in terms of convergence rates nor in terms of bounds on the eigenvalues.

6. Numerical results

In this section, we examine the DtN coarse space numerically and compare it to the standard coarse space based on plane waves, c.f. Sec. 1.

6.1. Framework and implementational details

We describe framework in which the preconditioner (7) is used and give implementational details for the numerical experiments.

As a solver we use a generalized minimal residual (GMRES) method without restart with the BNN method (7) as a preconditioner. The termination criterion

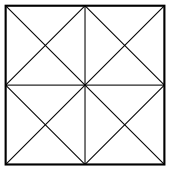


Figure 8: Mesh.

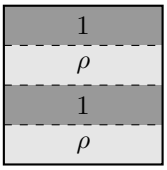


Figure 9: Problem 1.

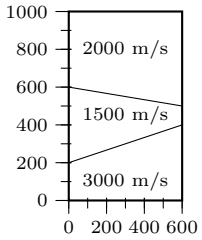


Figure 10: Problem 3.

is based on the error $\|u_h - u_i\|_\infty$, where u_h is the exact finite element solution and u_i is the iterative solution in step k. The system is considered to be solved in step i if $\frac{\|u_h - u_i\|_\infty}{\|u_h\|_\infty} < 10^{-7}$. A bar "–" is used in the tables, when the maximum number of iterations, here 400, is reached. The initial iterate has pseudorandom values drawn from the standard uniform distribution on the interval $(0, 1)$. The finite element part is implemented in FreeFem++ version 3.21 [44], the algebraic part in MATLAB version 7.10.0.499 (R2010a).

Due to the wave character of the solution, in all numerical experiments the grid has to be sufficiently fine in order for the discrete solution to be a good approximation of the continuous one. Additionally to the requirement of having a minimum number of points per wavelength, in order to avoid the *pollution effect* [2], not only kh, but also k^3h^2 needs to be bounded from above. Various modifications, see e.g. [2] and references therein, of the finite element method are known to reduce this effect, but not considered here.

We choose an overlap L of two mesh elements as defined in Definition 3.1. If nothing else is specified a decomposition into $N = n_S \times n_S$ squares is chosen, let n_i be the number of grid points on one side of one square subdomain. If $n_i = n_j$ for all $1 \le i, j \le N$, we define $n_{\text{loc}} := n_i$. In this case, the mesh is always of the type shown in Fig. 8. For decompositions using Metis [45], an arbitrary triangulation is chosen. We denote the number of grid points on one side of a square domain Ω by n_{glob}.

6.2. Model problems

Here we define the model problems that are used in the numerical experiments in sections 4, 5, and 6. They are all based on the Helmholtz equation (2). The first example [37] is the one that we investigate in most detail.

Problem 1 (Open cavity problem)**.** In Equation (2), let $\Omega := [0, 1]^2$, $\Gamma_D := \{0, 1\} \times [0, 1]$, and $\Gamma_R := [0, 1] \times \{0, 1\}$. The right-hand side f is a point source at $(0.5, 0.5)$. The wave number $k(\vec{x}) = \omega/c(\vec{x})$ is either constant, or the wave speed c is piecewise constant according to Fig. 9, where $\rho \in \mathbb{R}$, $\rho > 1$.

Problem 2 (Free space problem)**.** In Equation (2), let $\Omega := [0, 1]^2$, $\Gamma_R := \partial\Omega$, i.e. we simulate an unbounded region. The right-hand side f is a point source in the center $(0.5, 0.5)$ of the domain Ω. The wave number k is constant.

Problem 3 (Wedge problem)**.** This example mimics three layers with a simple heterogeneity. Let $\Omega = (0, 600) \times (0, 1000)\ m^2$ and $\Gamma_R = \partial\Omega$ in Equation (2). The right hand-side f is a point source located at $(300, 1000)$. The wave number is given by $k(\vec{x}) = \omega/c(\vec{x})$, where c is defined in Fig. 10.

6.3. Plane waves

We here introduce and shortly examine the coarse space based on plane waves. As there is no consensus in the literature on how to incorporate the plane wave based coarse space into a DDM, we choose one possibly non-optimal way. A plane wave $p^\theta : D \subset \mathbb{R}^d \to \mathbb{C}$ in direction $\vec{\theta}$ is a function of the form

$$p^\theta(\vec{x}) = \mathrm{e}^{\imath\bar{k}\theta\cdot\vec{x}}, \quad \vec{x} \in D,\ \vec{\theta} \in \mathbb{R}^d,\ \|\theta\|_2 = 1. \tag{13}$$

Here $\bar{k}$ is the mean value of the (possibly heterogeneous) wave number k on D. We decide to proceed as in the case of the DtN operator, and modify Algorithm 4.1. We ignore Lines 1 and 2 and specify the functions $\vec{u}_i^k$ in Line 4 directly. Note that in contrast to DtN, m_i is chosen *a priori*. If nothing else is specified, we use $m_i = 25$ modes per subdomain.

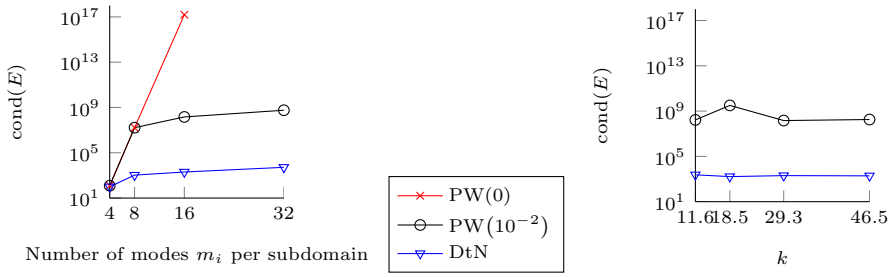
Definition 6.1 (Plane wave coarse space)**.** Let $1 \le i \le N$. For each $1 \le k \le m_i$ choose a direction $\theta_k \in \mathbb{R}^d$, $\|\theta_k\|_2 = 1$ and let $\vec{g}_i^k \in \mathbb{C}^{n_{\Gamma_i}}$ be the coefficient vector of the finite element approximation of p^{θ_k} defined in Equation (13) on the interface Γ_i. Let $\vec{u}_i^k$ be the extension of $\vec{g}_i^k$ in the sense of Def. 4.4.

It remains to be specified how the directions θ_k are chosen. As in [24], we define them via a uniform discretization of the unit circle into circular sectors:

$$\theta_k := \begin{pmatrix}\cos(t_k)\\ \sin(t_k)\end{pmatrix}, \quad \text{where } t_k = \frac{2\pi(k-1)}{m_i}, \quad 1 \le k \le m_i.$$

The matrix Z based on plane waves can become rank deficient for a couple of reasons [24]. The rank deficiency of Z causes in the worst case divergence of the whole iterative scheme. To avoid this problem, we adapt the filtering of the coarse space described in [24], but apply filtering to functions defined on the entire subdomains instead of only the edges. We choose a filtering tolerance ϵ and do the following: Let Z have the blocks W_i. For each $1 \le i \le N$, perform the QR factorization of W_i, and then construct W_i^* as the union of the columns q_j of W_i for which $|R_{jj}| > \epsilon$. This is a local procedure that is performed on each subdomain separately. We substitute Z by the matrix constructed from the W_i^*. A too small value of ϵ can cause the matrix Z to be still rank deficient. The authors of [24] propose to choose ϵ rather too large than too small, setting $\epsilon = 10^{-2}$. We denote the method where the plane wave coarse space with filtering tolerance ϵ is employed by PW(ϵ).

As an adaptive strategy for choosing the coarse space size is crucial, we here examine to what extent it is provided by the filtering procedure. Therefore, we consider Problem 1 with $n_{\mathrm{loc}} = 40$, $k = 29.3$ and a decomposition into 5×5 subdomains. The dimension of the coarse space depends strongly on



(a) $n_{\text{loc}} = 40$, $k = 29.3$.

(b) $m_i = 16$, $k^3h^2 \approx \frac{2\pi}{10}$.

Figure 11: Condition number of coarse matrix E. Problem 1, 5×5 subdomains.

the number of modes per subdomain m_i that are initially chosen, even if the additional modes hardly influence the convergence rate: If we choose $m_i = 16$, the coarse space dimension is 384; with $m_i = 32$ it is 459. However, the number of iterations is 13 versus 11 and hence did hardly change despite the larger, more expensive coarse space. Consequently, even though filtering provides some sort of adaptivity, it is sensitive to the number of modes that are initially chosen.

6.4. Conditioning of coarse matrix

The condition number of the coarse matrix $E = Z^\dagger AZ$ plays an important role. If it is too large, the iterative method might stagnate. Here we investigate to what extent the DtN coarse space suffers from conditioning problems.

The matrix Z is constructed from an orthonormal basis of eigenvectors defined on the interfaces of the subdomains. Their extensions to the interior of the subdomains are in general not orthogonal as the extension matrix $A_{II}^{-1}A_{I\Gamma}$ is not unitary and hence does not conserve orthogonality. Nevertheless, in the numerical experiments for Problem 1 the condition number for DtN behaves well: In Fig. 11a, we examine the dependence of the condition number on the coarse space size and compare it to the one for plane waves, with and without filtering. For the DtN, the condition number of E is only mildly affected by the coarse space dimension. The same is true for plane waves with filtering, but here the upper bound is significantly larger. In Fig. 11b, we investigate whether this behavior carries over to a broader range of wave numbers k. We choose a fixed number $m_i = 16$ of modes per subdomain. If k and h are varied such that k^3h^2 is constant, the condition number for DtN remains almost constant. So filtering of the coarse modes is not necessary here. Furthermore, the condition number is significantly lower than for the plane waves.

6.5. Performance for homogeneous open cavity problem

We study the performance of the DtN coarse space for Problem 1 with homogeneous wave number. In Tab. 2, the number of iterations for different k

		5×5 subdomains					10×10 subdomains				
n_{glob}	k	1-lev	DtN		PW(10^{-2})		1-lev	DtN		PW(10^{-2})	
100	18.5	80	15	(144)	8	(352)	144	18	(344)	7	(1152)
200	29.3	116	18	(224)	11	(467)	241	26	(460)	9	(1286)
400	46.5	156	29	(299)	–	(577)	327	51	(624)	13	(1708)
800	73.8	217	39	(508)	24	(609)	–	65	(936)	–	(2346)

Table 2: Number of iterations (dimension of coarse space) for Problem 1. Comparison of RAS method (6) without coarse space (1-lev), and with DtN and PW(10^{-2}) coarse spaces.

n_{loc}	k	m_i	DtN	PW
10	11.6	4	14	17
20	18.5	6	21	23
40	29.3	9	23	22
80	46.5	12	35	35
160	73.8	21	38	29

(a) Number of modes m_i computed from the DtN coarse space dimension

n_{loc}	k	m_i	DtN	PW
10	11.6	12	8	7
20	18.5	15	9	9
40	29.3	17	13	12
80	46.5	24	18	16
160	73.8	25	36	24

(b) Number of modes m_i computed from the PW coarse space dimension

Table 3: Comparison of number of iterations with identical coarse space size for DtN and PW(10^{-2}). 5×5 subdomains, Problem 1.

is shown. It increases slightly with k if k^3h^2 is constant. This could be due to the decreasing physical size Lh of the overlap, c.f. Tab. 5. Moreover, the dimension of the DtN coarse space depends linearly on the wave number k. The number of iterations for the one-level method doubles if the number of subdomains is doubled in both directions. With the DtN coarse space, the influence of the number of subdomains is not that strong, but still present.

In Tab. 2 we have seen that the number of iterations is smaller with the PW than with the DtN coarse space. However, it is not fair to compare these numbers as the dimensions of the coarse spaces differ significantly. Therefore, in Tab. 3 we compare the two methods enforcing the dimension to be the same by prescribing a fixed number of modes m_i on each subdomain also for DtN. These numbers m_i are the same on all subdomains and are computed by dividing the sizes in Tab. 2 by the number of subdomains. With this setting, DtN and PW then yield approximately the same convergence rates.

In Fig. 12, we examine whether the convergence rates depend on the value of the constant k^3h^2. There is no clear indication which value might be optimal, but a rather fine grid gives the highest number of iterations absolutely. The coarse space dimension depends on k but is independent of the grid width.

In Tab. 4, the mesh width is kept fixed and the wave number is varied. We see how the coarse space dimensions increases with k. The number of iterations

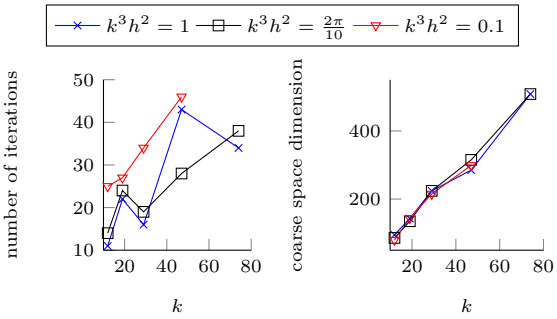


Figure 12: Testing different values of k^3h^2. Problem 1, 5×5 subdomains.

k	1-level	DtN	
5	106	88	(25)
10	115	68	(70)
15	117	61	(90)
30	133	31	(224)
45	169	36	(299)

Table 4: Dependence of number of iterations (coarse space dimension) on wave number for fixed mesh width. Problem 1, 5×5 subdomains, $n_{\text{loc}} = 120$.

	$n_{\text{loc}} = 20$, $L = 2$			$n_{\text{loc}} = 80$, $L = 2$			$n_{\text{loc}} = 80$, $L = 8$		
k	1-level	DtN		1-level	DtN		1-level	DtN	
1	73	56	(25)	94	81	(25)	66	39	(25)
5	64	43	(25)	96	78	(25)	55	37	(25)
10	68	21	(74)	106	49	(74)	66	22	(74)
20	84	32	(139)	107	32	(144)	86	33	(139)

Table 5: Dependence of number of iterations (coarse space dimension) on overlap/mesh width. Problem 1, 5×5 subdomains.

remains only constant if k is large enough. For k small, the coarse space built with Criterion 4.3 is so small that the number of iterations remains rather large.

In Tab. 5, two properties of the DtN coarse space and of Criterion 4.3 become visible: On the one hand, for small k, only one mode per subdomain is chosen and the number of iterations is hardly influenced by the coarse space. This is not a flaw of the coarse space itself, but due to Criterion 4.3; choosing more modes results in a stronger impact on convergence rates. For the homogeneous case this is not a problem as cases with very small wave number k can be solved by standard methods.

On the other hand, we study the influence of mesh refinement. If the mesh is refined twice and the overlap stays constant in terms of number of elements L, see the central columns of Tab. 5, the convergence rates deteriorate a lot; in the worst cases we get more than a factor 2 more iterations with about the same coarse space size. If however Lh, i.e. the physical size of the overlap, is constant, see the last three columns of Tab. 5, the number of iterations even decreases if the mesh is refined. This behavior is probably due to the transmission conditions that make the convergence rates depend on the size of the overlap [37]. Within this context, it might be worth to investigate more advanced transmission conditions, e.g. optimized ones [34].

		Number of subdomains							
n_{loc}	k	5×5		5×10		5×15		5×20	
10	11.6	15	(80)	18	(160)	20	(240)	22	(320)
20	18.5	15	(144)	15	(314)	16	(484)	16	(654)
40	29.3	18	(224)	18	(484)	19	(744)	20	(1004)
80	46.5	29	(299)	37	(624)	43	(949)	48	(1274)

Table 6: Dependence of number of iterations (coarse space dimension) on number of subdomains. DtN coarse space, Problem 1.

		$\rho = 5$				$\rho = 10$			
n_{loc}	ω	DtN		PW(10^{-2})		DtN		PW(10^{-2})	
10	11.6	18	(69)	8	(229)	19	(69)	8	(214)
20	18.5	23	(111)	10	(274)	23	(111)	11	(263)
40	29.3	31	(159)	13	(339)	35	(159)	16	(326)
80	46.5	33	(242)	–	(442)	40	(236)	–	(414)
160	73.8	47	(388)	–	(519)	57	(378)	42	(494)

Table 7: Number of iterations (coarse space dimension). Heterogeneous Problem 1, 5×5 subdomains.

Finally, we vary the number of subdomains in the y direction. For $n \times m$ square subdomains with $n < m$, we let $\Omega = [0, 1] \times [0, \frac{m}{n}]$. Results are given in Tab. 6. When the number of subdomains in one direction increases, the coarse space size grows approximately proportionally to it. This makes sense as our mode selection criterion is purely local. The slight differences stem from subdomains touching the boundary. Moreover, the number of iterations remains almost constant when the number of subdomains is increased.

6.6. Performance for heterogeneous open cavity problem

In this section, we study some small heterogeneous test cases for Problem 1. In Tab. 7, the iteration numbers for constant k^3h^2 are shown. For PW, for some cases convergence stagnated due to ill-conditioning despite the rather large filtering tolerance. Moreover, the adaptively chosen coarse space size for DtN is significantly smaller than that for PW. This also has a small effect on the convergence rates, with PW performing better. As in the homogeneous case, the coarse space size increases with the wave number.

In Tab. 8, we vary the contrast $\rho := k_{\max}/k_{\min}$. With increasing contrast, the convergence rates for the one-level method deteriorate. For DtN, even though the coarse space size decreases, the number of iterations grows only slightly. Only for larger contrast, the situation deteriorates. In the parts

ρ	1-level	DtN		PW(10^{-2})		PW(10^{-1})	
10^0	156	29	(299)	43	(577)	16	(505)
10^1	154	40	(236)	–	(414)	26	(346)
10^2	173	52	(236)	–	(388)	33	(320)
10^3	177	53	(236)	–	(379)	35	(315)

Table 8: Varying contrast. Number of iterations (coarse space dimension) for heterogeneous Problem 1, 5×5 subdomains, $n_{\text{loc}} = 80$, $\omega = 46.5$.

n_{loc}	ω	m_i	DtN	PW(10^{-2})	
10	11.6	3	18	20	(75)
20	18.5	5	20	24	(123)
40	29.3	7	31	40	(171)
80	46.5	10	38	55	(237)
160	73.8	16	57	89	(356)

Table 9: Fixed coarse space size. Number of iterations (coarse space dimension) for heterogeneous Problem 1, $\rho = 10$, 5×5 subdomains.

n_{glob}	k	1-level	DtN	
50	11.6	64	14	(116)
100	18.5	92	15	(168)
200	29.3	130	20	(257)
400	46.5	173	29	(381)
800	73.8	256	36	(645)

Table 10: Number of iterations (coarse space dimension) for homogeneous Problem 1. 25 subdomains with Metis [45].

of the domain where ρ is large, the problem is very close to the Laplacian and hence almost positive definite. As we have seen in Tab. 5, DtN does not work well for such situations since the coarse space is too small to enhance convergence. PW does not suffer from this problem, because the coarse space size is not chosen adaptively. Here, the filtering tolerance for PW has to be larger than 10^{-2} to avoid stagnation of convergence due to ill-conditioning of the matrix E. The convergence only stagnates at a certain error; consequently visibility of this effect depends on the desired accuracy of the iterative solution.

We now choose the same coarse space dimension for both DtN and PW, see Tab. 9, to verify that the better convergence rates for PW are due to the size of the coarse space. In contrast to the homogeneous case in Tab. 3, for the heterogeneous one DtN performs significantly better than PW when the number of modes chosen is the same, in particular for larger wave numbers.

6.7. Extension to other problems

In this section, we consider also the other examples defined in Sec. 6.2 to confirm that our results are valid for a broader range of examples.

Irregular decomposition. In all the previous experiments, we have used a decomposition into square subdomains to ensure reproducibility. Here we show that this is restriction is not necessary for the method to work. In Tab. 10 we consider Problem 1, where both the decomposition done with Metis [45] and the triangulation are now irregular. Compared to the regular case in Tab. 2, the

		5×5 subdomains			10×10 subdomains		
k	n_{glob}	1-level	DtN		1-level	DtN	
18.5	100	43	15	(144)	80	16	(364)
29.3	200	49	18	(224)	102	22	(460)
46.5	400	55	26	(315)	106	43	(660)
73.8	800	60	30	(514)	120	47	(956)

Table 11: Number of iterations (coarse space dimension) for Problem 2.

		15 subdomains			60 subdomains		
ω	n	1-level	DtN		1-level	DtN	
90	150×250	44	14	(267)	82	22	(541)
180	300×500	48	16	(514)	94	23	(1074)
360	600×1000	106	20	(968)	99	25	(2113)

Table 12: Number of iterations (coarse space dimension). Problem 3 decomposed with Metis

method behaves similarly. While the dimension of the coarse space increases slightly, the number of iteration is almost the same.

Free space problem. Here we examine Problem 2, where non-reflecting boundary conditions are imposed on the entire boundary. The iteration numbers for different partitions are reported in Tab. 11. The qualitative behavior is similar to the one observed for Problem 1 in Tab. 2, but the absolute number of iterations is lower, in particular for the one-level method.

Wedge problem. As a last example, we consider the wedge problem, Problem 3. The results are reported in Tab. 12. Also for this case, the 2-level method with the coarse space based on the DtN operator shows a good behavior. To be able to compare with the results for the unit square, note that the number of wave-lengths in the y-direction for the smallest angular frequency $\omega = 90$ corresponds to a wave number k varying between 30 and 60 for the unit square.

7. Conclusions

We have introduced and tested a two-level domain decomposition method for the efficient solution of the heterogeneous Helmholtz equation discretized with piecewise linear finite elements. The important new ingredient is the coarse space. Its construction is based on local eigenproblems involving the Dirichlet-to-Neumann operator and can be performed efficiently in parallel. The resulting method has been tested successfully on 2D model problems and compared to

the standard approach with a coarse space based on plane waves. The extension to 3D problems is straightforward.

The construction of the DtN coarse space inherently respects variations in the wave number, making it possible to treat heterogeneous Helmholtz problems. Moreover, it does not suffer from ill-conditioning as the standard approach based on plane waves, even if many coarse modes are chosen. This is an important property as ill-conditioning can cause the iterative method to stagnate.

Our construction is based on local eigenvalue problems which are more costly than an explicit plane wave coarse space. However, these eigenvalue solves can be performed in parallel and should not affect the scalability of the algorithm. In contrast to the plane wave based coarse space, where several parameters need to be adapted carefully to the problem under consideration, our completely automatic construction refrains from the need of parameter tuning.

Our analysis has shown that linear combinations of eigenvectors associated to eigenvalues with different signs that enter the coarse space cause problems if the coarse space is incomplete, possibly making convergence rates worse than those of the one-level method. Our mode selection criterion ensures that all important modes are present, hence guaranteeing good convergence rates.

References

[1] A. Moiola, E. A. Spence, Is the Helmholtz equation really sign-indefinite?, 2012. Preprint.

[2] I. M. Babuška, S. A. Sauter, Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?, SIAM Rev. 42 (2000) 451–484.

[3] O. G. Ernst, M. J. Gander, Why it is difficult to solve Helmholtz problems with classical iterative methods, in: I. Graham, T. Hou, O. Lakkis, R. Scheichl (Eds.), Numerical Analysis of Multiscale Problems, volume 83 of *Lecture Notes in Computational Science and Engineering*, Springer Berlin Heidelberg, 2012, pp. 325–363.

[4] Y. A. Erlangga, Advances in iterative methods and preconditioners for the Helmholtz equation, Arch. Comput. Methods Eng. 15 (2008) 37–66.

[5] M. Magolu Monga Made, Incomplete factorization-based preconditionings for solving the Helmholtz equation, Internat. J. Numer. Methods Engrg. 50 (2001) 1077–1101.

[6] D. Osei-Kuffuor, Y. Saad, Preconditioning Helmholtz linear systems, Appl. Numer. Math. 60 (2010) 420–431.

[7] M. Bollhöfer, M. J. Grote, O. Schenk, Algebraic multilevel preconditioner for the Helmholtz equation in heterogeneous media, SIAM J. Sci. Comput. 31 (2009) 3781–3805.

[8] M. J. Gander, F. Nataf, AILU for Helmholtz problems: a new preconditioner based on the analytic parabolic factorization, J. Comput. Acoust. 9 (2001) 1499–1506.

[9] B. Engquist, L. Ying, Sweeping preconditioner for the Helmholtz equation: hierarchical matrix representation, Comm. Pure Appl. Math. 64 (2011) 697–735.

[10] A. Bayliss, C. I. Goldstein, E. Turkel, An iterative method for the Helmholtz equation, J. Comput. Phys. 49 (1983) 443–457.

[11] Y. A. Erlangga, C. Vuik, C. W. Oosterlee, On a class of preconditioners for solving the Helmholtz equation, Appl. Numer. Math. 50 (2004) 409–425.

[12] A. Brandt, I. Livshits, Wave-ray multigrid method for standing wave equations, Electron. Trans. Numer. Anal. 6 (1997) 91.

[13] H. C. Elman, O. G. Ernst, D. P. O'Leary, A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations, SIAM J. Sci. Comput. 23 (2002) 1291–1315.

[14] A. Toselli, O. B. Widlund, Domain decomposition methods–algorithms and theory, Springer Verlag, 2005.

[15] M. J. Gander, H. Zhang, Domain decomposition methods for the Helmholtz equation: a numerical investigation, in: Domain Decomposition Methods in Science and Engineering XX, Lecture Notes in Computational Science and Engineering, Springer, San Diego, 2012.

[16] B. Després, Méthodes de décomposition de domaine pour les problèmes de propagation d'ondes en régime harmonique, Ph.D. thesis, Université de Paris IX (Dauphine), Paris, 1991.

[17] A. Toselli, Some results on overlapping Schwarz methods for the Helmholtz equation employing perfectly matched layers, Technical Report 765, Courant Institute of Mathematical Sciences, New York University, New York, 1998.

[18] A. Schädle, L. Zschiedrich, Additive Schwarz method for scattering problems using the PML method at interfaces, Domain Decomposition Methods in Science and Engineering XVI (2007) 205–212.

[19] C. C. Stolk, A rapidly converging domain decomposition method for the Helmholtz equation, J. Comput. Phys. 241 (2013) 240 – 252.

[20] F. Collino, S. Ghanemi, P. Joly, Domain decomposition method for harmonic wave propagation: a general presentation, Comput. Methods Appl. Mech. Engrg. 184 (2000) 171 – 211.

[21] M. J. Gander, F. Magoulès, F. Nataf, Optimized Schwarz methods without overlap for the Helmholtz equation, SIAM J. Sci. Comput. 24 (2002) 38–60.

[22] Y. Boubendir, X. Antoine, C. Geuzaine, A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation, J. Comput. Phys. 231 (2012) 262 – 280.

[23] C. Farhat, A. Macedo, M. Lesoinne, A two-level domain decomposition method for the iterative solution of high frequency exterior Helmholtz problems, Numer. Math. 85 (2000) 283–308.

[24] C. Farhat, P. Avery, R. Tezaur, L. Jing, FETI-DPH: a dual-primal domain decomposition method for acoustic scattering, J. Comput. Acoust. 13 (2005) 499–524.

[25] J.-H. Kimn, M. Sarkis, Restricted overlapping balancing domain decomposition methods and restricted coarse problems for the Helmholtz problem, Comput. Methods Appl. Mech. Engrg. 196 (2007) 1507–1514.

[26] J. Li, X. Tu, Convergence analysis of a balancing domain decomposition method for solving a class of indefinite linear systems, Numer. Linear Algebra Appl. 16 (2009) 745–773.

[27] R. Aubry, S. Dey, R. Löhner, Iterative solution applied to the Helmholtz equation: Complex deflation on unstructured grids, Comput. Methods Appl. Mech. Engrg. (2012).

[28] V. Dolean, F. Nataf, R. Scheichl, N. Spillane, Analysis of a two-level Schwarz method with coarse spaces based on local Dirichlet–to–Neumann maps, Comput. Methods Appl. Math. 12 (2012) 391–414.

[29] F. Nataf, H. Xiang, V. Dolean, N. Spillane, A coarse space construction based on local Dirichlet-to-Neumann maps, SIAM J. Sci. Comput. 33 (2011) 1623–1642.

[30] J. Fish, Y. Qu, Global-basis two-level method for indefinite systems. Part 1: convergence studies, Internat. J. Numer. Methods Engrg. 49 (2000) 439–460.

[31] D. Givoli, High-order local non-reflecting boundary conditions: a review, Wave Motion 39 (2004) 319 – 326.

[32] F. Ihlenburg, Finite element analysis of acoustic scattering, volume 132, Springer, 1998.

[33] X.-C. Cai, M. A. Casarin, F. W. Elliott Jr., O. B. Widlund, Overlapping Schwarz algorithms for solving Helmholtz's equation, Contemp. Math. 218 (1998) 391–399.

[34] M. J. Gander, Optimized Schwarz methods, SIAM J. Numer. Anal. 44 (2006) 699–731.

[35] J. Mandel, Balancing domain decomposition, Communications in Numerical Methods in Engineering 9 (1993) 233–241.

[36] Y. A. Erlangga, R. Nabben, Deflation and balancing preconditioners for Krylov subspace methods applied to nonsymmetric matrices, SIAM J. Matrix Anal. Appl. 30 (2008) 684–699.

[37] M. J. Gander, L. Halpern, F. Magoulès, An optimized Schwarz method with two-sided Robin transmission conditions for the Helmholtz equation, Internat. J. Numer. Methods Fluids 55 (2007) 163–175.

[38] J. Galvis, Y. Efendiev, Domain decomposition preconditioners for multiscale flows in high-contrast media, Multiscale Model. Simul. 8 (2010) 1461–1483.

[39] R. Nabben, C. Vuik, A comparison of deflation and the balancing preconditioner, SIAM J. Sci. Comput. 27 (2006) 1742–1759.

[40] R. Nabben, J. J. Tang, C. Vuik, Deflation acceleration for domain decomposition preconditioners, in: P. H. P. Wesseling, C.W. Oosterlee (Ed.), Proceedings of the 8th European Multigrid Conference September 27-30, (2005) Scheveningen The Hague, The Netherlands, Delft.

[41] Y. Saad, M. H. Schultz, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. and Stat. Comput. 7 (1986) 856–869.

[42] T. B. Jönsthövel, M. B. Van Gijzen, C. Vuik, A. Scarpas, On the use of rigid body modes in the deflated preconditioned conjugate gradient method, SIAM J. Sci. Comput. 35 (2013) B207–B225.

[43] R. A. Horn, C. R. Johnson, Matrix analysis, Cambridge University Press, 1990.

[44] F. Hecht, O. Pironneau, A. Le Hyaric, K. Ohtsuka, FreeFem++, Université Pierre et Marie Curie, 2007. `http://www.freefem.org/ff++/ftp/freefem++doc.pdf`.

[45] G. Karypis, V. Kumar, A fast and high quality multilevel scheme for partitioning irregular graphs, SIAM J. Sci. Comput. 20 (1998) 359–392.